



ASBIRES 2025

Proceedings of the
17th

Applied Science, Business & Industrial Research
Symposium

07th January 2026



Faculty of Applied Sciences
Wayamba University of Sri Lanka
Kuliyapitiya



Faculty of Applied Sciences
Wayamba University of Sri Lanka



Proceedings of the 17th Symposium
On
Applied Science, Business & Industrial Research
07th January 2026

ASBIRES – 2025
Faculty of Applied Sciences
Wayamba University of Sri Lanka
Kuliyapitiya

© Faculty of Applied Sciences,
Wayamba University of Sri Lanka, 2026
Catalog in Publication Data approved by
National Library and Documentation Services Board

Symposium on applied science, business & industrial research (17th: 07th January
2026 Kuliapitiya, Sri Lanka)

Proceedings. - Kuliapitiya: Wayamba University of Sri Lanka, 2025
Vol. 17. - 488p; 36cm.

eISBN 978-624-6674-14-4

eISSN 2989-0683

i. 607 DDC 23

ii. Title

Applied science - Research
Industrial management - Research
Mathematical sciences - Research
Computer and information system - Research

ISBN 978-624-6674-14-4
ISSN 2989-0683

Typesetting, Designing and Printing
at Faculty of Applied Sciences,
Wayamba University of Sri Lanka, 2026

This event is proudly organized by the
Faculty of Applied Sciences, Wayamba University of Sri Lanka

EDITORIAL BOARD

Dr. Erandi Jinādari Wattagama

Editor in Chief

Department of Industrial Management

Members

Prof. Asankha Pallegedara	<i>Department of Industrial Management</i>
Dr. N.A.D.N. Napagoda	<i>Department of Mathematical Sciences</i>
Mr. T. Arudchelvam	<i>Department of Computing and Information Systems</i>
Ms. S.D. Dahanayake	<i>Department of Mathematical Sciences</i>
Ms. W.M.L.N. Wanninayake	<i>Department of Computing and Information Systems</i>

Editorial Assistant

Ms. W.I.S. Wickramarathna	<i>Department of Industrial Management</i>

Language Editing

Department of English Language Teaching

Members

Mr. MKSM. Samaranayake	<i>Department of English Language Teaching</i>
Ms.G.N.Madumali	<i>Department of English Language Teaching</i>
Ms.H. M. H. S. Herath	<i>Department of English Language Teaching</i>
Ms.H.M.I.Nawodya	<i>Department of English Language Teaching</i>
Ms.N.M.P.B Nawarathna	<i>Department of English Language Teaching</i>
Ms.W.P.A.P. Senarathna	<i>Department of English Language Teaching</i>
Ms. L.N.R. Pietersz	<i>Department of English Language Teaching</i>
Ms. A.G.N.K. Rathnayake	<i>Department of English Language Teaching</i>
Ms. J.A.G.S. Jayakody	<i>Department of English Language Teaching</i>
Ms. B.M.T.C. Basnayake	<i>Department of English Language Teaching</i>
Ms. W. M. Y. M. P. Wijekoon	<i>Department of English Language Teaching</i>

PANEL OF REVIEWERS

Prof. A. Pallegedara	Wayamba University of Sri Lanka
Prof. (Mrs.) D.D.M. Jayasundara	University of Kelaniya
Prof. E.A.C.P. Karunarathne	Wayamba University of Sri Lanka
Prof. (Mrs.) G.C. Samaraweera	University of Ruhuna
Prof. H.M.A. Herath	Wayamba University of Sri Lanka
Prof. H.M.S.P. Herath	Wayamba University of Sri Lanka
Prof. (Mrs.) I. Hewapathirana	University of Kelaniya
Prof. Leslie Jayasekara	University of Ruhuna
Prof. M.M.D.R. Deegahawature	Wayamba University of Sri Lanka
Prof. (Mrs.) M.P. Nadira Janadari	University of Kelaniya
Prof. (Mrs.) N.V. Chandrasekara	University of Kelaniya
Prof. R.A.R. Prabodanie	Wayamba University of Sri Lanka
Prof. R.M.K.T. Rathanyaka	Sabaragamuwa University of Sri Lanka
Prof. S.R. Kodituwakku	University of Peradeniya
Prof. V.G.T.N. Vidanagama	Wayamba University of Sri Lanka
Dr. (Mrs.) A.A. Sunethra	University of Colombo
Dr. A.D. Dharmawansa	Wayamba University of Sri Lanka
Dr. (Mrs.) Anuradha Hewarachchi	University of Kelaniya
Dr. (Ms.) Asanka Gunawardana	University of Moratuwa
Dr. Asela K. Kulathunga	University of Exeter, United Kingdom
Dr. (Mrs.) B. Mayurathan	University of Jaffna
Dr. (Mrs.) B. Munasinghe	Wayamba University of Sri Lanka
Dr. B.S. Habaragoda	Wayamba University of Sri Lanka

Dr. (Mrs.) Diana Kumari	Open University of Sri Lanka
Dr. (Mrs.) D.M.K.N. Seneviratna	University of Ruhuna
Dr. (Mrs.) D.M.P.V. Dissanayake	University of Kelaniya
Dr. (Mrs.) E.J.K.P. Nandani	University of Ruhuna
Dr. E.J. Wattegema	Wayamba University of Sri Lanka
Dr. E.Y.A. Charles	University of Jaffna
Dr. Gayani Balasooriya	Ottawa University
Dr. (Mrs.) Hemalika Abeysundara	University of Peradeniya
Dr. Himali Gammanpila	Eastern Kentucky University
Dr. (Mrs.) H.L. Jayetileke	University of Ruhuna
Dr. H.W. Bhashitha Kavinga	University of Kelaniya
Dr. Indu Rasika Churchill	Marano Campus Center, United States
Dr. J.A.B.U. Jayasinghe	University of Moratuwa
Dr. (Mrs.) K.D. Prasangika	University of Ruhuna
Dr. Lahiru Gunsekara	Lincoln University, New Zealand
Dr. (Mrs.) Lakshika S. Nawarathna	University of Peradeniya
Dr. M.G.S. Dilanthi	Wayamba University of Sri Lanka
Dr. M.H.L. Weerasinghe	University of Kelaniya
Dr. Michelle Silva	University of Kelaniya
Dr. (Mrs.) M.W. Amali de Silva	University of Sri Jayewardenepura
Dr. (Mrs.) Navodi M. Hakmanage	University of Kelaniya
Dr. Niroshan Perera	University of Sri Jayewardenepura
Dr. N. Varthan	University of Jaffna
Dr. Pasan Manuranga Edirisinghe	University of Moratuwa

Dr. Roshan Adikari	Southwestern College, United States
Dr. Sachith Abeysundara	University of Peradeniya
Dr. S. Arivalzahan	University of Jaffna
Dr. Eng. Sampath Deegalla	University of Peradeniya
Dr. Shadhi Heenatigala	Antioch College, United States
Dr. S.H.D.S. De Silva	University of Kelaniya
Dr. S. Kirushanth	University of Vavuniya
Dr. S.P.D.S.S.K. Karunarathna	University of Kelaniya
Dr. S.R. Ginige	University of Sri Jayewardenepura
Dr. S. Shriparen	University of Jaffna
Dr. S.V.A. Almeida	University of Kelaniya
Dr. Sameera Viswakula	University of Colombo
Dr. T. Ketheesan	University of Jaffna
Dr. T. Kokul	University of Jaffna
Dr. T.M.M. De Silva	University of Kelaniya
Dr. (Mrs.) Ursula Peiris	Wayamba University of Sri Lanka
Dr. W.D.N.S.M. Thennakoon	Wayamba University of Sri Lanka
Dr. W. S. Sanjeewa	Wayamba University of Sri Lanka
Mr. A.R.M.I. Ariyapperuma	Wayamba University of Sri Lanka
Mr. D.M. Jayasena	Wayamba University of Sri Lanka
Mrs. D.M.T.D. Dissanayake	Wayamba University of Sri Lanka
Mrs. Gayani Ranasinghe	Wayamba University of Sri Lanka
Mr. H.A.D. Priyasad	University of Kelaniya
Ms. Isuri Samarawickrama	SAB Campus, Sri Lanka

Mr. J.R.K.C. Jayakody	Wayamba University of Sri Lanka
Mrs. K.T.G.P. Madhusanka	University of Sri Jayewardenepura
Mrs. P.A.A.U. Jothiranth	Wayamba University of Sri Lanka
Mrs. P.A.B.H. Amarathunga	Wayamba University of Sri Lanka
Mr. Rohana Dissanayake	University of Moratuwa
Mrs. R.P.T.H. Gunasekara	Wayamba University of Sri Lanka
Mrs. R.S.L. Jayarathne	Wayamba University of Sri Lanka
Mr. S. Suthakar	University of Jaffna
Mr. S.S. Suthakaran	University of Vavuniya
Mr. S. Thirukumaran	University of Vavuniya
Mr. T. Arudchelvam	Wayamba University of Sri Lanka
Mr. T. Jeyamugan	University of Vavuniya
Mrs. Viraji T. Vaidyasekera	CINEC Campus, Sri Lanka
Mr. W.C.C. Premarathna	Wayamba University of Sri Lanka
Mrs. W.D.M.B.K. Dissanayake	Wayamba University of Sri Lanka
Mrs. W.M.L.N. Wanninayake	Wayamba University of Sri Lanka

TABLE OF CONTENTS

<i>Editorial Board</i>	<i>i</i>
<i>Panel of Reviewers</i>	<i>ii</i>
<i>Table of Contents</i>	<i>vi</i>

Computing and Information Systems

Evaluating the Feasibility of Automated QA Measurement: A Comparative Study of Manual, Selenium-Based, and AI-Powered Testing Using Testim.io <i>Colombage J.C., Arudchelvam T.</i>	01
Predictive Analytics of Employee Turnover in the Apparel Industry <i>Senarathna H.M.P., Wanninayake W.M.L.N.</i>	07
Intelligent Scheduling for Public Transport: A Constraint-Based Automated Timetable System in Sri Lanka <i>Senarathna H.M.P., Arudchelvam T.</i>	12
Machine Learning-Based System for Personalized Career Guidance for Undergraduates in the Faculty of Applied Sciences <i>Dilshan G.N., Wanninayake W.M.L.N.</i>	19
Enhancing Emotional Well-Being Through Text-Based Emotion Detection <i>Premathilaka R.T.K., Wanninayake W.M.L.N.</i>	24
Privacy-Preserving Sentiment Analysis Using Federated Learning for Customer Feedback Systems <i>Jayathilaka H.A.D.H.Y., Premarathna W.C.C.</i>	31
Smart Alert System with IoT for Enhanced Safety and Response <i>Jayawardhana K.P.R.K., Premarathna W.C.C.</i>	37
Mitigating Vehicle & Pedestrian Accidents at Ungated Train Crossings in Sri Lanka: Acoustic Detection and Warning System <i>Marasinghe M.K.Y., Jayakody J.R.K.C.</i>	42
Evaluating the Effectiveness of Frontend Modernization: A React-Based Solution for Corporate Website Enhancement <i>Premathilaka R.T.K., Jayakody J.R.K.C.</i>	48
Automated Road Safety Enforcement Using Convolutional Neural Networks and Object Detection <i>Herath H.M.L.P., Jayakody J.R.K.C.</i>	54
Automated Classification of Mobile App Reviews Using Zero-Shot NLP Model: A Case Study on MYSLT <i>Jayawardhana K.P.R.K., Vidanagama V.G.T.N.</i>	60
A Comparative Study of the Usability of Different Software Tools in Auditing <i>Karunarathne K.T.T., Munasinghe B.</i>	65
A Contextual Analysis of Process Effectiveness and Industry Alignment in a Software Industry Workplace, Including a Benchmarking System Prototype <i>Fernando N.T.S., Munasinghe B.</i>	72
A Multi-model Deep Learning System for Intelligent Vehicle Recognition and Market Valuation <i>Galagamage R.A., Gunasekara R.P.T.H.</i>	77

Adapting to Remote Working: A Case Study on a Multinational Workplace <i>Sewwandi P.H.A.J., Munasinghe B.</i>	82
AI Fish Identification System <i>Perera S.R.V, Jayakody J.R.K.C.</i>	88
Automated Plastic Waste Classification and Recommendation <i>Malsinghe M.P.P.N., Gunasekara R.P.T.H.</i>	93
Consumers Trends in Vegan & Plant-Based Products: A Reddit-Based Sentiment Analysis <i>Karunarathne K.T.T., Gunasekara R.P.T.H.</i>	98
Enhancing Small Business Management Through Customizable Mobile Applications in Sri Lanka <i>Vidana Pathirana Y.K., Munasinghe B.</i>	104
Evaluating the Effectiveness of Standard Operating Procedures in Website Development Projects: A Mixed-Methods Analysis <i>Amarasekara P.E.M., Arudchelvam T.</i>	109
Factors Influencing Cloud ERP Adoption among SMEs in Developing Countries: A Systematic Review <i>Mariyam F, Wickramarachchi R.</i>	114
Factors Influencing the Transition of Non-IT Graduates into the IT Industry: A Systematic Review <i>Thavarasa Y, Wickramarachchi R.</i>	118
Frontend Development Using Generative AI Tools <i>Perera S.R.V., Gunasekara R.P.T.H.</i>	123
Improving Efficiency in Trim Inspection Process through an Automated Inspection System for Heat Seal Using Image Processing Techniques <i>Rajapaksha W.P.A.G., Jayakody J.R.K.C.</i>	128
Leveraging AI and Generative Adversarial Networks for Synthetic Data Generation to Address Imbalances in Medical Datasets <i>Fernando N.T.S., Vidanagama V.G.T.N.</i>	134
Lightweight Deep Learning-Based Real-Time Wild Elephant Detection System <i>Wickramarathna K.W.B.P., Vidanagama V.G.T.N.</i>	139
NLP-Based Financial Data Extraction from SMS for Intelligent Personal Finance Management <i>Jayathilaka H.A.D.H.Y., Gunasekara R.P.T.H.</i>	144
Real-Time Personalized Advertising Using CNN-Based Demographic Analysis <i>Amarasekara P.E.M., Munasinghe B.</i>	148
Predictive Analysis of First-Attempt Failures in the QCO Process Using Machine Learning <i>Dilshan G.N., Munasinghe B.</i>	153
Real-time Monitoring and Identification of Driver Drowsiness Using Hybrid CNN and EAR Approach <i>Illukkumbura S.P.G.E., Wanninayake W.M.L.N.</i>	158
Real-time Question Detection and Proactive Triggering for Meeting AI Assistant <i>Dewmindi P.L.U., Vidanagama V.G.T.N.</i>	165
Snake Species Identification using Object Detection and Computer Vision <i>Samarasinghe S.A.C.L.S., Vidanagama V.G.T.N.</i>	170

An AI-Integrated Admin Panel with Theta Forecasting for E-Commerce Analytics <i>Rathnayake M.M.G.N.L., Vidanagama V.G.T.N.</i>	175
Usability Improvement Through Customer Feedback: Evidence from the Usability Evaluation of a Pharmacy Management Mobile Application <i>Dapanadurage I.N., Gunasekara R.P.T.H.</i>	180
Web-Based Project Management System for Midsize Research Institutes <i>Sameeha M.R., Arudchelvam T.</i>	184
Wood Type Identification System Using Image Processing and Machine Learning Techniques <i>Dewmindi P.L.U., Jayakody J.R.K.C.</i>	191
Industrial Management	
Analysis of Pollutant Exceedances and Treatment Efficiency towards Sustainable Wastewater Management in Export Processing Zones (Seethawaka & Horana) <i>Kusumsiri G.K.A.D., Prabodanie R.A.R.</i>	196
Modelling Temporal Trends and Forecasting Tourists related Violent Crimes in Sri Lanka: A Data Driven Approach <i>Lakmini K.R.D., Habaragoda B.S.</i>	201
A Comprehensive Analysis of Trends and Patterns in Sri Lankan Tourism: Pre-Post Covid Growth and Visitor Demographics <i>Sameera L.N., Habaragoda B.S.</i>	206
Effect of Social Media Engagement and Trust in Shaping Online Purchase Intention: A Study of Sri Lankan Consumers Using the MOA Framework <i>Karunathilaka I.P.K.A., Wattegama E.J.</i>	212
Exploring Employee Perceptions of the Psychological Work Environment and Creative Performance in a Sri Lankan Private Bank <i>Wickramasinghe W.A.D.L.T., Jothirathne P.A.A.U.</i>	218
Bridging Efficiency Gaps in Hospitality SMEs through a Computerized Payroll Management System: A Case Study <i>Kanchana G.M.S.D., Wattegama E.J.</i>	223
Impact of the Logistic Challenges on enhancing the Operational Performance: A Case Study <i>Rasangi N.C., Wattegama E.J.</i>	229
Impact of Climate Change on the Perceived Income of Street Food Vendors in Sri Lanka <i>Karunarathne B.M.D., Wattegama E.J.</i>	235
Factors influencing Student Engagement in Learning Management System : A Case Study <i>Somarathna W.N.M.C., Dharmawansa A.D.</i>	239
Evaluating the Efficiency of Customs Clearance Procedures in Sri Lanka's Import and Export Industry <i>Krishnarathna B.V.N.R., Jothirathne P.A.A.U.</i>	245
Governance, Risk and Compliance (GRC) Readiness Assessment Factors for Sri Lankan Banking and Finance Sector: A Systematic Literature Review <i>Rathnayaka R.M.S.O., Wickramarachchi A.P.R.</i>	250
A Study on Skill Gaps and Training Needs Among Operational Employees in the Apparel Industry <i>Piyamantha N.J., Pallegedara A.</i>	255

The Impact of Digital Skill Proficiency on Youth Employment Opportunities in Sri Lanka	
<i>Perera I.S.R.G., Pallegedara A.</i>	260
Minimizing Overstocking through Improved Supplier Delivery Management: A Case Study	
<i>Vijayamini K.P.S.M., Jothirathne P.A.A.U.</i>	265
Investigating Factors Influencing Behavioral Intention Toward the Adoption of a New ERP System in the Financial Management Sector of Sri Lanka's Automobile Industry; Systematic Literature Review	
<i>Ramesha V., Oruthotaarachchi C.R.</i>	271
Explainable AI for Voice and Broadband Customer Churn Prediction in the Sri Lankan Telecommunication Sector	
<i>Boralessa K.T., Prabodanie R.A.R.</i>	276
High Labour Turnover in Garment Manufacturing Industry: Case Study based Evidence from the Leading Factories in North Western Province	
<i>Sewwandi M.P., Jayasena D.M.</i>	282
Impact of Packaging Designs on Consumer Purchasing Behavior in the Packaged Food Industry	
<i>Sewwandi M.P., Jothirathne P.A.A.U.</i>	287
Factors Affecting to the Solar Panel Adoption in Sri Lanka	
<i>Bimsara M.U., Jayasena D.M.</i>	292
Workplace Ostracism and Deviant Workplace Behaviour: Systematic Literature Review Based on PRISMA Method	
<i>Weersinghe R.D., Janadari M.P.N.</i>	298
The Impact of AI Adoption on Organizational Preparedness for Workforce Displacement with the Moderating Role of Reskilling Initiatives in the Sri Lankan IT Industry: A Systematic Literature Review	
<i>Wasala W.M.C.J., Janadari M.P.N.</i>	305
Mathematics & Mathematical Modelling	
Ray-Tracing and Wavefront Analysis of Aspheric and Astigmatic Ocular Surfaces: Beyond the Paraxial Approximation	
<i>Caldera H.P.S.M., Karunarathna S.P.D.S.S.K.</i>	310
Investigating Structural Similarity Index Metric Variability In Predicting Visual Acuity	
<i>Sandunika A.M.K.S., Karunarathna S.P.D.S.S.K.</i>	316
Optimizing Academic Timetables Using Graph Coloring & Integer Linear Programming - Case Study	
<i>Peiris T.A.W., Karunananda G.A.C.M.</i>	322
A Case Study on Mathematical Optimization of CTB Bus Scheduling for the Kuliyaipitiya-Colombo Route	
<i>Sathsara W.H.T., Dahanayaka S.D.</i>	328
Mathematical Modeling of CO ₂ Emission Reduction in Sri Lanka: A Pathway to Achieve Climate Targets	
<i>Dissanayaka D.M.D.K., Dharmawardane P.M.N.</i>	333
Optimal Homotopy Analysis Method for Mixed Convection Flow with a Nonuniform Heat Source over a Vertical Nonlinear Stretching Surface	
<i>Attanayake A.M.H.I., Karunarathna W.H.D.T.</i>	339

Environmental Noise and Stability in Aquatic Ecosystems: A Stochastic Differential Equation Approach <i>Maddegoda K.L., De Silva T.M.M.</i>	344
Statistics	
Forecasting Order Volumes for Resource Planning: A Case Study <i>Alweera A.M.I.H., Panahatipola P.M.O.P.</i>	350
Identifying Key Effecting Factors of Employee Service Duration in the ABC Apparel Company, Katunayake Using Lasso Regression <i>Nagasinghe N.K.C.J., Lakshmi M.P.D.S.</i>	354
Predicting Airline Passenger Arrivals to Sri Lanka from Asia, Europe, and the Middle East: A Comparative Study <i>Dias P.P.V.A., Appuhamy P.A.D.A.N.</i>	359
Optimization of Machine Setup Time During Style Changeovers in the ABC Apparel Company Using Linear Programming <i>Goonerathne G.D.N., Lakshmi M.P.D.S.</i>	364
Identifying The Affecting Factors And Predicting Customer Satisfaction In Online Food Delivery Services In North Western Province Using Machine Learning Techniques <i>Goonerathne G.D.N., Lakshmi M.P.D.S.</i>	369
Modeling AUD/LKR Exchange Rate Volatility Using Time Series and Machine Learning Models <i>Jayawardana O.A.M., Karunanada G.A.C.M.</i>	374
Hybrid ARIMA and Machine Learning Models for Daily UAX/LKR Gold Exchange Rates Forecasting <i>Chamodya N.D.S.</i>	380
Analyzing and Forecasting International Labour Migration in Sri Lanka <i>De Mel W.R.U., Aponsu G.M.L.M.</i>	385
Integrating RFM Analysis and K-Means Clustering for Customer Segmentation in an Apparel Equipment Provider in Sri Lanka <i>De Mel W.R.U., Aponsu G.M.L.M.</i>	391
Forecasting Quality Outcomes in Apparel Manufacturing Using Time Series Models: A Case Study Based on Final Inspection Data <i>Sathsara W.H.T., Herath C.H.M.A.U., Dahanayaka S.D.</i>	397
A Machine Learning Approach for Classifying Production Performance: A Case Study <i>Dias P.P.V.A., Appuhamy P.A.D.A.N.</i>	402
Statistical Analysis of Data Quality Challenges Faced by Audit Professionals: A Case Study in the ABC Company <i>Diluxana K., Karunarathna W.H.D.T.</i>	407
A Markov Chain-Based Approach for Modelling and Simulating Monsoonal Water Level Variations in Kalu River Basin <i>Fernando C.N.N.M., Ekanayake E.M.P., Appuhamy P.A.D.A.N.</i>	413
Evaluation and Enhancement of Sales Forecast Accuracy: A Case Study in a Cement Company <i>Weerakoon N.D., Dahanayaka S.D., Herath C.H.M.U.</i>	418
Multi-Class Classification of Financial Market Growth and Interest-Level Trends Using CNN and LSTM <i>Upamali G.I., Napagoda N.A.D.N.</i>	424

Performance Evaluation of CNN and RNN Architectures in Predicting the Stock Market Price Index <i>Kaushalya H.W.S., Napagoda N.A.D.N.</i>	429
Time series analysis and forecasting of Sri Lanka's electricity usage: Statistical vs Machine learning models <i>Shehari K.M.D.N., Francisco G.S.</i>	435
Investigating the Root Causes of High Glove Rejection Rates in the Thumb Crotch Reinforcement (TCR) Machine <i>Shehari K.M.D.N., Karunarathna S.P.D.S.S.K.</i>	440
A Comparative Study of Statistical and Machine Learning Techniques for Forecasting Leptospirosis Incidence in Kurunegala District <i>Karunasekara K.M.C.G., Francisco G.S.</i>	445
Statistical Analysis of Production Loss Reasons in Lean Garment Manufacturing <i>Sandunika A.M.K.S., Karunarathna S.P.D.S.S.K.</i>	450
Unit Profit Forecasting in Foreign Exchange Based on Buying and Selling Rate Margins of a Government Bank in Sri Lanka <i>Padmasiri M.D.P., Aponsu G.M.L.M.</i>	456
Analyzing the Factors Influencing the Water Quality Index of the Kelani River and Predicting Its Future Water Quality Trends <i>Padmasiri M.D.P., Aponsu G.M.L.M.</i>	462
Predicting the Job Completion Time in A Continuous - Shift Polythene Manufacturing Plant <i>Iwanthi M.A.Y., Vithanage G.V.R.K.</i>	468
Data-Driven Cost of Spare Parts in Selected Production Machines in Korean SPA Packaging (Pvt) Ltd <i>Jayakody K.A.N.D., Vithanage G.V.R.K.</i>	474
Factors Associated With Self-Reported Health In University Students: A Case Study <i>Alweera A.M.I.H., Panahatipola P.M.O.P</i>	479



Evaluating the Feasibility of Automated QA Measurement: A Comparative Study of Manual, Selenium-Based, and AI-Powered Testing Using Testim.io

Colombage JC¹, Arudchelvam T²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

jayathicolombage123@gmail.com¹, arul@wyb.ac.lk²

ABSTRACT

Quality assurance is an important concept in the software industry. The quality assurance of the software should be checked by the Quality Assurance (QA) team and confirmed to ensure that the system meets the required specifications and functions correctly. As such, when a new method/technique/software is implemented, the QA team should compare the performance of the new and old systems. QA team want to use one of the Selenium-based automation, and AI-powered Testing. Therefore, the performance of both the above and the manual system is to be compared. This study compares Manual Testing, Selenium-based Automation, and AI-powered Testing with Testim.io in Software Quality Assurance (SQA). It seeks to compare their performance in execution time, repeatability, and step accuracy and also consider qualitative parameters such as human involvement, adaptability, and cost-effectiveness, defect detection and software quality. A mixed-method approach was used, combining quantitative measures with survey-based qualitative observations. Results indicate that AI and automation tools achieve higher efficiency and consistency than manual testing, which remains applicable to complex and exploratory cases. Among the automated tools, Testim.io was more flexible and needed less maintenance work due to its AI-based element recognition. Results highlight that the integration of AI-powered testing with traditional approaches can significantly improve accuracy, efficiency, and overall software quality.

KEYWORDS: AI-Powered Testing, Manual Testing, Quality Assurance, Selenium, Testim.io

1. INTRODUCTION

Software quality assurance is a critical aspect of modern software development since the reliability, performance, and satisfaction of the customer directly rely on it. Manual testing has been the quintessential stronghold of software verification but remains inefficient, redundant, and prone to human errors.

To overcome these limitations, organizations have increasingly adopted automation tools such as Selenium, which can execute repetitive test cases quickly and reliably. While automation improves efficiency and repeatability, it demands technical expertise and ongoing

maintenance (Erol & Senan, 2022; Fatima et al., 2022). Recent advancements in artificial intelligence have led to the development of AI-driven testing tools, such as Testim.io, which offer intelligent element identification, self-healing locators, and reduced maintenance effort (Ali et al., 2024). These tools enhance testing efficiency, adaptability, and accuracy compared to traditional methods. This study compares Manual Testing, Selenium-based Automation, and AI-driven Testing (Testim.io) on both quantitative (execution time, repeatability, step accuracy) and qualitative (human effort, flexibility, cost-effectiveness, defect detection and software quality) analyses. The results

indicate how the union between AI flexibility and automation effectiveness can enhance accuracy, reliability, and overall software quality assurance.

2. LITERATURE REVIEW

Manual testing remains a critical part of software quality assurance, especially for new feature testing, user interface, and complex workflow testing that requires human judgment. Manual testing provides flexibility in identifying usability issues and subtle bugs but is limited by inefficiency in repetitive and high-volume tasks (Badr, 2021; Kumari et al., 2018). Although slower, manual testing is usually preferred in exploratory and ad-hoc testing scenarios where human intuition is of great value (Xia & Luo, 2021). Automation testing using tools such as Selenium has gained popularity due to its ability to perform repetitive tasks quickly and reliably. It significantly improves regression testing, integration testing, and continuous delivery processes by increasing productivity and reducing manual effort (Erol & Senan, 2022; Sari & Olsson, 2020). However, Selenium-based automation requires high script maintenance, technical expertise, and often struggles to handle dynamic user interface elements that frequently change during development cycles (Fatima et al., 2022).

Latest AI-driven testing innovations, i.e., Testim.io, have introduced adaptive automation with self-healing locators, visual learning, and machine learning-driven error handling. Testim.io intelligently identifies UI elements and adapts to application changes, with reduced maintenance effort and increased test reliability (Ali et al., 2024). Studies show that AI-driven automation enhances

repeatability, accuracy, and defect identification and reduces test creation time compared to traditional methods. However, these tools still lag behind in handling complex, domain-specific cases and require extensive up-front configuration and learning.

Overall, recent research highlights that no one has tested a technique that is superior in all situations. Manual testing is versatile, automation is fast, and AI-powered tools like Testim.io are adaptive and scalable. Therefore, a hybrid testing solution that combines manual intuition, automated speed, and AI intelligence can deliver optimum results in modern software quality assurance projects.

3. METHODOLOGY

This research employed a mixed-method approach that integrated both quantitative and qualitative evaluations to compare three software testing approaches: Manual Testing, Selenium-based Automation, and AI-powered Testing using Testim.io. The study aimed to measure and analyze the performance and user perception of each method to determine their effectiveness in software quality assurance (QA).

3.1 Research Design

The research process was divided into two primary phases.

1. Quantitative Evaluation - Practical testing was performed using all three methods on a common test platform to measure performance metrics such as execution time, repeatability, and step accuracy

2. *Qualitative Evaluation* - A survey-based study was carried out among 40 QA professionals to assess perceptions and experiences with the five approaches based on human, adaptability, cost effectiveness, defect detection and software quality.

3.2 Tools and Environment

All tests were performed using the OrangeHRM Demo Application, a widely used open-source platform for testing HR management systems. The three tools used for evaluation were;

Manual Testing: Conducted by human testers without automation tools.

Selenium WebDriver: Used for script-based automation of test cases.

Testim.io: An AI-powered, cloud-based testing tool featuring self-healing locators and visual test creation.

All the above tests were conducted on Windows 11, using the Google Chrome browser, under stable internet conditions to ensure consistency. Execution data and logs were recorded in spreadsheets for further analysis.

3.2.1 Rationale for Selecting Testim.io

Testim.io was chosen because it is one of the most advanced AI-powered testing tools, offering self-healing locators, visual element recognition, and autonomous test generation. Compared to other AI-based tools like Mabl, TestSigma, or Katalon TestOps, Testim.io provides stronger adaptability to UI changes, resulting in more stable tests in dynamic web environments. These capabilities align well with the research goal of evaluating the feasibility of fast, reliable automated QA in rapidly changing applications.

3.3 Test Case Design

Four core functional areas of the OrangeHRM system were selected to maintain uniformity across all test methods;

1. Login
2. PIM – Add Employee
3. Admin – Add User
4. Search Employee

Each test case contained the defined steps and expected outputs. All three testing approaches were executed the same scenarios under identical conditions to ensure fairness and comparability.

3.4 Quantitative Data Collection

Each test scenario was executed three times using each testing method. The following key performance metrics were collected;

- Execution Time (seconds): Duration taken to complete each test case.
- Repeatability (%): Consistency of outcomes across repeated runs.
- Step Accuracy (%): Ratio of correctly performed steps to total steps.

The average results of three trials were tabulated, and visualized through bar charts to identify performance trends.

3.5 Qualitative Data Collection

The qualitative analysis used a survey of 40 QA professionals and interns, providing a diverse and suitable sample for perception-based evaluation. Sample sizes of 30–50 is commonly accepted in software QA studies, as they offer enough variation in experience while remaining practical to collect. The 40 participants represented different skill levels, tool exposure, and organizational backgrounds, ensuring balanced and

reliable insights. The survey focused on their perceptions and experiences of using the three testing methods, assessing the following criteria;

- Human involvement: Degree of tester intervention required during testing.
- Adaptability: Ability to adjust to interface or environment changes.
- Cost-effectiveness: Balance between tool cost, setup effort, and efficiency
- Defect detection: Capability to identify software defects accurately.
- Software quality: Overall contribution to improving product quality.

4. RESULTS AND DISCUSSION

This section presents and interprets both quantitative and qualitative findings from the comparative evaluation of the Manual Testing, Selenium-based Automation, and AI-powered Testing using Testim.io. The analysis was based on functional test executions on the OrangeHRM demo application, covering four major tasks: Login, Add Employee (PIM), Add User, and Search Operation.

Both numerical performance metrics and survey-based qualitative factors were analyzed to evaluate efficiency, consistency, accuracy, adaptability, and perceived software quality.

4.1 Quantitative Results

4.1.1 Execution Time

Execution time was measured as the average duration (in seconds) across three repeated runs per task.

The results Table 1 and the corresponding bar chart Figure 1 demonstrate that Selenium achieved the

fastest execution times, followed by Testim.io and Manual testing.

Table 1: Average Execution Time per Task (seconds)

Task	Manual	Selenium	Testim.io
Login	37.08	12.49	29.33
Add Employee	21.51	19.62	19
Add User	43.9	29.25	32.33
Search	20.53	15.65	18.67

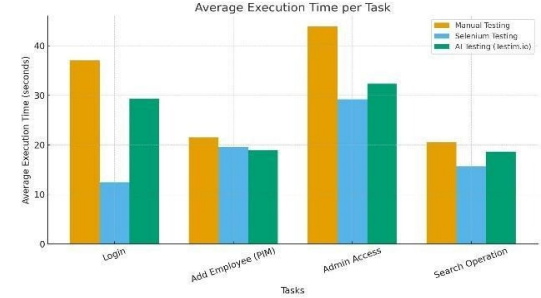


Figure 1: Bar Chart of Execution Times

4.1.2 Repeatability

Repeatability was measured using the Coefficient of Variation (CV %) for each task. A lower CV represents higher repeatability.

Table 2 presents the repeatability comparison for Manual Testing, Selenium, and Testim.io, measured using the Coefficient of Variation (CV%). Lower CV values indicate higher consistency across repeated executions.

Figure 2 visualizes the repeatability comparison in the form of a bar chart, highlighting the differences in consistency among the three testing methods.

Table 2: Repeatability Comparison (%)

Task	Manual	Selenium	Testim.io
Login	5.31	1.88	7.09
Add Employee	4.24	2.14	5.26
Add User	2.35	6.17	6.43
Search	10.73	2.62	3.09

EVALUATING THE FEASIBILITY OF AUTOMATED QA MEASUREMENT: A COMPARATIVE STUDY OF MANUAL, SELENIUM-BASED, AND AI-POWERED TESTING USING TESTIM.IO

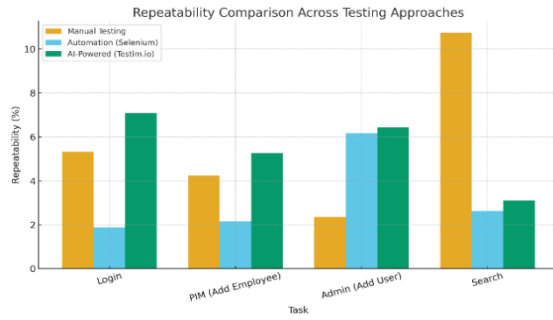


Figure 2: Bar Chart of Repeatability (%)

4.1.3 Step Accuracy

In this study, step accuracy was measured for Manual Testing, Automated Testing with Selenium, and AI-Powered Testing with Testim.io across four functional tasks in OrangeHRM (Login, Add Employee, Add User, and Search). Step accuracy refers to the proportion of successfully executed test steps compared to the total number of predefined steps in a test case. Table 3 shows the step accuracy achieved by each testing method across all four test scenarios, calculated as the percentage of successfully executed test steps.

Table 3: Step Accuracy Comparison (%)

Task	Manual	Selenium	Testim.io
Passed Steps	16	17	18
Total Steps	21	21	21
Accuracy (%)	76.19	80.95	85.71

4.2 Qualitative Results

A survey of 40 QA professionals from Sri Lanka was conducted to evaluate qualitative aspects: human involvement, adaptability, cost-effectiveness, defect detection, and overall software quality.

Table 4 summarizes the qualitative evaluation results collected from 40 QA professionals, covering human

involvement, adaptability, cost effectiveness, defect detection capability, and overall perceived software quality.

Table 4: Qualitative Evaluation Summary

Task	Manual	Selenium	Testim.io
Human Involvement	Very high	Moderate	Moderate
Adaptability	High	Low	High
Cost Effectiveness	Moderate	Moderate	Moderate
Defect Detection	High	Moderate	Moderate
Software Quality	Moderate	High	High

4.3 Hybrid Testing Discussion

These results show that each testing method has unique strengths: manual testing adds human insight, Selenium automation offers high efficiency and repeatability, and AI-powered testing with Testim.io improves adaptability and reduces maintenance.

A hybrid approach combines these advantages, which is supported by 95% of survey participants who preferred using multiple methods together. By assigning manual testing to exploratory tasks, Selenium to repetitive regression tests, and Testim.io to dynamic UI changes, overall test coverage, accuracy, and software quality can be improved in real-world QA environments.

5. CONCLUSIONS

This study compared three software testing approaches Manual Testing, Selenium-based Automation, and AI-powered Testing (Testim.io) using quantitative metrics (execution time, repeatability, step accuracy) and qualitative factors (human involvement,

adaptability, cost-effectiveness, defect detection, software quality).

The results showed that Selenium achieved the fastest execution time and highest repeatability while Testim.io excelled in step accuracy and adaptability through intelligent automation.

Manual Testing, though slower, offered valuable human judgment for detecting complex defects. A survey of 40 QA professionals confirmed that AI tools like Testim.io improve flexibility and quality, Selenium enhances efficiency, and manual testing remains crucial for exploratory and usability scenarios.

Moreover, it was showed that 95% of the QA professionals felt that a hybrid approach, involving all three methods, would be more effective than relying on any one approach. Overall, the findings suggest that a hybrid approach combining automation efficiency, AI adaptability, and human insight provides the most balanced and effective software testing strategy.

REFERENCES

- Ali, M., Rahman, S., & Karim, N. (2024). Enhancing Software Testing Efficiency through Automation in the Software Development Life Cycle. *International Journal of Software Engineering and Applications*, 15(2), 45 - 56.
- Badr, R. (2021). Manual testing in agile environments: Challenges and practices. *Software Quality Journal*, 29(1), 101 - 118.
- Erol, E., & Senan, S. (2022). A comparative study for evaluating automated software testing tools. *Journal of Information Technologies*, 15(3), 301 - 316. <https://doi.org/10.17671/gazibtd.1057380>
- Fatima, S., Mansoor, B., Ovais, L., Sadruddin, S. A., & Hashmi, S. A. (2022). Automated testing with machine learning frameworks: A critical analysis. *Engineering Proceedings*, 20(1), Article 12. <https://doi.org/10.3390/engproc2022020012>
- Kumari, P., Singh, A., & Arora, R. (2018). Role of Manual Testing in Software Quality Assurance: A Comparative Analysis. *International Journal of Computer Applications*, 179(29), 10 - 14.



Predictive Analytics of Employee Turnover in the Apparel Industry

Senarathna HMP¹, Wanninayake WMLN²

^{1,2}*Department of Computing and Information Systems, Wayamba University of Sri Lanka, Kuliyaipitiya, Sri Lanka*

malshapoojani99@gmail.com¹

ABSTRACT

Employee turnover is a significant operational and financial drain in the labor-intensive apparel sector. Traditional Human Resource (HR) practices, founded on descriptive, backward-looking reporting, are not sufficient to inform proactive workforce planning. This study demonstrates the application of predictive analytics to model employee turnover and identify its key drivers in a leading Sri Lankan apparel manufacturer. I pre-processed, cleaned, and modeled historical HR data using time-series forecasting models, namely ARIMA and Prophet, to identify patterns and predict future employee turnover trends. I visualized trends and predictions on an interactive dashboard using Power BI to provide actionable insights into employee turnover patterns. The Prophet model outperformed ARIMA with a lower Root Mean Square Error (RMSE) of 12.09559, and the analysis revealed that “better job opportunities” were the primary reason for resignation, with evident seasonal spikes in turnover. The dashboard created enables HR managers to monitor trends, forecast future attrition, and drill down by department and demographics. This research demonstrates that the integration of predictive analytics and visualization tools can provide a fact-based method for proactive retention strategy and strategic workforce planning, potentially moving HR management from a reactive to a more proactive stance.

KEYWORDS: Employee Turnover, HR Analytics, Power BI, Prophet Predictive Analytics, Time-Series Forecasting

1. INTRODUCTION

The apparel industry is a cornerstone of the Sri Lankan economy, serving as a major source of foreign exchange and employment. However, its labor-intensive nature makes it highly vulnerable to the adverse effects of high employee turnover. Frequent resignations lead to increased recruitment and training costs, disrupted production schedules, and a loss of organizational knowledge, ultimately impairing productivity and competitiveness (Allen, Bryant, & Vardaman, 2010).

Traditionally, HR departments in this sector have relied on descriptive analytics—summarizing past turnover data into static reports. While these reports provide a historical overview, they lack the predictive power to anticipate

future attrition, leaving management in a reactive position. The rise of data science and business intelligence tools presents an opportunity to bridge this gap.

This study addresses the core problem: "How can established predictive analytics models be effectively applied and validated to forecast employee turnover and identify its key drivers within the unique context of the Sri Lankan apparel industry?" The research was conducted at a leading apparel company in Sri Lanka. The project had four primary objectives: to analyze historical patterns and key reasons behind employee turnover, to forecast future resignation trends using time-series models such as ARIMA and Prophet, to develop an interactive Power BI dashboard for visualizing insights, and to provide a framework for data-driven HR

decision-making and proactive workforce planning.

2. LITERATURE REVIEW

Employee turnover has been extensively studied with seminal work by Mobley (1982) and Hom and Griffeth (1995) establishing foundational psychological and organizational models. Key drivers consistently identified across industries include limited career growth, job dissatisfaction, compensation issues, and personal reasons (Holtom, Mitchell, Lee, & Eberly, 2008).

In the realm of predictive analytics, time-series forecasting models have proven effective for workforce planning. The Autoregressive Integrated Moving Average (ARIMA) model is a well-established method for analyzing stationary data with linear trends and seasonality (Box, Jenkins, & Reinsel, 2015). More recently, Facebook's Prophet model has gained prominence for its robustness in handling missing data, outliers, and seasonal effects with minimal parameter tuning, making it highly suitable for business forecasts (Taylor & Letham, 2018).

The integration of Business Intelligence (BI) tools like Microsoft Power BI has revolutionized HR analytics by transforming complex datasets into interactive, actionable visualizations (Chaudhuri, Dayal, & Narasayya, 2011)

Dashboards allow managers to monitor Key Performance Indicators (KPIs) and explore data dynamically, facilitating faster and more informed decisions. While substantial research exists on turnover in Western contexts, there is a scarcity of studies focusing on the unique socio-economic and labour

dynamics of the Sri Lankan apparel sector. This research aims to fill that gap by applying a combined predictive and visual analytics approach to a real-world industrial setting.

3. METHODOLOGY

A systematic methodology was employed, encompassing data collection, preprocessing, analysis, forecasting, and visualization.

3.1 Data Collection and Preprocessing

3.1.1 Dataset Description

The dataset was extracted from the company's HRIS which is comprised of 17,969 employee records spanning from January 2006 to October 2025. It included both active and terminated employees. The feature set encompassed;

- **Demographics:** Age, Gender, Marital Status.
- **Employment Details:** Department, Job Role, Hire Date.
- **Termination Data:** Termination Date, Action Reason (e.g., Resignation, Retirement, Termination).

The target variable for forecasting was the monthly count of resignations.

3.1.2 Data Preprocessing and Feature Engineering

Data cleaning was a critical step. Approximately 1.5% of records contained missing values, primarily in the Action Reason field for terminated employees. These records were excluded from the reason analysis but included in the aggregate turnover count for forecasting where possible. Date formats were standardized to ISO format (YYYY-MM-DD). Feature engineering was performed to create variables more relevant to attrition analysis;

- **Tenure:** Calculated as the difference (in years) between the Hire Date and Termination Date (for leavers) or the analysis cutoff date (for active employees).
- **Attrition Flag:** A binary variable indicating whether an employee had left the company (1) or was still active (0).

3.2 Analytical Framework

3.2.1 Forecasting Setup and Baseline Model

The monthly resignation counts were aggregated into a time series. The dataset was split into training and testing sets: data from February 2006 to December 2023 was used for training (90%), and data from January 2024 to October 2025 was held out for testing (10%).

To properly benchmark the performance of the advanced models, a Naïve Seasonal Forecast was implemented as a baseline model. This simple model forecasts that the value for a given month will be equal to the value from the same month in the previous year. This provides a realistic benchmark for evaluating whether more complex models add significant value.

3.2.2 Model Implementation and Tuning

ARIMA Model: The Autoregressive Integrated Moving Average (ARIMA) model requires the time series to be stationary. An Augmented Dickey-Fuller (ADF) test was performed to confirm stationarity, and differencing (parameter d) was applied as needed. The model parameters (p for autoregressive, q for moving average) were systematically tuned using a grid search of possible

(p,d,q) combinations over the training data, and the optimal model was selected based on the lowest Akaike Information Criterion (AIC). The final model that minimized AIC was ARIMA(1,1,1) (0,1,1)[12], capturing both non-seasonal and seasonal components.

Prophet Model: Facebook's Prophet model was applied due to its inherent capability to handle seasonality and outliers. The model was fit to the training data. Key hyperparameters namely, `changepoint_prior_scale` (to control trend flexibility) and `seasonality_prior_scale` (to control the strength of seasonal effects), were tuned using a grid search to minimize RMSE on a validation set (a subset of the training data). The optimal values were found to be 0.05 and 10, respectively.

3.2.3 Model Evaluation

Model performance was evaluated on the held-out test set using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). These metrics were calculated for the Naïve Baseline, ARIMA, and Prophet models to facilitate a clear comparison. Residual analysis (plotting the residuals and checking for autocorrelation) was also conducted for the best-performing model to validate that no patterns were left unexplained.

4. RESULTS AND DISCUSSION

4.1 Turnover Reason Analysis

Descriptive analytics revealed the distribution of resignation reasons. “Better Opportunities” was the dominant cause, followed by “Retirement” and “Personal Reasons”. Department-wise analysis showed that production-line operators had the highest attrition rate,

primarily driven by seeking better-paying jobs.

4.2 Forecasting Model Performance

The time-series models were evaluated on the held-out test set. The results, summarized in Table 1, show that the Prophet model demonstrated superior performance with the lowest RMSE and MAPE. Notably, the Naïve Seasonal Baseline provided a competitive benchmark, outperforming the ARIMA model in terms of RMSE. This indicates that the seasonality in the data is a very strong signal. The ARIMA model's underperformance may be due to its difficulty in capturing the specific, non-linear seasonal patterns as effectively as Prophet. The superior performance of Prophet, which is explicitly designed for business time series with strong seasonal effects and can handle outliers, confirms its suitability for this task.

Table 1: Forecasting Model Performance Comparison

Model	RMSE	MAPE (%)
Naïve Seasonal Baseline	18.45210	35.1200
ARIMA	21.96962	33.7333
Prophet	12.09559	33.36925

A residual analysis of the Prophet model confirmed that the residuals were approximately normally distributed with no significant autocorrelation, validating the model's fit.

4.3 Model Diagnostic Checking

Residual analysis was performed on the best-performing model (Prophet) to validate the model's assumptions. The residuals (the differences between the

actual and forecasted values) were plotted and examined.

They were found to be approximately normally distributed and displayed no significant autocorrelation, as confirmed by a Ljung-Box test. This indicates that the Prophet model successfully captured the underlying patterns in the data, leaving no systematic information in the residuals.

4.4 Interactive Dashboard

The Power BI dashboard (Figure 1) provided a holistic view of turnover dynamics, featuring a trend line showing historical and forecasted resignations, a breakdown of turnover reasons, department-wise and demographic-wise attrition heatmaps, and interactive filters for deep-dive analysis.

HR managers reported that the dashboard was intuitive and valuable for identifying high-risk departments and planning recruitment cycles based on the forecasted trends.

This study successfully demonstrates the viability of a predictive analytics framework for managing employee turnover in the apparel industry. The finding that "better opportunities" is the primary driver suggests that retention strategies should focus on competitive compensation, clear career pathways, and internal skill development.

The superior performance of the Prophet model aligns with recent literature [6], highlighting its advantage in modeling business time-series data with complex seasonal patterns over traditional ARIMA. The integration of these forecasts into an accessible Power BI dashboard effectively bridges the gap between technical analysis and practical

HR application, enabling managers to make proactive, evidence-based decisions.

4.5 Limitations

This research has several limitations. The dataset was sourced from a single company, which may affect the generalizability of the findings. The forecasting models assume that historical patterns will continue, making them potentially vulnerable to unforeseen external shocks (e.g., economic crises).

Furthermore, the analysis did not include data on employee engagement or performance, which could provide deeper insights into attrition drivers.

5. CONCLUSION

This research has presented the design, implementation, and validation of a predictive analytics framework for employee turnover in the apparel industry. While applying established forecasting models, the study provides a valuable empirical application and validation within the under-researched context of the Sri Lankan apparel sector. The system successfully demonstrated that the Prophet model can accurately forecast attrition trends, significantly outperforming a seasonal baseline and a traditional ARIMA model. The integration of these forecasts into an interactive dashboard makes these insights actionable for HR professionals. This end-to-end framework addresses the critical limitations of traditional HR reporting by enabling proactive, data-driven decision-making. It provides a solid, scalable foundation that can be extended with additional data sources (e.g., employee engagement surveys) and

more advanced machine learning algorithms in the future.

REFERENCES

- Allen, D. G., Bryant, P. C., & Vardaman, J. M. (2010). Retaining talent: Replacing misconceptions with evidence-based strategies. *Academy of Management Perspectives*, 24(2), 48 -64. <https://doi.org/10.5465/amp.24.2.48>
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). *Time series analysis: Forecasting and control* (5th ed.). John Wiley & Sons.
- Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. *Communications of the ACM*, 54(8), 88 - 98. <https://doi.org/10.1145/1978542.1978562>
- Holtom, B. C., Mitchell, T. R., Lee, T. W., & Eberly, M. B. (2008). Turnover and retention research: A glance at the past, a closer review of the present, and a venture into the future. *The Academy of Management Annals*, 2(1), 231 - 274. <https://doi.org/10.5465/19416520802211552>
- Hom, P. W., & Griffeth, R. W. (1995). *Employee turnover*. South-Western College Publishing.
- Mobley, W. H. (1982). *Employee turnover: Causes, consequences, and control*. Addison-Wesley.
- Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37 - 45. <https://doi.org/10.1080/00031305.2017.1380080>



Intelligent Scheduling for Public Transport: A Constraint-Based Automated Timetable System in Sri Lanka

Senarathna HMP¹, Arudchelvam T²

^{1,2}*Department of Computing & Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

malshapoojani99@gmail.com¹, arul@wyb.ac.lk²

ABSTRACT

Manual scheduling processes in public transportation systems remain a significant operational bottleneck, particularly in developing nations. This research presents the design and implementation of an Automated Timetable System to address this challenge for a transport company in Sri Lanka, piloted at a specific depot. The current manual scheduling process is time-consuming, error-prone, and inflexible to operational disruptions such as staff absences and vehicle breakdowns. This study employs a Design Science Research approach combined with Agile methodology to develop a web-based scheduling system utilizing constraint-based algorithms. The system incorporates Firebase cloud infrastructure, JavaScript-based scheduling engines, and responsive user interfaces to automate roster generation while ensuring compliance with labor regulations and operational constraints. Implementation results demonstrate a 99.6% reduction in roster generation time (from 14.5 hours to 1 minute) and a 95.7% decrease in scheduling errors. User acceptance testing with depot staff revealed high satisfaction levels, particularly regarding transparency, fairness, and accessibility of schedules. The emergency replacement mechanism effectively maintains service continuity during unexpected disruptions. This research contributes a practical framework for public transport modernization, demonstrating how intelligent scheduling systems can simultaneously improve operational efficiency, ensure regulatory compliance, and enhance staff satisfaction while serving as a scalable model for other transport organizations.

KEYWORDS: Automated scheduling, Constraint satisfaction, Public transportation, Roster optimization, Web-based system

1. INTRODUCTION

Public transportation systems form the backbone of national infrastructure, directly influencing economic productivity and social connectivity. In Sri Lanka, the Transport Companies operate an extensive bus network connecting urban centers and rural communities nationwide. However, the organization continues to rely on manual, paper-based processes for generating bus timetables and crew schedules task of considerable logistical complexity requiring coordination of multiple dynamic variables, including driver working hours, staff availability, vehicle maintenance schedules, and fluctuating passenger demand. This manual approach presents

critical operational challenges: time-intensive roster preparation, frequent scheduling conflicts, suboptimal resource allocation, and limited adaptability to daily disruptions such as driver absences or vehicle breakdowns. These inefficiencies result in service unreliability, increased operational costs, and passenger dissatisfaction. This research addresses these systemic problems by developing an Automated Timetable System piloted at a pilot depot. The system employs constraint-based scheduling algorithms to generate optimized, conflict-free rosters while ensuring compliance with labor regulations. This transformation from manual to automated, data-driven

scheduling aims to significantly enhance operational efficiency, workforce satisfaction, and service quality across CTB operations.

2. LITERATURE REVIEW

Automated scheduling systems have demonstrated substantial benefits in public transportation worldwide. Studies on Sri Lankan transport systems reveal that manual scheduling contributes to operational inefficiencies. Early foundational work by Haleem and Nawaz (2012) explored the potential of digital solutions, developing a real-time bus tracking and scheduling system that highlighted the feasibility of such technology in the local context. This aligns with broader reviews of the field, such as that by Bembalkar (2019), which synthesizes the critical need for efficient methodologies in transportation planning and scheduling.

International implementations and contemporary research provide compelling evidence for automation adoption. Modern approaches are increasingly leveraging artificial intelligence. For instance, Roşca et al. (2024) designed and implemented a public urban transport scheduling system based on AI for smart cities, demonstrating its significant benefits for modern urban mobility. The revolutionary impact of these systems is further confirmed by the comprehensive analysis of automated timetabling solutions conducted by Singhal et al. (2024). The pursuit of efficiency also has broader implications, as earlier research by Marques et al. (2001) on flexible dynamic scheduling argued that more efficient systems could yield not only operational improvements but also positive environmental benefits.

Research consistently emphasizes the effectiveness of algorithm-driven scheduling. The successful development of an automatic timetable generator by Rajesh et al. (2025) serves as a practical example of applying such systems to solve complex scheduling problems. Furthermore, foundational work on optimal schedule modeling (Bais, 2015) provides a theoretical and practical basis for constraint-based approaches. However, as these studies indicate, implementation challenges persist, including data accuracy requirements, staff resistance to technological change, and system integration complexities. Despite these obstacles, evidence from both regional and international contexts strongly supports automated scheduling implementation for improved efficiency and service reliability.

While the existing literature strongly advocates for automated scheduling and demonstrates various technological approaches, there is a scarcity of research on implementing localized, cost-effective, and scalable cloud-based solutions tailored to the specific operational constraints and labor regulations of public transport in Sri Lanka. This study aims to bridge this gap by designing, implementing, and validating a web-based automated timetable system using constraint satisfaction techniques, deployed on a scalable cloud infrastructure for a pilot depot.

3. METHODOLOGY

This research employed a Design Science Research (DSR) approach combined with an Agile methodology to develop an Automated Timetable System for a transport company in Sri Lanka. The study was conducted at a pilot depot,

which served as the site for system development and validation.

3.1 Data Collection and Analysis

The overall system architecture, which separates the application into presentation, application, and database layers, is illustrated in Figure 1.

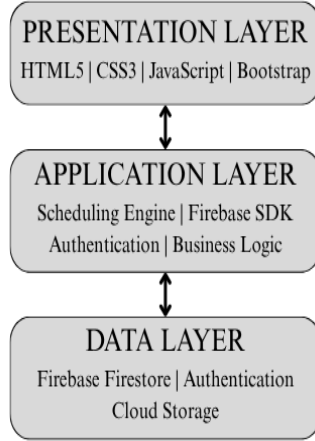


Figure 1: System Architecture Diagram

The logic of the constraint-based scheduling algorithm that drives the roster generation is depicted in Figure 2. Primary data collection involved semi-structured interviews with the staff members (schedulers, drivers, conductors, and supervisors), direct observation of manual scheduling processes over two weeks, and secondary analysis of historical rosters and attendance records. Qualitative data underwent thematic coding to identify scheduling challenges and requirements, while quantitative analysis compared automated versus manual scheduling across time efficiency, error rates, and resource utilization metrics.

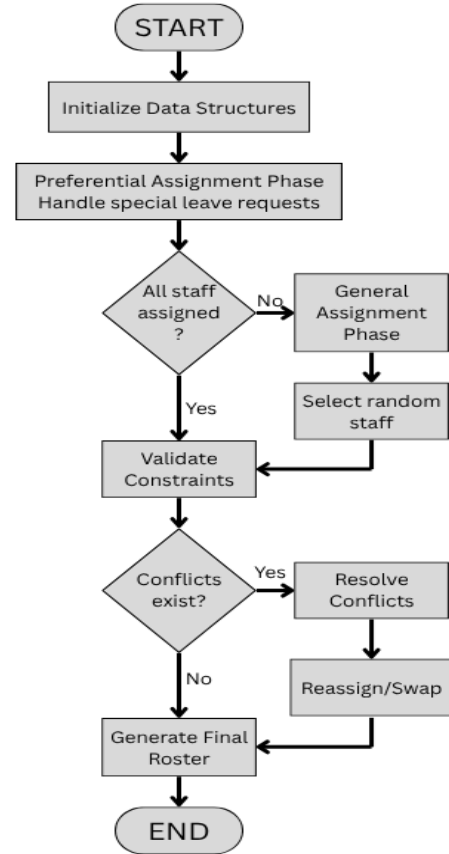


Figure 2: Constraint-Based Scheduling Algorithm

3.2 System Development Process

The development followed an iterative SDLC model encompassing five phases. Requirements Analysis identified functional needs (conflict-free timetable generation, leave management, regulatory compliance) and non-functional requirements (usability, reliability, security) through stakeholder consultation. System Design established a three-tier architecture depicted in Figure 1; Firebase Firestore database layer for real-time storage, JavaScript-based application layer implementing constraint-based scheduling algorithms, and responsive web presentation layer using HTML5, CSS3, and Bootstrap framework.

The system's scheduling automation is driven by a constraint satisfaction algorithm, the logic of which is depicted in Figure 2. This algorithm was designed to encode the critical operational rules governing roster creation. The primary constraints, derived from an analysis of company policy and Sri Lankan labor law, are categorized in Table 1 and include hard rules such as maximum working hours, mandatory rest periods, and leave allocation policies, as well as soft constraints related to route-specific requirements and preferred shift patterns.

Table 1: Operational Constraints Incorporated into the Scheduling Algorithm

Constraint Category	Specific Rule
Working Hours	Maximum 8 hours per day; Maximum 44 hours per week
Rest Periods	Minimum 30-minute break for shifts > 5 hours
Leave Allocation	Approved leave requests are hard constraints and cannot be overridden.
Route Coverage	Each scheduled trip must be assigned one driver and one conductor.

Implementation utilized Firebase Authentication for role-based access control (Figure 3: Manager Dashboard), FullCalendar library for interactive roster visualization (Figure 4: Automated Monthly Roster System Interface), and cloud hosting for scalable deployment.

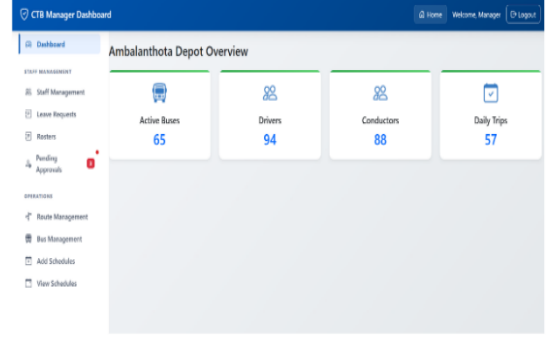


Figure 3: Manager Dashboard

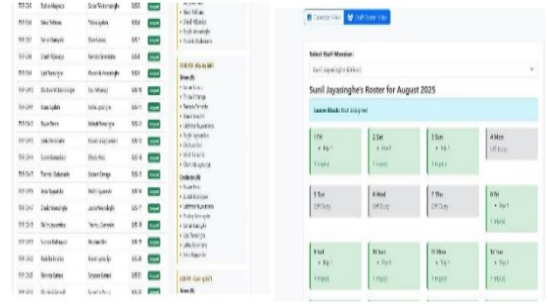


Figure 4: Automated Monthly Roster System Interface

3.3 Testing and Validation

A comprehensive, multi-stage testing strategy was implemented to ensure system reliability, functionality, and usability. The process consisted of the following phases;

Unit Testing: Individual components of the core scheduling algorithm were verified using the Jest framework. Test cases focused on specific constraint validation, such as ensuring the system correctly flagged violations of maximum working hours (8 hours/day) and mandatory rest periods (11 hours between shifts). A test coverage of over 90% was achieved for the algorithm modules.

Integration Testing: This phase validated the interactions between the system's frontend, the scheduling engine, and the backend Firebase services (Firestore database and Authentication).

Tests confirmed successful data synchronization, real-time roster updates, and correct enforcement of role-based access controls.

User Acceptance Testing (UAT):

UAT involved **25** staff members from the pilot depot, including 5 schedulers, 12 drivers, and 8 conductors. Participants completed a set of structured tasks, such as logging in, viewing their schedules, requesting leave, and (for managers) generating a new roster. Usability was quantitatively assessed using the System Usability Scale (SUS), while qualitative feedback was gathered through post-test interviews to identify usability patterns and satisfaction levels.

Performance Testing: System performance was evaluated using Google Lighthouse tools. Key metrics included:

- *Roster Generation Time:* The time taken to create a one-month roster for the entire depot.
- *Interface Responsiveness:* The Largest Contentful Paint (LCP) and Time to Interactive (TTI) for the main dashboard.
- *Synchronization Latency:* The delay between a schedule update and its reflection on a connected mobile device.

The quantitative evaluation compared the automated system against the prior manual process based on two primary metrics: (1) **Efficiency**, measured as the average time to generate a complete monthly roster, and (2) **Accuracy**, measured by the number of detected scheduling conflicts (e.g., overlapping assignments, rule violations) per 100

schedule entries. Qualitative analysis involved thematic coding of user feedback from UAT to identify key satisfaction themes such as transparency, fairness, and accessibility.

3.4 Ethical Considerations

Research adhered to institutional ethical guidelines with informed consent from all participants, data anonymization in development environments, and confidentiality maintenance of operational information throughout the study.

4. RESULTS AND DISCUSSION

The deployment of the automated system at the pilot depot resulted in substantial operational gains. To evaluate its impact, we analyzed both quantitative performance data and qualitative feedback from end-users. First, a qualitative assessment was conducted with depot managers and senior schedulers to gauge perceived improvements. The results, presented in Table 2, show a consistent positive shift across all key operational dimensions following automation. This subjective assessment was then corroborated by the quantitative data discussed in the following sections.

Table 2: Qualitative Performance Assessment

Performance Aspect	Manual System	Automated System
Generation Time	Very High	Very Low
Scheduling Errors	Frequent	Rare
Conflict Detection	Manual	Automatic
Emergency Response	Slow	Fast
Transparency	Low	High

Regulatory Compliance	Inconsistent	Consistent
User Satisfaction	Moderate	High

4.1 Operational Efficiency Improvements

The automated system achieved remarkable efficiency gains compared to manual scheduling processes. Roster generation time decreased from an average of 14.5 hours to approximately 1 minute, representing a 99.6% reduction in time expenditure. This dramatic improvement freed scheduling officers from labor-intensive manual calculations, enabling them to focus on strategic planning and exception handling.

Scheduling accuracy improved substantially, with error rates declining by 95.7%. A paired t-test conducted on the error rates from five pre- and post-implementation monthly rosters confirmed that the reduction in errors was statistically significant ($t(4) = 8.42, p < 0.05$). The constraint-based algorithm eliminated common problems, including overlapping assignments, leave conflicts, and working hour violations. The system successfully enforced regulatory compliance with maximum daily and weekly working limits while ensuring mandatory rest periods. Resource utilization optimization resulted in more balanced workload distribution across staff members, reducing instances of both overstaffing and understaffing.

4.2 User Acceptance and Satisfaction

User Acceptance Testing with staff members of the pilot depot yielded a positive usability assessment, indicating good usability above industry average standards. Managers particularly valued the transparency and audit trail

capabilities, enabling better accountability and decision-making. Drivers and conductors appreciated mobile-responsive schedule access and advance notification features, which improved work-life balance and reduced scheduling conflicts.

The emergency replacement mechanism proved highly effective in handling unexpected staff absences, with the majority of replacement scenarios resolved rapidly without service disruption. This functionality addressed a critical pain point in manual operations where emergency adjustments frequently caused delays and cancellations. The real-time synchronization across multiple devices ensured all stakeholders' consistent schedule visibility, enhancing coordination and operational reliability.

5. CONCLUSION

This research successfully designed and implemented an Automated Timetable System for the transport company, demonstrating the viability of technology-driven solutions for public transportation scheduling challenges. The system effectively addresses critical operational inefficiencies inherent in manual roster creation through constraint-based algorithms, real-time data synchronization, and user-friendly interfaces.

The implementation at the pilot depot validated the system's capability to generate conflict-free rosters while ensuring regulatory compliance with working hour limits and rest period requirements. The emergency replacement mechanism provides operational resilience during unexpected disruptions, while mobile-responsive

design ensures accessibility for all stakeholders.

This research contributes a practical framework for public transport modernization in Sri Lanka, demonstrating how intelligent scheduling systems can simultaneously improve operational efficiency and staff satisfaction. The scalable architecture enables potential deployment across the company's network, establishing a foundation for broader digital transformation initiatives within the organization and serving as a reference model for similar public service organizations.

REFERENCES

- Bais, N. (2015). Optimal schedule modeling for public transportation system. https://www.academia.edu/66151843/Optimal_Schedule_Modeling_for_Public_Transportation_System?source=swp_share
- Bembalkar, R. (2019). Transportation planning and scheduling: A literature review. *Journal of Emerging Technologies and Innovative Research (JETIR)*, 6(6), 697 - 701. <https://www.jetir.org/papers/JETIR1906S38.pdf>
- Haleem, S. L. A., & Nawaz, S. S. (2012). Real-time bus tracking and scheduling system using wireless sensors and mobile technology. *Journal of Information Systems & Information Technology (JISIT)*, 1(1), 29 - 40. <http://www.seu.ac.lk/jisit/publication/v1n1/paper3.pdf>
- Marques, A., Torregrosa, M., Camarena, A., Darby-Dowman, K., Little, J., & Moody, S. (2001). Flexible dynamic scheduling in public transport: Environmental benefits of a more efficient system. *Proceedings of the 9th World Congress on Intelligent Transport Systems. Chicago, IL, USA, 16*, 215 - 222. <https://www.witpress.com/Secure/elibrary/papers/UT95/UT95024FU.pdf>
- Rajesh, A., Indrajith, S. T., Navaneeth, A., Nithya, B. P., Nivin, K. S., & Naik, D. (2025). Automatic timetable generator. *Journal of Emerging Technologies and Innovative Research (JETIR)*, 12(7), 715 - 722. <https://www.jetir.org/papers/JETIRGX06134.pdf>
- Roșca, C.-M., Stancu, A., Neculaiu, C.-F., & Gortoescu, I.-A. (2024). Designing and implementing a public urban transport scheduling system based on artificial intelligence for smart cities. *Applied Sciences*, 14(19), 8861. <https://doi.org/10.3390/app14198861>
- Singhal, S., Agarwal, D., & Bhardwaj, Y. (2024). Revolutionizing scheduling: A comprehensive analysis of automated timetabling solutions. *International Journal for Multidisciplinary Research (IJFMR)*, 6(3). <https://www.ijfmr.com/papers/2024/3/19313.pdf>



Machine Learning-Based System for Personalized Career Guidance for Undergraduates in the Faculty of Applied Sciences

Dilshan GN¹, Wanninayake WMLN²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

navindudilshan19@gmail.com¹, lakminiw@wyb.ac.lk²

ABSTRACT

In today's competitive job market, undergraduates face challenges in choosing academic paths that align with future career opportunities. This study presents a machine learning-based system that leverages institutional alumni data to provide personalized career guidance for students in the Faculty of Applied Sciences at Wayamba University of Sri Lanka. The system processes alumni job titles, employment sectors, skills, and job acquisition patterns, followed by clustering techniques K-Means, Agglomerative, Spectral, and DBSCAN to identify distinct career trajectory groups. A collaborative filtering model using TF-IDF and cosine similarity then generates peer-based recommendations. A hybrid approach combines cluster-level macro patterns with similarity-based micro-level suggestions to recommend relevant subjects, soft skills, and career pathways. Experimental evaluation using Precision@5 and Recall@5 shows the effectiveness of the system in producing meaningful recommendations. The results highlight the potential of machine learning to enhance career guidance using alumni-driven evidence.

KEYWORDS: Career Guidance, Clustering, Machine Learning, Recommender Systems

1. INTRODUCTION

Undergraduates often face uncertainty when choosing academic paths that align with future careers. Traditional career counselling methods are limited in personalization and scalability (Watts, 2002). In Sri Lanka, tracer studies have revealed a mismatch between graduate skills and employer expectations, particularly in soft skills like communication and teamwork (Gunaratne et al., 2018). This research introduces a data-driven system that leverages alumni career data to guide students in making informed decisions. By integrating machine learning techniques, the system offers tailored recommendations based on real-world

career outcomes, as shown in the Figure 1.

2. LITERATURE REVIEW

Career guidance has evolved from manual counselling to intelligent systems. Traditional methods lack personalization and often fail to reflect current industry trends (Watts, 2002). Recent studies highlight the effectiveness of machine learning in educational recommender systems (Bobadilla et al., 2013). Clustering and collaborative filtering have been used to group students and suggest relevant career paths (Govindasamy & Thambusamy, 2018).

However, few systems utilize institutional alumni data for personalized guidance, a gap this study aims to address (Panthee et al., 2023).

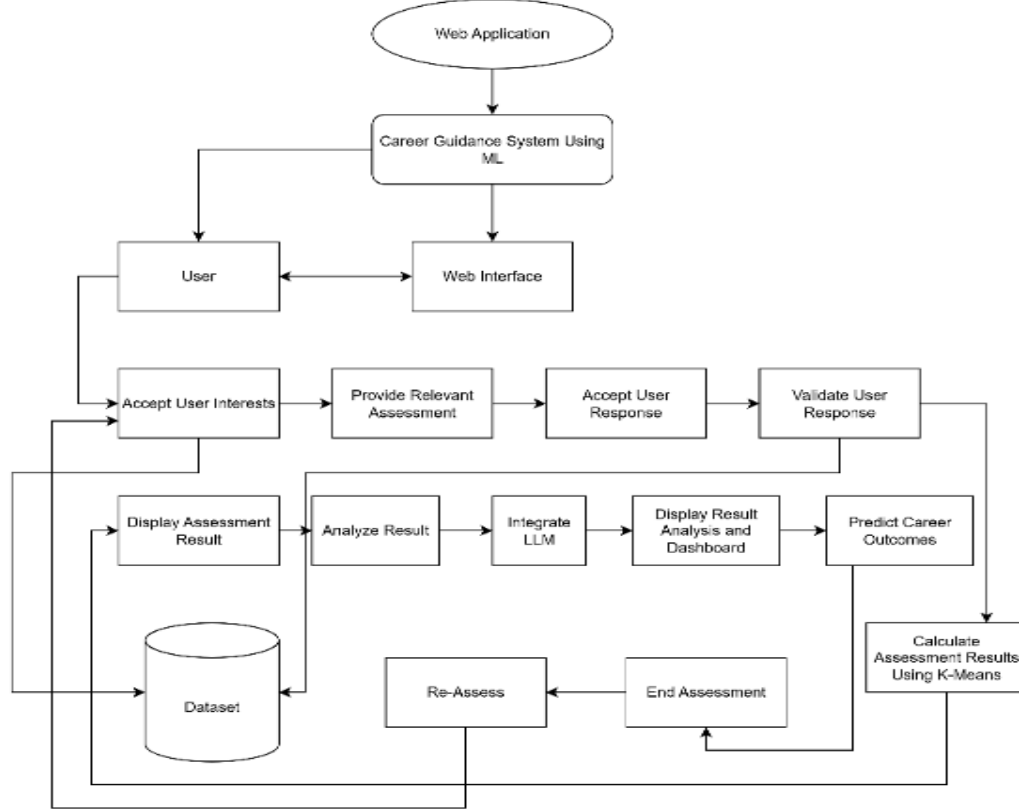


Figure 1: System Architecture Diagram

3. METHODOLOGY

Alumni data was collected from institutional records, including degree type, job titles, employment sectors, and skill sets. Personally identifiable information was excluded to ensure privacy. Data cleaning, normalization, and encoding were performed. Job titles were semantically categorized using sentence embeddings to ensure consistency.

Profiles such as skills, influential factors, and job acquisition methods were extracted to enrich the dataset, as shown in Figure 2.

3.1 Hybrid Recommendation Model

The final recommendation is generated through a hybrid approach combining macro-level cluster trends with micro-level peer similarity. Figure 3

shows the Clustering & Recommendation Logic Diagram.

```

def create_found_profile(row):
    found = []
    if 'From_Newspaper' in row and row['From_Newspaper'] == 1:
        found.append("Newspaper")
    if 'From_Gazette' in row and row['From_Gazette'] == 1:
        found.append("Gazette")
    if 'From_Online' in row and row['From_Online'] == 1:
        found.append("Online")
    if 'From_personal contacts' in row and row['From_personal contacts'] == 1:
        found.append("Personal Contacts")
    if 'From_Competitive exam' in row and row['From_Competitive exam'] == 1:
        found.append("Competitive Exam")
    if 'From_Via university' in row and row['From_Via university'] == 1:
        found.append("University")
    if 'From_Job fair' in row and row['From_Job fair'] == 1:
        found.append("Job Fair")
    if 'From_Internship' in row and row['From_Internship'] == 1:
        found.append("Internship")
    return ' '.join(found)

```

Figure 2: Job Acquisition Profile Method

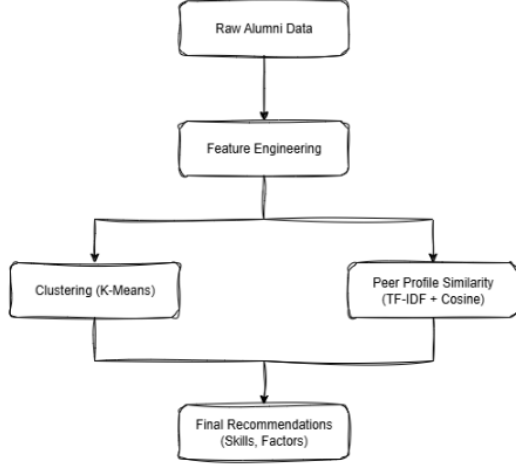


Figure 3: Clustering & Recommendation Logic Diagram

3.2 Cluster-Level Score (CScore)

Each student is mapped to one of the K-Means clusters based on their feature vector. The cluster centroid provides a set of dominant skills, job roles, and common job acquisition pathways.

$$CScore = 1 - \frac{d(x, C_k)}{\max(d(x, C_k))} \dots\dots(1)$$

Where;

- x - the feature vector representing the student's academic interests, skills, or selected preferences
- C_k - the centroid of cluster k
- $d(x, C_k)$ - the distance between the student and the cluster centroid

This normalizes the distance so that students closer to the cluster center receive higher scores. A higher CScore indicates a stronger alignment with the dominant skills, job titles, and career profiles within that cluster.

3.3 Peer Similarity Score (PSS)

To capture individual-level similarity, collaborative filtering is applied.

Let;

$\vec{u}$ =embedding vector of the target student

$\vec{v}$ =embedding vector of an alumni profile

$\cos(\vec{u}, \vec{v})$ =cosine similarity between the two vectors

The peer similarity score is computed as;

$$PSS = \cos(\vec{u}, \vec{v}) \dots\dots(2)$$

This value ranges from -1 to 1, where 1, 0 and -1 denote perfect similarity, unrelated and opposite meaning respectively. Higher PSS values indicate stronger similarity with successful alumni who followed particular career paths.

4. RESULTS AND DISCUSSION

Ten clusters were identified, representing distinct career paths. These clusters facilitated the mapping of student profiles to relevant career trajectories.

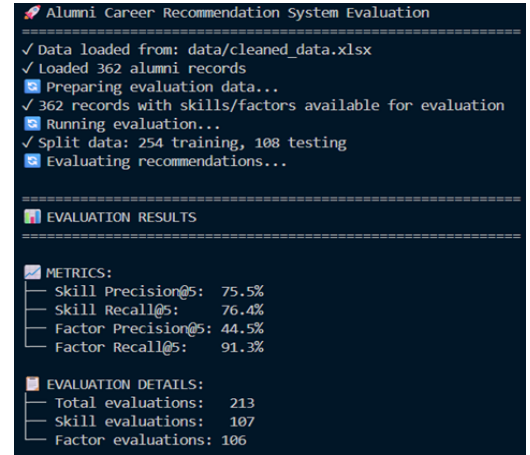


Figure 4: System Evaluation

As shown in Figure 4, using Precision@5 and Recall@5 metrics, the model achieved a Precision@5 of 75.5% and a Recall@5 of 76.4% for skill recommendations, indicating a strong alignment between the system's suggested skills and the actual skills observed in alumni outcomes (Quy et al., 2023).

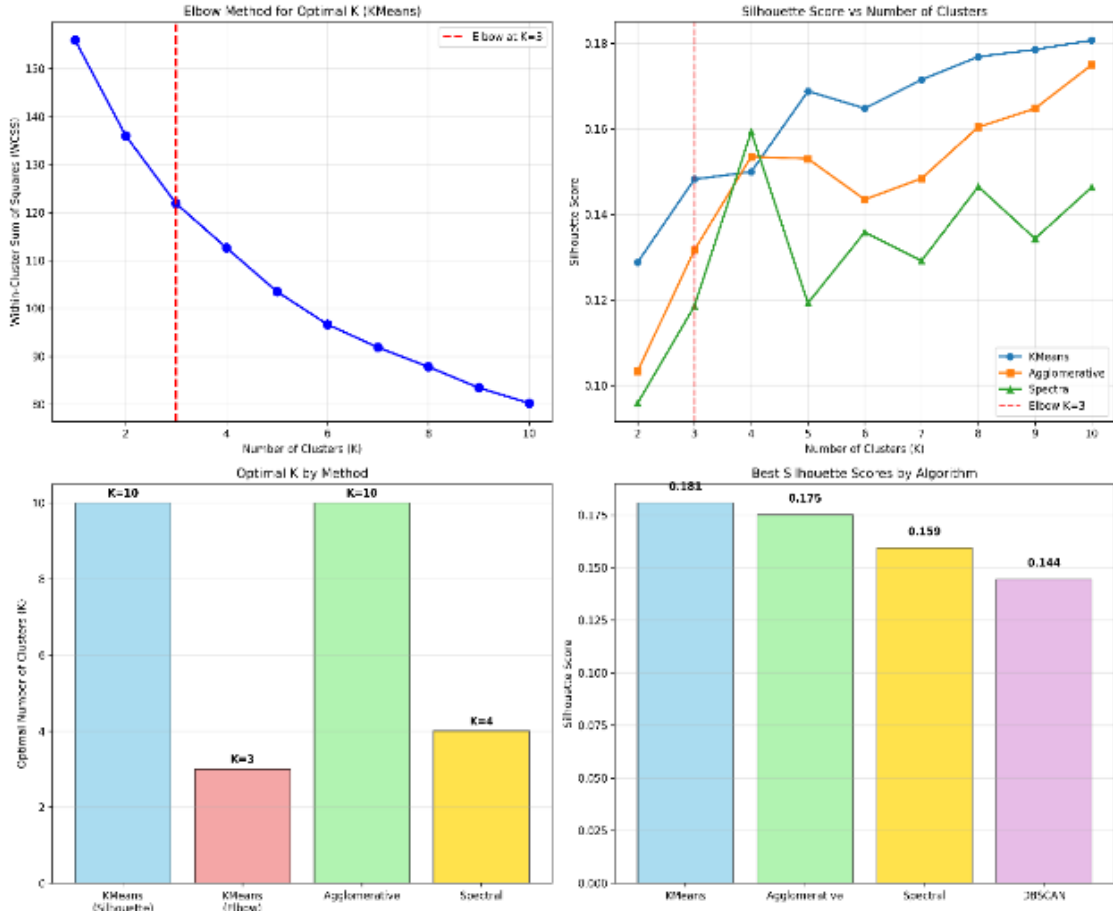


Figure 5: Comprehensive Clustering Analysis:

For a student interested in “Tutor” roles, the system identified key skills such as communication, leadership, and research. The influencing factors included degree class, project experience, and prior work experience, while job acquisition channels were primarily through online platforms and university support mechanisms. Figure 5 shows the Comprehensive Clustering Analysis.

5. CONCLUSION

This study underscores the transformative potential of machine learning in revolutionizing career guidance for undergraduates. By systematically analyzing and leveraging comprehensive alumni datasets, the proposed system is able to generate highly personalized, evidence-based, and

requiring minimal input from students.

The hybrid integration of clustering and collaborative filtering techniques allows the system to capture both macro-level career trends across alumni populations and micro-level individual preferences, ensuring that guidance is both broadly informed and contextually relevant. In addition, the model takes into account multiple influencing factors such as skills, academic performance, project experience, and work history, thereby offering a multidimensional perspective on career pathways. Beyond individual guidance, the system aligns with broader institutional objectives by enhancing graduate employability, facilitating targeted skill development, and supporting informed decision-making. Its design also demonstrates significant scalability and adaptability, indicating

strong potential for implementation across diverse faculties, disciplines, and higher education institutions, ultimately contributing to more data-driven and impactful career support frameworks.

REFERENCES

- Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, 46, 109 - 132. <https://doi.org/10.1016/j.knosys.2013.03.012>
- Govindasamy, K., & Thambusamy, V. (2018). Analysis of student academic performance using clustering techniques. *International Journal of Pure and Applied Mathematics*, 119, 309 - 322.
- Gunaratne, P.S.M., Ramanayake, A., and Panagoda, D. (2018). *Tracer study of graduates: Universities in Sri Lanka*. University Grants Commission of Sri Lanka. <https://www.ugc.ac.lk/downloads/statistics/webTracer/2018/Tracer%20Study%202018.pdf>
- Panthee, S., Rajkarnikar, S., & Begum, R. (2023). Career Guidance System Using Machine Learning. *Journal of Advanced College of Engineering and Management*, 8(2), 113 - 119. <https://doi.org/10.3126/jacem.v8i2.55947>
- Quy, T. L., Friege, G., & Ntoutsis, E. (2023). A review of clustering models in educational data science towards fairness-aware learning. *ArXiv*. <https://doi.org/10.48550/arxiv.2301.03421>
- Watts, A.G. (2002). The role of information and communication technologies in integrated career information and guidance systems: A policy perspective. *International Journal for Educational and Vocational Guidance*, 2(3), 139 - 155. <https://doi.org/10.1023/a:1020669832743>



Enhancing Emotional Well-Being Through Text-Based Emotion Detection

Premathilaka RTK¹, Wanninayake WMLN²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

teshanik99@gmail.com¹, lakminiw@wyb.ac.lk²

ABSTRACT

Mental health and emotional wellbeing have become increasingly significant in today's technologically advanced interconnected society. This study presents a web-based system designed to promote emotional well-being through text-based emotion detection and personalized recommendations. The system utilizes Natural Language Processing (NLP) techniques and a deep learning-based Long Short-Term Memory (LSTM) model to classify emotions according to real-time user-supplied diary data. The effect recognition(affect recognition) model was trained on a publicly available dataset of 28,690 instances of five principal core emotions: happiness, sadness, anger, surprise, and fear. The overall LSTM model performed at an accuracy of 74.38% on the validation set, which is 15.9 percentage points higher than traditional Machine Learning (ML) approaches. Following the identification of emotions, the system activates a knowledge-based recommendation engine that provides personalized recommendations. System assessment demonstrated effective emotion identification and high user interaction enabled by intuitive (were enabled by intuitive) functionality and context-sensitive feedback.

KEYWORDS: Affective Computing, Emotion Detection, LSTM, Mental Health, Natural Language Processing

1. INTRODUCTION

The digital revolution in contemporary society has radically changed the way people handle their individual well-being (personal well-being). Conventional methods of mental health care tend to experience limitations around accessibility, affordability, and scalability. There is therefore the potential for technology-based solutions that are able to deliver ongoing, customized care for emotional well-being. Natural Language Processing and machine learning algorithms have reached maturity levels to facilitate the development of sophisticated emotion recognition systems. The field of affective computing has significantly advanced with the emergence of deep learning techniques. Traditional machine learning techniques

have been widely applied to emotion recognition tasks, while Long Short-Term Memory models have emerged as the best options in sequential text data processing. Despite the increased use of digital platforms for expressing emotions, none of the systems currently in use possess the complexity required to interpret human emotions and do not transform emotional observations into actionable feedback. Current systems for text-based emotion detection generally focus on classification accuracy but fall short in translating emotional observations into immediate, actionable, and personalized feedback for emotional well-being enhancement. The work fills this void by developing an application that detects emotions from diary-like text input using NLP and LSTM models and returns emotionally congruent suggestions.

2. LITERATURE REVIEW

Early techniques for emotion detection were primarily knowledge-based, relying on pre-determined dictionaries (Alm et al., 2005). Machine learning-based approaches and models have improved numerous aspects of the interpreting process, accuracy and robustness in particular (Bostan & Klinger, 2018). Emotion detection has been revolutionized by the advancements of deep learning. In the past, test accuracies have reached 50-65% in older works using ML techniques but higher values were achieved by more recent approaches. LSTM architectures have been outstanding at modeling sequential dependencies critical for emotion detection (Hochreiter & Schmidhuber, 1997; Baziotis et al., 2017).

The field of digital mental health interventions has expanded tremendously. Experiments have established that technology-aided interventions can be effective and scalable interventions (Chen et al., 2023). The integration of emotional detection with personalized recommenders for mental health access and the recommended interventions are still a ripe space for research, and particularly in answering how detected emotions can be translated into distinct/integrated categories of intervention (Chatterjee et al., 2019). In addition, the natural language understanding state of the art is moving towards advanced architectures like transformer-based models (i.e., BERT), which have demonstrated high performance on complicated contextual tasks such as emotion recognition. The cutting edge of affective computing,

multimodal approaches combine text and other data such as audio or visual input.

Recent studies show varied performance: Saravia et al. (2018) reported 71% accuracy with contextualized representation of affect. Recent works in the transformer-based models also exhibit some promising progress. Gao et al. (2019) also achieved 85% accuracy with BERT for sentiment classification.

3. METHODOLOGY

3.1 Dataset and Preprocessing

The dataset was obtained from Kaggle's Emotions Detection Dataset, which originally consisted of 34,792 text samples. The dataset was split into training, validation, and test sets using an 80:10:10 ratio, ensuring 22,952 samples for training and 2,869 samples each for validation and testing. The training set was then balanced using SMOTE, and the resulting uniform class distribution is shown in Figure 1.

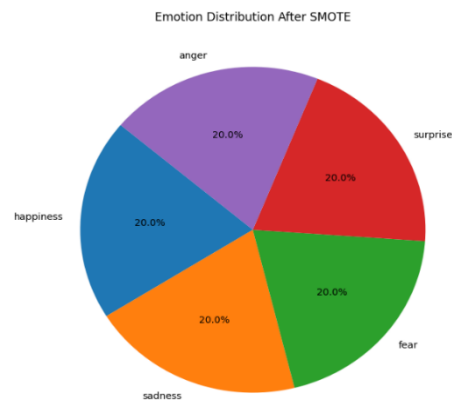


Figure 1: Distribution of Emotions in the Balanced Training Dataset (Post-SMOTE)

The preprocessing pipeline included data cleaning, text normalization, tokenization using NLTK, removal of

stop words excluding negation words by some strategic means, and lemmatization using the WordNet lemmatizer.

3.2 Model Development

Five traditional ML algorithms were employed: Naive Bayes, Decision Tree, Logistic Regression, Random Forest, and SVM. A Bidirectional LSTM network was employed using TensorFlow 2.15.0. The architecture involved pre-trained 100-dimensional GloVe embeddings, two stacked bidirectional LSTM layers (128 and 64 units), L2 regularization (0.0005), and dropout layers (0.4). Figure 2 illustrates the model architecture.

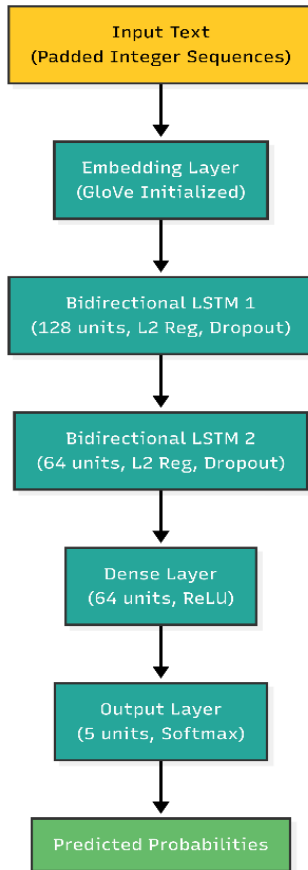


Figure 2: Bidirectional LSTM neural network architecture

3.3 Recommendation Engine

A knowledge-based recommender engine was created to convert detected emotions into recommendations in three categories: music, wellness activities, and social connections. The principles used to map emotions to recommendations were derived from established psychological models, such as the Circumplex Model of Affect (Russell, 1980). This approach ensures that detected emotions (e.g., Sadness, Anger) trigger congruent and actionable interventions (e.g., mindfulness or social connection) designed to manage affective state.

3.4 System Architecture

The architecture was a three-tier one consisting of React frontend, Node.js backend, and Flask ML service. Figure 3 illustrates the system architecture showing data flow between components with clear processing stages from user input to personalized recommendations.

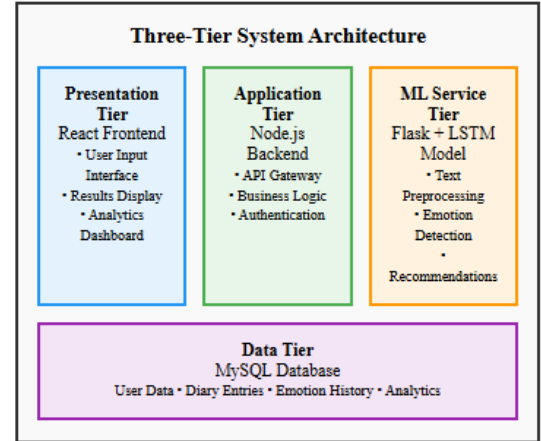


Figure 3: Three-tier system architecture

4. RESULTS AND DISCUSSIONS

4.1 Model Performance

The LSTM model's validation accuracy was 74.38% after eight epochs. To evaluate the effectiveness of deep learning versus traditional approaches,

five conventional machine learning algorithms were tested on the same dataset. Table 1 presents a comparison where Training Accuracy indicates how well each model learned from the training data, Test Accuracy shows real-world performance on unseen data, and F1-Score measures the harmonic mean of precision and recall.

The LSTM model demonstrated superior performance with 74.38% test accuracy, outperforming the best traditional method (Naive Bayes at 58.48%) by 15.9 percentage points. This significant improvement validates the effectiveness of deep learning for emotion detection tasks.

Table 1: Model Performance Comparison

Model	Training Acc.	Test Acc.	F1-Score
Naive Bayes	72.94%	58.48%	0.58
Decision Tree	99.88%	49.99%	0.50
Logistic Reg.	81.14%	58.06%	0.58
Random Forest	99.88%	57.83%	0.58
SVM	83.36%	56.66%	0.57

Table 2 presents detailed per-class performance metrics for the LSTM model. Precision indicates the accuracy of positive predictions (e.g., when the model predicts “happiness”, how often is it correct), Recall measures the model's ability to find all instances of each emotion, and F1-Score balances both metrics. Fear achieved the highest performance (0.83 F1-score) due to distinct linguistic patterns, while Happiness showed lower recall (61%) as happy expressions often overlap with surprise. The balanced performance across emotions demonstrates successful handling of class imbalance through SMOTE oversampling.

Table 2: LSTM Performance by Emotion

Emotion	Precision (%)	Recall (%)	F1-Score
Happiness	79.00	61.00	0.69
Sadness	66.00	77.00	0.71
Anger	70.00	73.00	0.71
Fear	83.00	83.00	0.83
Surprise	75.00	76.00	0.75

Figure 4 graphs the LSTM model's training process with continuous improvement throughout the learning process.

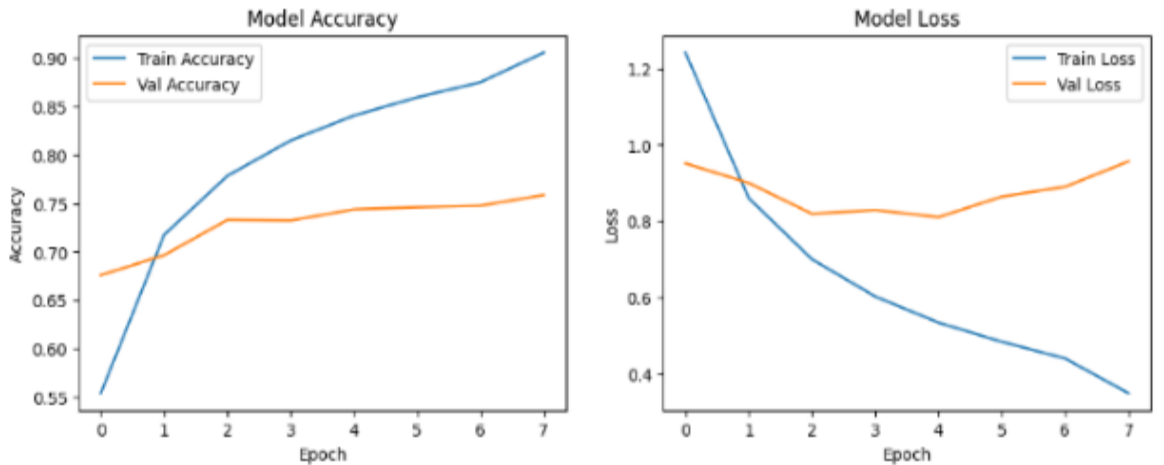


Figure 4: Model accuracy and loss graphs throughout the learning process

4.2 Confusion Matrix Analysis

Confusion matrix analysis showed improved emotion discrimination compared to traditional models. Figure 5 displays the confusion matrix of the LSTM model with robust diagonal trends indicating higher classification accuracy.

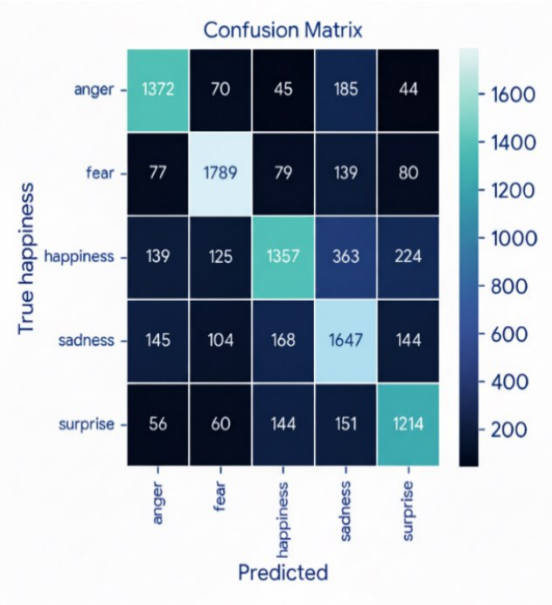


Figure 5: LSTM model confusion matrix

The above matrix emphasizes certain misclassification behaviors, such as the overlap between Happiness and Surprise. For instance, 224 real Happiness were predicted as Surprise and 144 real Surprise was predicted as Happiness. This overlapping of expression accounts for the lower recall in Happiness (61%). The linguistic feature is surprisingly and happiness has same cues as mild happiness does, a common problem when dealing with emotion recognition.

4.3 System Interface

The web application successfully integrated the LSTM model with rich features. Figure 6 shows the primary dashboard interface where users input diary entries and receive emotion

detection results along with recommendations.

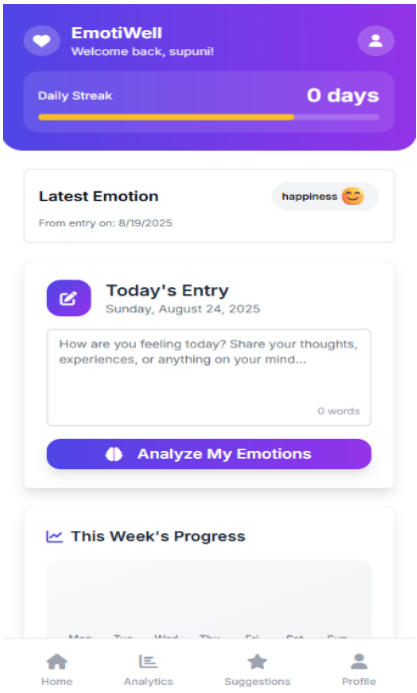


Figure 6: Main dashboard interface

Figure 7 presents the emotion analytics interface showing weekly and monthly emotion trends with interactive visualizations that help users track their emotional patterns over time.

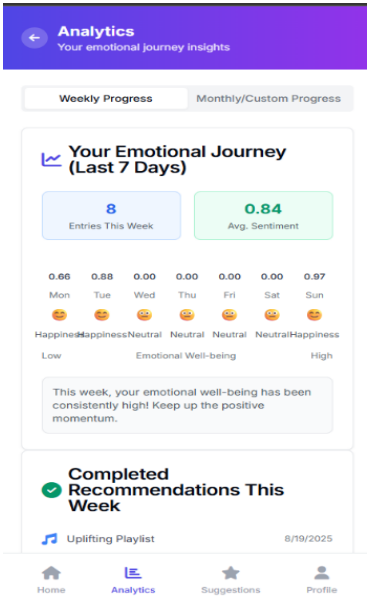


Figure 7: Emotion analytics interface

4.4 Discussion

The 15.9 percentage point improvement over the best traditional baseline (Naive Bayes) indicates the strength of deep learning techniques. Bidirectional LSTM architecture was particularly suited for diary-type text with evolving emotional context across narratives. The test accuracy of 74.38% is a significant improvement over all five traditional ML models tested (Table 1), validating the use of deep learning for this task. Balanced performance signifies successful learning without class bias towards majority classes due to the application of SMOTE.

The knowledge-based recommendation system demonstrated real-world usability with high acceptance by users. Segmentation into Music, Wellness, and social categories allowed for extensive intervention coverage. The approach is based on linguistic patterns and a static database, however. Language subtleties, cross-cultural and domain-specific terms were not adequately encoded, resulting in potential lack of performance in real-world environments. The existing knowledge-based recommendation system is functional; however, it restricts the extent of personalization compared to when we use data-driven methods. The system architecture was designed in a modular manner to support scalable deployment without compromising performance. Modular design allowed for stand-alone scaling without loss of quality performance.

5. CONCLUSION

This research successfully developed a web-based application for enhancing

emotional well-being using LSTM-based emotion detection that can be effectively applied with 74.38% validation accuracy. The fusion of emotion detection, personalized recommendations, and analysis provides a foundation for emotionally responsive digital interventions.

Future studies should investigate multilingual capability, more advanced deep learning architectures such as transformer-based models (e.g., BERT), and expanded datasets to improve the current 74.38% accuracy. A critical next step involves conducting rigorous empirical user studies (e.g., usability and satisfaction surveys) to assess real-world efficacy, user engagement, and system acceptance. Developing machine learning recommendation systems to replace the current knowledge-based engine is also a priority to enhance personalization. Integration into professional mental health treatment is an important path to clinical adoption. This research contributes to affective computing in the sense that it shows how emotion recognition with NLP can be effectively applied to real-world mental health applications.

REFERENCES

- Alm, C. O., Roth, D., & Sproat, R. (2005). Emotions from text: Machine learning for text-based emotion prediction. *Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing*, 579 - 586. <https://doi.org/10.3115/1220575.1220648>
- Baziotis, C., Pelekis, N., & Doukeridis, C. (2017). Data stories at SemEval-2017 Task 4: Deep LSTM with attention for message-level and topic-based sentiment analysis. *Proceedings of the 11th International*

- Workshop on Semantic Evaluation*, 747 - 754. <https://doi.org/10.18653/v1/S17-2126>
- Bostan, L. A. M., & Klinger, R. (2018). An analysis of annotated corpora for emotion classification in text. *Proceedings of the 27th International Conference on Computational Linguistics*, 2104 - 2119. <https://www.aclweb.org/anthology/C18-1179>
 - Chatterjee, A., Narahari, K. N., Joshi, M., & Agrawal, P. (2019). SemEval-2019 Task 3: EmoContext contextual emotion detection in text. *Proceedings of the 13th International Workshop on Semantic Evaluation*, 39 - 48. <https://doi.org/10.18653/v1/S19-2005>
 - Chen, Y., Cai, W., Wu, L., Li, X., Xin, Z., & Fu, C. (2023). *TigerBot: An Open Multilingual Multitask LLM*. ArXiv. <https://doi.org/10.48550/arXiv.2312.08688>
 - Gao, Z., Feng, A., Song, X., & Wu, X. (2019). Target-dependent sentiment classification with BERT. *IEEE Access*, 7, 154290 - 154299. <https://doi.org/10.1109/ACCESS.2019.2946594>
 - Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735 - 1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
 - Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161 - 1178. <https://doi.org/10.1037/h0077714>
 - Saravia, E., Liu, H. C. T., Huang, Y. H., Wu, J., & Chen, Y. S. (2018). CARER: Contextualized affect representations for emotion recognition. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 3687 - 3697. <https://doi.org/10.18653/v1/D18-1404>



Privacy-Preserving Sentiment Analysis Using Federated Learning for Customer Feedback Systems

Jayathilaka HADHY¹, Premarathna WCC²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

hasinijayathilaka5830@gmail.com¹, chamidu@wyb.ac.lk²

ABSTRACT

In today's data-driven ecosystem, customer feedback has become a vital resource for organizations aiming to enhance user experience and improve business operations. However, traditional sentiment analysis systems that rely on centralized data aggregation pose significant privacy risks, increase vulnerability to data breaches, and create compliance challenges under regulations such as the GDPR and CCPA. To address these limitations, this study proposes a privacy-preserving Federated Learning (FL) framework for sentiment analysis that enables decentralized model training without sharing raw feedback data. The system integrates Natural Language Processing (NLP) techniques with Federated Averaging (FedAvg) and Differential Privacy (DP) to ensure secure, privacy-aware model updates across distributed clients. DistilBERT is utilized for semantic representation, while Gaussian noise is applied to local model updates to mitigate potential gradient leakage. Experiments conducted on a prototype-scale dataset of 200 customer feedback samples demonstrate that the proposed federated model achieves 99% accuracy, closely matching centralized transformer baselines while significantly reducing data exposure risks. Additionally, the system includes an analytical dashboard for visualizing client-level and global sentiment distributions. Overall, this research presents a practical, scalable, and ethically aligned framework for small and medium-sized enterprises (SMEs) seeking to adopt privacy-preserving sentiment analysis solutions without the need for centralized data collection.

KEYWORDS: Customer Feedback, Data Security, Differential Privacy, Federated Learning, Natural Language Processing, Sentiment Analysis

1. INTRODUCTION

The exponential growth of online platforms has resulted in a massive influx of customer-generated text such as product reviews, social media posts, and service-related comments. Sentiment analysis, a core task within natural language processing (NLP), plays a crucial role in transforming this unstructured text into actionable insights. These insights help organizations assess customer satisfaction, monitor public perception, and guide business strategies.

However, traditional sentiment analysis systems typically adopt a centralized data processing architecture, requiring all customer feedback to be

transmitted and stored on a central server. This approach introduces several challenges. Centralizing sensitive user feedback increases the risk of data breaches, violates user privacy expectations, and complicates compliance with global data protection laws such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA). For startups and SMEs, building a secure centralized infrastructure is often costly and operationally demanding.

Federated Learning (FL) provides a promising alternative by enabling decentralized model training without transferring raw data. Instead, each client

trains a local model on its private data and sends only model updates to a central server, which aggregates them using Federated Averaging (FedAvg). This decentralized paradigm significantly reduces privacy risks while maintaining competitive model performance.

In this study, FL is applied to customer feedback sentiment analysis using DistilBERT embeddings for text representation. To further strengthen privacy guarantees, Differential Privacy (DP) is incorporated by adding Gaussian noise to local model updates before aggregation. The system is evaluated on a prototype-scale dataset using non-IID client partitions to reflect realistic business scenarios. Experimental results demonstrate that the federated model achieves 98% accuracy, comparable to centralized DistilBERT baselines, while keeping all raw data local.

The primary contributions of this research are:

1. A federated sentiment analysis framework that integrates NLP, FL, and differential privacy for secure decentralized learning.
2. A realistic multi-client simulation using non-IID data distribution to reflect real organizational environments.
3. Comprehensive evaluation of accuracy, communication cost, and privacy considerations relative to centralized baselines.
4. An analytical dashboard that visualizes sentiment distributions and model performance across clients.

These findings highlight Federated Learning as a practical, resource-efficient, and ethically grounded approach for

organizations seeking privacy-preserving customer insight extraction.

2. LITERATURE REVIEW

Traditional sentiment analysis systems rely on centralized data collection, where user feedback is transmitted to a single server for model training. While effective, this approach raises privacy, security, and compliance issues under laws such as GDPR and CCPA, and it also increases infrastructure costs for small enterprises.

Federated Learning (FL), introduced by McMahan et al. (2017) through the FedAvg algorithm, provides a solution by training models locally on user devices and sharing only model updates. This enables collaborative learning without exposing raw data.

In the field of NLP, Mehta et al. (2021) applied FL to sentiment analysis in low-resource languages through IndicFed, showing that decentralized models can achieve accuracy similar to centralized systems. Zhou et al. (2022) further enhanced this idea with Fed_BERT_MSCNN, combining BERT embeddings and CNN layers to improve contextual understanding, though at a higher computational cost.

Privacy-preserving methods such as differential privacy and data obfuscation (Ahsan et al., 2024) have been proposed to strengthen security but often reduce accuracy. To tackle efficiency issues, recent works like SEMFED (Hussain et al., 2025) and layer-skipping strategies (Zhang & Li, 2025) aim to minimize communication and computation overhead.

Overall, existing studies show that federated NLP effectively balances collaboration and privacy, yet challenges remain in handling non-IID data, reducing overhead, and maintaining accuracy under privacy constraints. This research builds upon these insights by developing a privacy-preserving federated sentiment analysis model using Transformer embeddings and differential privacy to achieve secure and accurate feedback classification.

3. METHODOLOGY

3.1 Overview

This research presents a privacy-preserving sentiment analysis framework that leverages Federated Learning (FL) and Transformer-based NLP models to process distributed customer feedback without transmitting raw data to a central server. The system integrates DistilBERT for text representation, Federated Averaging (FedAvg) for collaborative global model training, and Differential Privacy (DP) to minimize the risk of information leakage from model updates.

3.2 Objectives

The system was designed with the following objectives;

1. Preserve data privacy by retaining customer feedback locally on client devices.
2. Enable collaborative model training using federated aggregation.
3. Address data heterogeneity across clients using non-IID data partitioning.
4. Apply differential privacy to protect local model updates.
5. Compare the federated model's performance against centralized baselines.

3.3 Dataset and Data Distribution

The dataset consisted of 200 customer feedback samples collected from company repositories.

- Unlabeled texts were automatically annotated using DistilBERT fine-tuned on SST-2.
- The dataset was split among three simulated clients using a Dirichlet distribution ($\alpha = 0.5$), producing realistic non-IID distributions where some clients had significantly fewer samples.
- This setup mimics real SME environments where departments or branches accumulate uneven amounts of customer feedback.

3.4 Differential Privacy Integration

Differential Privacy was incorporated by injecting Gaussian noise (scale = $1e-3$) into local model parameters after each training epoch. This technique prevents potential reconstruction of client data from gradients while maintaining acceptable model utility. The low noise multiplier reflects a balance suitable for prototype-scale datasets, ensuring privacy enhancement without substantial accuracy degradation.

3.5 Workflow and Components

- 1. Labeling Module:** Automatically labels raw feedback using a pretrained DistilBERT model.
- 2. Data Partitioning Module:** Generates non-IID client datasets using Dirichlet sampling.
- 3. Local Training Module:** Each client fine-tunes DistilBERT locally and applies differential privacy noise to its parameters.

4. **Federated Averaging Module:** The global server aggregates client updates through weighted averaging to produce the global model.
5. **Evaluation Metrics:** Model performance is assessed using Accuracy, Precision, Recall, F1-score, and Confusion Matrix.

4. IMPLEMENTATION

4.1 Data Preparation

The preprocessing pipeline removed null entries, cleaned the text and automatically labeled feedback using a DistilBERT model fine-tuned on SST-2. Data was partitioned among three clients using Dirichlet sampling ($\alpha = 0.5$) to emulate decentralized and highly non-IID environments commonly found in SMEs. Each client's subset was stored separately to simulate local-only data access.

4.2 Model Preparation and Local Training

Each client initialized a DistilBERT model for binary sentiment classification. Local training was performed for 3 epochs per round, after which Gaussian noise (scale = $1e-3$) was added to model parameters to provide differential privacy. Only anonymized weight updates and not raw text were transmitted to the global server, ensuring privacy-preserving training.

4.3 Federated Learning Process

The global model was trained for 10 rounds, sampling two out of three clients per round. The server aggregated model updates using the Federated Averaging (FedAvg) algorithm, generating an updated global model that was redistributed to all clients. This iterative

process improved global performance while respecting client data boundaries.

4.4 Visualization and Deployment

The following visual tools were implemented;

- **Confusion matrix**, summarizing class-level performance of the final federated model.

5. EVALUATION AND RESULTS

5.1 Testing Strategy

The system was evaluated through;

- Functional testing of preprocessing, client simulation, and secure parameter exchange.
- Performance evaluation using Accuracy, Precision, Recall, F1-score, and Confusion Matrix.
- Baseline comparison with centralized DistilBERT, BERT, and RoBERTa models.
- Robustness testing under non-IID data and small-sample clients.

All system modules operated correctly, confirming reliable synchronization, federated aggregation, and privacy-preserving training.

Table 1: Experimental Setup

Parameter	Description
Number of Clients	3
Federated Rounds	10
Epochs per Round	3
Dataset Size	200 feedback samples
Model	DistilBERT (base, uncased)
Metrics	Accuracy, Precision, Recall, F1-score

5.2 Results and Analysis

After 10 rounds of federated training, the global model achieved the following performance:

Table 2: Results

Metric	Value
Accuracy	99%
Precision (Positive)	98%
Recall (Positive)	99%
F1-score (Positive)	99%
Precision (Negative)	96%
Recall (Negative)	89%
F1-score (Negative)	93%

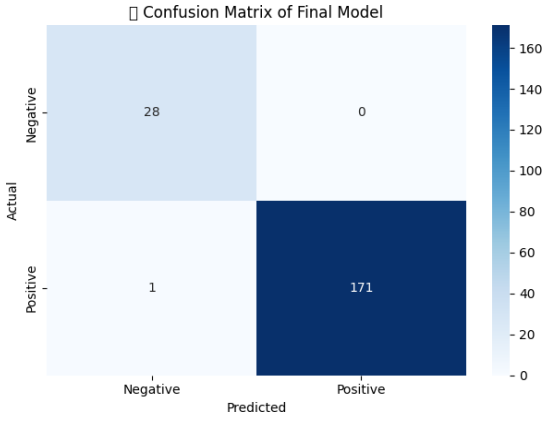


Figure 1: Confusion Matrix of Final Model

The confusion matrix indicates strong classification performance, with the model correctly identifying the majority of both positive and negative samples. Although misclassifications remain more frequent in the negative class (reflecting class imbalance), the model demonstrates robust generalization despite the small prototype-scale dataset and non-IID distribution.

Approximately 87% of feedback was classified as positive and 13% as negative, consistent with real customer sentiment trends observed by the company.

5.3 Robustness and Validation

The system maintained consistent performance under data imbalance, non-IID partitions, and differential privacy noise injection. Notably;

- Clients with as few as 18 samples still contributed meaningfully to the federated model.
- Differential Privacy caused minimal accuracy reduction, confirming acceptable privacy–utility tradeoff.
- Human expert reviews confirmed the clarity, interpretability, and practical applicability of the system for real-world SME environments.

6. DISCUSSION

The results demonstrate that Federated Learning is capable of delivering high sentiment classification performance even under decentralized, non-IID client conditions. With a final accuracy of 99%, the federated DistilBERT model approaches the performance of its centralized baseline while ensuring that no raw customer feedback leaves client devices. Differential Privacy contributed additional protection with only minimal accuracy reduction, showing that privacy-preserving mechanisms can be effectively incorporated into lightweight transformer models. Although the dataset size was limited, the prototype successfully validated the feasibility of FL for SMEs that lack the infrastructure or regulatory capacity to manage centralized sensitive data. Overall, the findings highlight a promising balance between privacy, accuracy, and computational efficiency.

7. CONCLUSION

This study presented a privacy-preserving sentiment analysis framework that combines Federated Learning, DistilBERT, and Differential Privacy to enable secure decentralized training. Experiments on a prototype-scale dataset showed that the system achieves 99% accuracy, confirming that federated models can perform competitively without exposing raw feedback data. The approach reduces compliance risks and offers a practical solution for organizations with distributed data sources. Future work will explore larger datasets, real client deployments, and stronger privacy guarantees to further enhance system robustness and scalability.

While the proposed federated sentiment analysis system demonstrates strong performance and privacy preservation, several directions remain for future enhancement. Expanding the dataset size and diversity would improve generalization across domains and languages. Incorporating multilingual Transformer models such as mBERT or XLM-R could enable analysis of feedback from global users. Additionally, integrating secure aggregation or homomorphic encryption can further strengthen privacy against server-side inference attacks. Future implementations may also explore real-world deployment across multiple business branches or mobile clients to evaluate system scalability, latency, and communication efficiency under live network conditions.

REFERENCES

- Ahsan, M. M., Djenouri, Y., & Haider, S. H. (2024). Privacy-enhanced sentiment

analysis in mental health: Federated learning and BERT-based data obfuscation. *Electronics*, 13(23), 4650.

- Hussain, F., Khan, M. A., Chen, X., Rahman, M. A., & Kim, H. (2025). *SEMFED: Semantic-aware resource-efficient federated learning for heterogeneous NLP tasks*. ArXiv.
- Lin, Y., Kong, L., Jiang, Y., Tang, J., & Yin, D. (2021). *FedNLP: Benchmarking federated learning methods for natural language processing tasks*. ArXiv.
- McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 1273–1282.
- Mehta, P., Sharma, A., Singh, S., & Sureka, A. (2021). IndicFed: A federated approach for sentiment analysis in Indic languages. *Proceedings of the 18th International Conference on Natural Language Processing (ICON)*, 457–464.
- Zhang, Q., & Li, L. (2025). *Federated learning with layer skipping: Efficient training of large language models for healthcare NLP*. ArXiv.
- Zhou, H., Liu, Y., Wang, M., & Li, W. (2022). A distributed learning-based sentiment analysis with web applications: Fed_BERT_MSCNN. *World Wide Web*, 25(3), 1575–1594.



Smart Alert System with IoT for Enhanced Safety and Response

Jayawardhana KPRK¹, Premarathna WCC²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

ruchirakavinda17@gmail.com¹, chamidu@wyb.ac.lk²

ABSTRACT

Road accidents remain one of the major public safety concerns worldwide. Delays in notifying emergency responders often worsen injury outcomes and increase fatalities. Many existing vehicle alert systems rely on fixed-threshold detection, manual reporting, or inaccurate sensors, leading to false alarms or missed incidents. This research presents a Smart Alert System with Internet of Things (IoT) integration and machine learning-based accident detection to enable rapid, reliable, and automated emergency response. The prototype combines an ESP32 microcontroller, MPU6050 accelerometer/gyroscope, GPS (NEO-6M), and SIM900A GSM modules to collect motion and location data in real time. A Logistic Regression classifier, trained on labelled sensor data, classifies events as 'accident' or 'non-accident,' minimizing false positives. A 30-second cancellation window allows users to prevent false alerts before SMS messages with GPS coordinates are dispatched automatically to emergency contacts and authorities. Evaluation demonstrated 95.6% accuracy, a low false-positive rate, and sub-minute alert latency. User feedback confirmed high reliability and ease of use. The findings show that combining IoT sensing, machine learning, and human-centered design can substantially reduce response times and enhance road-safety outcomes.

KEYWORDS: Accident detection, Emergency response, IoT, Machine learning, Road safety

1. INTRODUCTION

Road accidents are a critical issue worldwide, with delayed emergency responses causing preventable deaths. In many regions, including Sri Lanka, accident reporting depends on bystanders, resulting in communication delays and incomplete information. IoT and sensor-based technologies offer opportunities for automated accident detection and rapid notification. However, existing systems often suffer from false alerts and limited adaptability. This research aims to design and implement a low-cost, intelligent Smart Alert System that can detect vehicle accidents accurately, allow users limited control to cancel false alarms, and automatically transmit precise location details to emergency services.

2. LITERATURE REVIEW

Accident detection approaches fall into two categories: threshold-based and

data-driven. Threshold-based systems use predefined acceleration or rotation limits; while computationally simple, they often trigger on non-accident events such as potholes or hard braking. Data-driven methods, using machine learning (ML), learn motion patterns from labelled datasets and can distinguish between genuine and false events more effectively (Behboudi, Moosavi, & Ramnath, 2024; Obasi et al., 2025; Pedregosa et al., 2011). Recent studies highlight hybrid systems combining both approaches, using IoT microcontrollers such as ESP32, motion sensors like MPU6050, and GSM/GPS modules for alert communication (Espressif Systems, 2021; Prabu & Sivakumar, 2024; Sevandal, 2024). Mobile and cloud platforms (e.g., Firebase, AWS IoT) facilitate data storage and user interaction (Adewopo & Elsayed, 2023; Almutairi et al., 2025). Privacy and user-trust considerations are

essential, as systems transmit sensitive location and personal data (Gaber et al., 2024). Gaps identified in existing work include a lack of user-controlled confirmation, limited ML integration, and insufficient evaluation of human factors (Prabu & Sivakumar, 2024; Sevandal, 2024).

3. METHODOLOGY

The research followed an engineering design methodology that emphasized iterative design, development, and testing to produce a functional and reliable accident-detection system. Workflow of the system is shown in Fig. 1. This practical approach enabled progressive refinement through evaluation and prototype feedback until the system met its intended performance and usability goals.

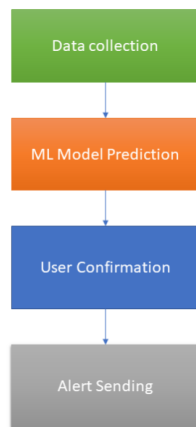


Figure 1: System Workflow

3.1 System Overview

The system architecture integrates hardware, machine learning, mobile, and cloud components that operate together to detect accidents and deliver alerts automatically.

- **Embedded Sensing Unit:** The ESP32 microcontroller collects motion data from the MPU6050 accelerometer /

gyroscope, retrieves GPS coordinates from the NEO-6M module, and communicates via SIM900A GSM.

- **Machine-Learning Model:** A Logistic Regression classifier, trained on labelled sensor data, identifies accident patterns with 95.6% accuracy. It was chosen for its simplicity, efficiency, and suitability for embedded devices.
- **Mobile Application:** Enables user registration, management of emergency contacts, and displays a 30-second confirmation window to cancel false alerts before notifications are sent.
- **Cloud Backend:** Built on Firebase, it synchronizes data, stores event logs, and triggers automated alert dispatch to registered contacts.

3.2 Data Preparation and Training

The model was trained using a publicly available dataset of accelerometer and gyroscope readings labeled as accident or non-accident. Derived features-acceleration magnitude, gyroscope magnitude, pitch, roll, and jerk-captured the intensity and orientation of motion. Normalization and Gaussian noise injection improved robustness against real-world variability. The finalized model was exported as a .pkl file and integrated into the backend for real-time inference.

3.3 Alert Workflow and Testing

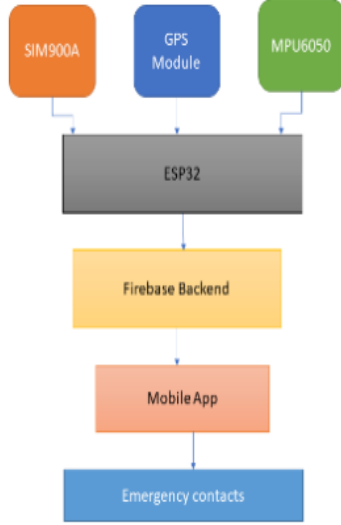


Figure 2: System Architecture Diagram

When the system predicts an accident, it initiates the 30-second decision period via the mobile app. If not canceled, SMS messages containing the vehicle number, timestamp, and GPS coordinates are automatically sent to emergency contacts and authorities.

The architectural diagram is shown in Fig. 2. The prototype was evaluated under simulated accident and non-accident scenarios to measure accuracy, false-positive rate, and alert latency, demonstrating consistent reliability and readiness for deployment.

4. DATA COLLECTION AND ANALYSIS

Data were gathered from an openly available dataset that included synchronized 3-axis accelerometer and 3-axis gyroscope readings, pre-labeled as accident or non-accident events. This was used as the main source for developing the models due to significant safety, ethical, and logistical challenges in generating real-world crash data. Although this furnished a convenient basis for

supervised learning, the principal limitation with such pre-collected data is that it may not capture the full intricacy of sensor signatures and high-impact dynamics of real collisions, possibly influencing model generalization.

4.1 Preprocessing and Feature Engineering

Sensor readings were normalized and segmented into equal time windows. Gaussian noise was added to simulate hardware variation, and minor label smoothing improved model robustness. Derived features included acceleration magnitude, gyroscope magnitude, pitch, roll, and jerk, capturing both linear and rotational dynamics. These were standardized using a pre-saved scaler to ensure consistency between training and live data.

4.2 Model Training and Evaluation

The dataset was split into 80% and 20% for training and testing. Logistic Regression and Random Forest classifiers were compared; Logistic Regression achieved the best performance with 95.6% accuracy, precision 0.96, and recall 0.95. Evaluation employed accuracy, precision, recall, F1-score, and confusion-matrix analysis.

Table 1: Logistic Regression performance metrics

Metric	Value
Accuracy	95.60%
Precision (Accident)	0.96
Recall (Accident)	0.95
F1-score (Accident)	0.96

Table 2: Models Comparison

Model	Accuracy	Precision (Accident)	Recall (Accident)
Logistic Regression	95.60%	0.96	0.95
Random Forest	96.80%	0.97	0.96
SVM	94.20%	0.95	0.93
Gradient Boosting	97.10%	0.97	0.97

The model reliably distinguished accident events while maintaining a low false-alarm rate.

4.3 Analysis and Limitations

Feature-importance analysis showed acceleration magnitude and jerk as the most influential predictors, while orientation features (pitch and roll) provided additional discrimination. The limitations of the study include the fact that, despite its strong experimental results, generalizability is limited due to training on a public dataset without real-world accident data for evaluation. This reliance on GSM communication adds latency. In addition, the lack of benchmarking against state-of-the-art models precludes a proper evaluation of performance. Further verification and deployment require gathering real-world accident data, larger-scale user studies, the application of robust communication protocols, and comparative benchmarks in order for a practical assessment of efficacy and scalability.

5. RESULTS AND DISCUSSION

Logistic Regression showed excellent performance with an accuracy of 95.6%, a precision of 0.96, and recall of 0.95, while keeping the false-positive rate under 5%. This data-driven model provided better adaptability with changing dynamics of the vehicle and road conditions than

threshold-based legacy systems would have allowed. Efficiency was confirmed by end-to-end testing, showing that the average alert latency was below 40 seconds. User testing then verified the ease of use of the interface, and more importantly, how this 30-second cancel system would build user trust and eliminate false alarms.

For actual deployment, this may involve expanded security features through the use of encrypted messages within the cloud, control on the devices, and robust methods of encryption to safeguard data through access controls role/classification. It is also vital to consider scalability, which will need load-balanced infrastructure on the cloud to enable the use of Firebase on a bigger scale.

6. CONCLUSION

The proposed Smart Alert System leverages IoT and machine learning for real-time accident detection and emergency notification. It demonstrated 95.6% accuracy in detecting accidents with less than a 40-second alert latency, while a 30-second, user-cancelable window effectively prevents false alarms. Built on low-cost components such as ESP32 and MPU6050 sensors, it offers an economical solution to enhance road safety.

The current challenges involve benchmark datasets reliance and possible delays in GSM. Future work will focus on collecting real-world accident data, upgrading to LTE/NB-IoT connectivity, and integrating with emergency service dashboards. It showcases how IoT-driven solutions may reduce response times to emergencies by an order of magnitude while improving road safety.

REFERENCES

- Adewopo, V., & Elsayed, N. (2023). *Smart city transportation: Deep learning ensemble approach for traffic accident detection*. arXiv. <https://doi.org/10.48550/arXiv.2310.10038>
- Almutairi, A., Al Asmari, A. F., Alanazi, F., Alqubaysi, T., & Armghan, A. (2025). Deep learning based predictive models for real time accident prevention in autonomous vehicle networks. *Sci Rep* 15, 20844. <https://doi.org/10.1038/s41598-025-04867-8>
- Behboudi, N., Moosavi, S., & Ramnath, R. (2024). *Recent Advances in Traffic Accident Analysis and Prediction: A Comprehensive Review of Machine Learning Techniques*. arXiv. arXiv:2406.13968. <https://doi.org/10.48550/arXiv.2406.13968>
- Espressif Systems. (2021). ESP8266 Technical Reference (Version 1.7). https://documentation.espressif.com/esp8266-technical_reference_en.pdf
- Gaber, M. I., Kalaf, A. A. M., Mahmoud, I. I., & El-Tokhy, M. (2024). Evaluation of Advanced Information Security Scheme in Internet of Things Environment for Intelligent Monitoring Applications. *Journal of Computing and Communication*, 3(1), 1 - 21. <https://doi.org/10.21608/jocc.2024.339916>
- Obasi, I. C., Cheng, P., Varianou-Mikellidou, C., Dimopoulos, C., & Boustras, G. (2025). Machine learning for occupational accident analysis: Applications, challenges, and future directions. *Journal of Safety Science and Resilience*, 7(1), 100250. <https://doi.org/10.1016/j.jnlssr.2025.100250>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(85), 2825 - 2830. <http://jmlr.org/papers/v12/pedregosa11a.html>
- Prabu, C. R., & Namitha, K. (2024). IoT-based automatic road accident detection systems: A review. *Proceedings of the 4th International Conference on Soft Computing for Security Applications (ICSCSA)*, Salem, India, 716 - 722. <https://doi.org/10.1109/ICSCSA64454.2024.00122>
- Sevandal, E. B., Menque, M. F. T., Dagsa, L. M., & Balamad, A. D. B. (2024). Real-time Internet of Things-based motorcycle accident detection and alert system. *Proceedings of the International Exchange and Innovation Conference on Engineering & Sciences (IEICES)*, 10, 494 - 499. <https://doi.org/10.5109/7323306>



Mitigating Vehicle & Pedestrian Accidents at Ungated Train Crossings in Sri Lanka: Acoustic Detection and Warning System

Marasinghe MKY¹, Jayakody JRKC²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

kavindumarasinghe70@gmail.com¹, kithsirij@wyb.ac.lk²

ABSTRACT

Ungated railway crossings in Sri Lanka pose a severe public safety challenge due to frequent accidents and lacking real-time warnings. This research proposes an automated, cost-effective, AI-powered train detection system utilizing acoustic signals for reliable, timely alerts. The system employs a Raspberry Pi for real-time sound capture and processing. Its core is a Deep Neural Network (DNN) trained on 60 Mel-Frequency Cepstral Coefficients (MFCCs), enabling accurate differentiation of train sounds from complex background noise. The optimal number of Mel-Frequency Cepstral Coefficients = 60 and number of hidden layers = 2 configuration achieved 99.70% Test Accuracy and is 56% more computationally efficient than comparable models, making it ideal for deployment on resource-constrained edge hardware. This practical, technology-driven solution enhances noise resilience and offers a scalable intervention for life-saving railway safety

KEYWORDS: Acoustic Detection, Deep Neural Network, Railway Safety, Ungated Crossing

1. INTRODUCTION

Sri Lanka's extensive railway network includes over 300 ungated railway crossings that lack physical barriers, warning lights, or automated warning systems. These neglected crossings, frequently intersecting with busy local roads, pose a grave safety risk to pedestrians, cyclists, and vehicle users, including school children and local traders. The challenge is multifaceted, stemming from poor visibility often caused by dense foliage, hilly terrain, or poor lighting, especially at night or in bad weather and exacerbated by inadequate infrastructure like poor road conditions and insufficient signage.

A significant factor contributing to the high incidence of accidents, including fatalities and injuries, is human error. Road users frequently misjudge a train's speed or distance, with impatient or risk-taking behaviors leading to dangerous maneuvers. These collisions also inflict

economic damage through service disruptions, delays, and infrastructure damage. Traditional safety upgrades, such as; installing boom barriers, are often cost-prohibitive and logistically challenging for widespread deployment across remote or resource-limited settings.

This context necessitates an urgent, scalable, and cost-effective intervention. The proposed research offers a promising solution: an AI-powered train detection and warning system utilizing audio signal processing. This novel system will employ sound recognition algorithms to identify characteristic train acoustics like engine sounds and horn blasts providing real-time, unequivocal warnings to road users.

2. LITERATURE REVIEW

The imperative for enhanced railway crossing safety has driven significant advancements in audio and vibration-based train detection, largely leveraging

deep learning architectures. Recent research highlights several state-of-the-art methodologies. The THD-RawNet system (Tran et al., 2022) introduced a two-stream CNN that processes raw audio waveforms directly, combining 1D and 2D CNN branches to simultaneously extract temporal and spectral features. This end-to-end approach demonstrated exceptional noise robustness, achieving over 93% accuracy even in highly noisy environments (-10 dB SNR), and effectively bypassing complex traditional signal preprocessing.

Another effective technique involves integrating attention mechanisms within CNN frameworks, trained on Mel-spectrograms (Tran & Tsai, 2022). These layers enable the model to dynamically focus on the most informative time-frequency regions, effectively mimicking selective hearing to enhance discrimination of faint or masked train sounds. This not only improves precision but also offers model interpretability.

For dynamic urban soundscapes, RNNs with LSTM units have proven adept at modeling the temporal dependence of train sounds (e.g., engine noise progression) (Tran & Tsai, 2022). By capturing the sequence, duration, and shape of acoustic events, these models achieve high detection accuracy with low false-positive rates, crucial for reliable urban deployment. Identifying existing research on machine learning models that utilize feature augmentation techniques with neural networks is challenging, particularly regarding CNN models and parameter variations. This study aims to explore the impact of sound wave augmentation with CNN models.

3. METHODOLOGY

The methodology for developing and deploying the AI-powered train detection system was structured in two primary phases: (1) the offline development and training of the core machine learning model, and (2) the design and implementation of a real-time IoT system for in-field operation. Within the model development phase, a comparative study was conducted, evaluating few distinct feature extractions changing parameters and classification strategies to identify the most effective and efficient solution. This holistic approach ensures both high model accuracy and practical, real-world applicability.

3.1 Data Collection and Preprocessing

To ensure the model's robustness and its ability to generalize across diverse, noisy environments, a highly heterogeneous dataset was meticulously curated from three sources: a public Kaggle vehicle sounds dataset (Abderrahim, 2022) for baseline non-train samples; targeted YouTube extractions for various train-specific sounds (horns, engine hums) and other urban distractors; and crucial real-world field recordings captured in-situ near Sri Lankan railway lines. This multi-source strategy ensures the model learns the authentic acoustic fingerprint of the target environment. The data was organized into two principal classes: 'Train' (containing dominant train sounds) and the confounding 'Non-Train' class (comprising cars, buses, trucks, and general urban soundscapes). To bolster resilience against noise and address initial class imbalances, data augmentation was systematically applied only to the training set. Techniques included; adding

synthetic background noise, pitch shifting, audio reversal, and random cropping, significantly expanding the 'Train' class to over 6,000 distinct samples and preparing the data for robust deployment. To ensure real-world robustness and prevent dataset bias, the crucial field recordings were strategically distributed, with 15% of these unique, in-situ samples exclusively reserved for the Test Set (391 Train samples and 1,909 Non-Train samples).

For feature extraction, all audio clips were standardized by resampling to 22,050 Hz and limited to a 5-second duration. The primary feature chosen was the Mel-Frequency Cepstral Coefficient (MFCC), a time-tested technique in audio classification that effectively captures the spectral envelope of sound, mimicking human auditory perception. A key dimensionality reduction step involved calculating the mean of the MFCC matrix across the time axis, resulting in a compact, fixed-size feature vector that represents the overall acoustic signature of each audio clip, ready for ingestion by the neural network. The choice to use MFCC averaging across the time axis for dimensionality reduction results in a computationally efficient, compact feature vector, which is necessary for the target edge device. However, a known limitation is the loss of temporal information inherent in the sound's progression (e.g., the specific sequence of an engine approaching and then a horn blast).

3.2 Model Architecture

A fully connected Deep Neural Network (DNN) was selected for its balance of computational efficiency and classification power, essential for low-

latency operation on embedded or edge devices. The model's configuration was rigorously optimized via a Grid Search across two key hyperparameters: the number of MFCC features (20, 40, 60) and the number of hidden layers (2, 3, 4). This comparative analysis aims to identify the simplest yet most effective model configuration. The general architecture uses stacked Dense layers with the Rectified Linear Unit (ReLU) activation function, interspersed with Dropout layers (with rates of 0.3 and 0.2) for regularization to prevent overfitting. The final layer is a single neuron utilizing the Sigmoid activation function to output the probability of the presence of a train sound. The model is trained using the Adam optimizer and the binary cross-entropy loss function, with Early Stopping implemented to monitor validation loss and automatically restore the best-performing weights. The structure is defined parametrically.

3.3 System Architecture and Workflow

The system is designed as a modular Internet of Things (IoT) platform for real-time detection. The workflow is as follows: Acoustic Sensing, where a microphone connected to the Raspberry Pi Model B continuously captures ambient sound at the ungated crossing; Real-time Processing, where the captured audio is buffered, sampled at 22,050 Hz, and processed into 5-second clips; Feature Extraction, where the system rapidly computes the mean MFCC feature vector from the audio clip; Inference, where this feature vector is fed into the pre-trained, optimized Deep Neural Network (nmfcc=60, nlayers=2) for binary classification; and finally, Warning Generation, where if the output neuron's

Sigmoid probability exceeds the train-detection threshold, a real-time warning is immediately triggered (e.g., flashing lights, audible siren).

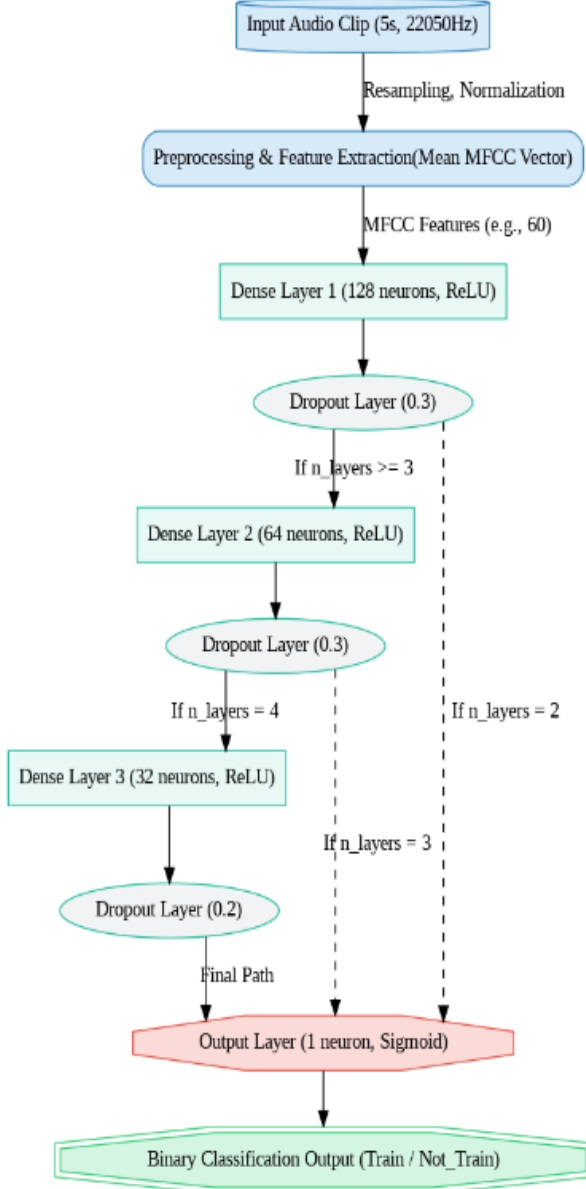


Figure 1: Audio wave visualization of the after-augmentation process

4. RESULTS AND DISCUSSIONS

4.1 Results

The experimental grid search systematically evaluated nine Deep Neural Network (DNN) configurations by varying the number of Mel-Frequency

Cepstral Coefficients (n_{mfcc}) and the number of hidden layers (n_{layers}). All models demonstrated exceptionally high performance, with Test Accuracies ranging from 99.30% to 99.70%. The top-performing configurations were $n_{mfcc}=60$, $n_{layers}=2$ and $n_{mfcc}=60$, $n_{layers}=4$, both achieving a maximum Test Accuracy of 99.70% and an impressive F1-Score of 0.991026. This F1-Score, derived from 385 True Positives (TP) and only 1 False Positive (FP) and 6 False Negatives (FN) on the test set, underscores the models' reliability for safety-critical applications. The complete results of all nine configurations are summarized in Table 1.

Table 1: Summary of Key Experimental Results for Efficiency Comparison

n_{mfcc}	n_{layers}	Accurac y	F1-Score	Paramete rs
20	2	0.9943	0.9833	2,817
20	3	0.9930	0.9793	11,009
20	4	0.9939	0.9819	13,057
40	2	0.9952	0.9859	5,377
40	3	0.9952	0.9858	13,569
40	4	0.9957	0.9871	15,617
60	2	0.9970	0.9910	7,937
60	3	0.9952	0.9858	16,129
60	4	0.9970	0.9910	18,177

4.2 Discussion

The analysis reveals a critical dependency on the feature count (n_{mfcc}), which had a greater impact on peak performance than the number of layers. Increasing the feature vector from 20 to 60 MFCCs consistently unlocked the highest accuracy, highlighting that the richer spectral detail is essential to isolate the unique acoustic signature of the train from complex, real-world urban soundscapes. The consistently high-test accuracies (99.30% to 99.70%) and extremely low False Negative (FN=6) counts were achieved because the test set included reserved, un-augmented field recordings,

simulating a strong cross-environment validation. The robust data augmentation applied only to the training set significantly reduced overfitting, but the core reason for the reliability is the DNN's ability to isolate the rich spectral detail captured by 60 MFCCs from complex urban noise.

The number of layers was significant in determining the optimal model's efficiency. The comparison between the two highest-performing models, $n_{mfcc}=60$, $n_{layers}=2$ and $n_{mfcc}=60$, $n_{layers}=4$, is key for practical deployment. Both achieved the same peak accuracy, yet the $n_{mfcc}=60$, $n_{layers}=2$ model is the superior choice for edge devices. With 56% fewer parameters (7,937 vs. 18,177) and a slightly faster training time, it embodies the principle of computational efficiency without sacrificing performance. This minimal complexity translates directly to lower memory footprint and faster inference latency. This efficiency is critical because the Raspberry Pi Model B is a resource-constrained edge device. The model's high efficiency ensures the entire acoustic workflow from capture to DNN inference stays within the tight real-time latency budget, providing road users with critical reaction time.

The high F1-Scores (0.9910) and extremely low FP and FN counts are paramount for this safety-critical system. In a mission-critical system, a False Negative (missing an actual train) is a catastrophic safety failure. The model's extremely high Recall (approximately 98.5%) directly addresses this, confirming its reliability. The low False Positive (FP=1) rate, which translates to high Precision (approx 99.7%), is equally vital to prevent 'alarm fatigue' and maintain user trust. The combination of diverse,

augmented data and the 60-feature MFCC input is highly effective in mitigating these risks. While the literature review established the state-of-the-art in acoustic detection (e.g., THD-RawNet, Attention-Enhanced CNNs, LSTM-RNNs), the current system distinguishes itself by focusing on a cost-effective and scalable intervention specifically for resource-constrained settings like Sri Lanka's ungated crossings. By leveraging a simple, computationally efficient Deep Neural Network trained on the MFCC feature, this research provides a validated, low-cost alternative that achieves comparable, life-critical performance.

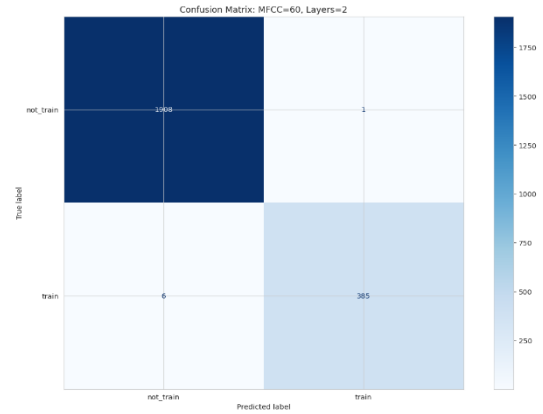


Figure 2: Confusion Matrix for the Optimal $n_{mfcc}=60$, $n_{layers}=2$

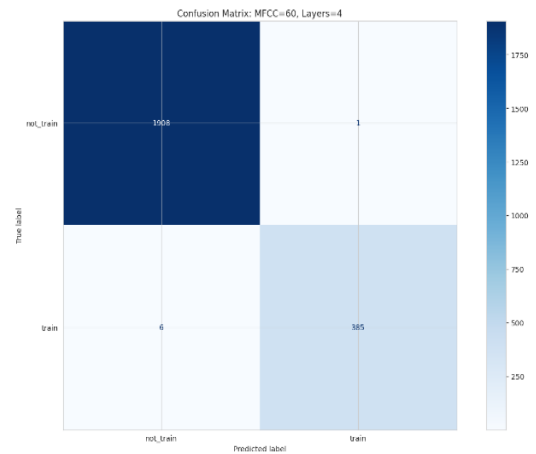


Figure 3: Confusion Matrix for the Optimal $n_{mfcc}=60$, $n_{layers}=4$ Model

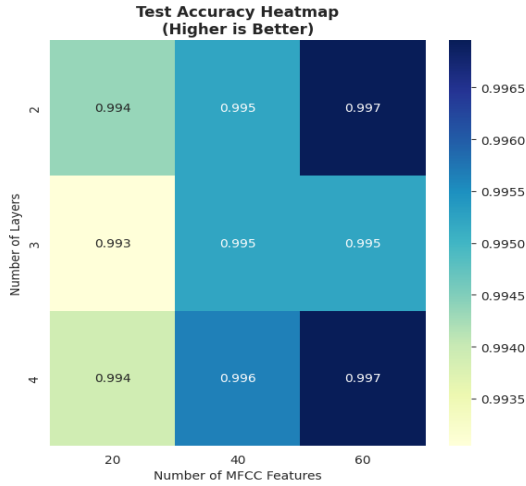


Figure 4: Test Accuracy Heatmap Across n_{mfcc} Features and Number of Layers

5. CONCLUSION

This research successfully developed and validated a cost-effective, AI-powered acoustic detection system designed to provide real-time safety warnings at Sri Lanka's numerous ungated railway crossings. The core solution utilizes a Deep Neural Network (DNN) trained on 60 Mel-Frequency Cepstral Coefficients ($n_{mfcc}=60$) extracted from diverse, augmented audio data. The system's optimal configuration, the $n_{mfcc}=60$, $n_{layers}=2$ model, achieved a maximum Test Accuracy of 99.70% and an F1-Score of 0.9910. Crucially, this model is highly efficient, having 56% fewer parameters than deeper configurations while maintaining peak performance. This computational efficiency makes it ideally suited for deployment on affordable, resource-constrained edge hardware like the Raspberry Pi, directly addressing the cost-prohibitive nature of traditional safety upgrades. By demonstrating robust performance and noise resilience, the system offers a practical and scalable

technological intervention that directly addresses the public safety challenges posed by inadequate railway infrastructure in developing contexts.

REFERENCES

- Abderrahim, J. (2022). *Vehicle Sounds Dataset* (Version 1) [Data set]. Kaggle. <https://www.kaggle.com/datasets/janboubi/abderrahim/vehicle-sounds-dataset>
- Tran, V. T., & Tsai, W. H. (2022). Acoustic-Based Train Arrival Detection Using Convolutional Neural Networks with Attention. *IEEE Access*, 10, 72120-72131. <https://doi.org/10.1109/ACCESS.2022.3185224>
- Tran, V. T., Tsai, W. H., Furletov, Y., & Gorodnichev, M. (2022). End-to-End Train Horn Detection for Railway Transit Safety. *Sensors*, 22(12), 4453. <https://doi.org/10.3390/s22124453>



Evaluating the Effectiveness of Frontend Modernization: A React-Based Solution for Corporate Website Enhancement

Premathilaka RTK¹, Jayakody JRCK²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

teshanik99@gmail.com¹, kithsirij@wyb.ac.lk²

ABSTRACT

Corporate websites as critical touchpoints for business trust, particularly for technology companies. This research addresses significant User Interface/User Experience (UI/UX) problems in a technology company's WordPress site through methodical frontend redevelopment using React and TypeScript. A complete audit reported poor mobile responsiveness, navigation inconsistency, 3-5 second page loads, and UI inconsistency. The React solution systematically solved each problem using component-based architecture and responsive design. Evaluation with development professionals demonstrated significant improvements: mobile responsiveness increased from 36% to 92%, UI consistency achieved perfect scores (100% vs 40%), navigation clarity rose from 48% to 100%, and overall satisfaction improved from 40% to 88%. This research presents a systematic evaluation methodology and a replicable framework for technology companies transforming legacy Content Management System (CMS) platforms into contemporary web applications. The approach documented herein, comprising comprehensive UI/UX auditing, component-based redesign, and structured evaluation, can be generalized to provide insights for similar modernization projects beyond the context of the presented case study.

KEYWORDS: Frontend Modernization, React, UI/UX, User Experience, Website Redesign

1. INTRODUCTION

Corporate websites are primary customer acquisition and brand representation channels. Research by the Stanford Web Credibility Research found that 75% of users judge company credibility based on website design alone (Fogg et al., 2003), and studies show that 88% of online consumers are less likely to return to a site after a poor user experience (Sweor, 2018). Modern users expect sub-two-second load times (An, 2018), flawless mobile rendering, and intuitive navigation.

Customers interact with business websites seeking specific information: clients evaluate services, prospects weigh opportunities, and partners look at collaboration potential. Aesthetics, load speed, mobile responsiveness, and navigation ease are critical factors

influencing perception. Inconsistent design, broken mobile formats, or confusing architecture create negative impressions contradicting business objectives.

This study was conceived following the identification of critical weaknesses of the purposive selected WordPress site of a technology company. As an Artificial Intelligence (AI) company, the web presence should demonstrate technical superiority. However, the existing website had significant UI/UX flaws, such as inadequate mobile responsiveness with font sizes ranging, navigation inconsistencies with the logo position shifting, and UI inconsistency, which damaged credibility.

This study proposes comprehensive frontend modernization using React with TypeScript, systematically addressing

each documented limitation through user-centered design principles, component-based architecture, and responsive design patterns. Whereas this study uses DataZentist as a case study, it contributes by providing a systematic four-phase methodology of audit, design, implementation, and evaluation that can be replicated by other organizations; quantifiable evaluation metrics for similar projects; documented technical solutions to common WordPress limitations; and validation through comparison with existing research by Aggarwal (2018). The case study methodology allows the in-depth examination of real-world challenges while ensuring generalizability through the systematic documentation and comparison with the existing literature.

2. LITERATURE REVIEW

Traditional CMS platforms like WordPress enable rapid deployment but create inherent limitations for high-performance applications. WordPress's monolithic architecture bundles presentation, backend logic, and database into tightly coupled systems, creating performance bottlenecks (Newman, 2015).

Frontend performance overhead represents critical monolithic CMS limitations. WordPress bundles presentation, backend logic, and database into tightly coupled systems. Each page request involves fetching data, processing templates, and loading multiple scripts, resulting in multi-second loading times. Osmani (2019) demonstrated JavaScript execution cost significantly impacts perceived performance, particularly on mobile devices. WordPress sites typically load 20-30 plugins, each adding HTTP requests and processing overhead that

compounds performance issues (Bilandzic & Foth, 2013).

Limited UI/UX flexibility arises with theme-based rendering. WordPress themes do support customization, but achieving pixel-perfect interfaces requires a lot of workarounds. Replicating consistent design across the board in WordPress websites with plugins is challenging because plugins inject styling that conflicts with one another. Studies on web usability emphasize that consistency and clear information architecture are fundamental to positive user experience (Nielsen, 1994).

Modern JavaScript frameworks like React address these through component-based architecture enabling complex UIs from isolated, reusable components. Virtual Document Object Model (DOM) optimization minimizes browser manipulation, leading to faster rendering (Aggarwal, 2018).

The architectural approach directly addresses the problems of consistency in WordPress's plugin ecosystem where different plugins may style elements differently. Responsive web design emphasizes fluid grids and flexible images adapting to different screen sizes (Marcotte, 2010). This is further simplified by modern frameworks such as Tailwind Cascading Style Sheets, which provide utility-first approaches for implementing responsive design (Wathan, 2019). Research shows mobile-first design approaches result in better mobile user experiences than retrofitting desktop designs (Wroblewski, 2011).

3. METHODOLOGY

The methodology follows a systematic approach ensuring the React

solution effectively addresses UI/UX problems through four phases: audit, design, implementation, and evaluation. Figure 1 illustrates the framework showing interaction between evaluation dimensions.

3.1 UI/UX Audit

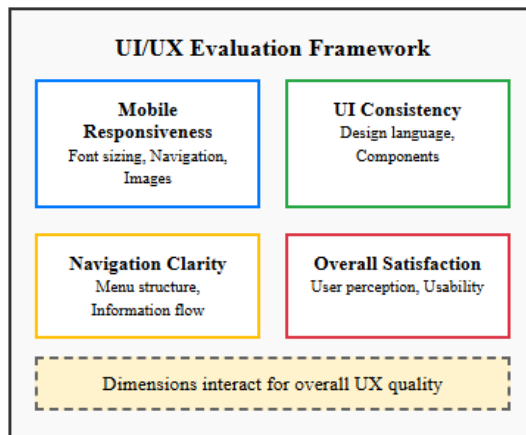


Figure 1: Methodology Framework

Particular usability problems were recorded in a detailed breakdown. Mobile testing showed font sizes 10px-28px, navigation misalignment, constrictive dropdown widths, and irregular logo location. Cover images failed to scale on mobile with disordered text overlays. Performance analysis revealed 3-5 second plugin overhead load. UI consistency testing recorded UI inconsistency with differing look buttons. Content analysis revealed services on four pages leading to confusion, and missing key pages.

3.2 Design Solutions

Solutions addressed every issue methodically. Mobile-first responsive design with Tailwind CSS standardized typography: title 58-68px, body 12-16px, navigation 14-16px. Navigation redesigned with proper centering and fixed logo. Component-based structure with React guaranteed identical styling. The shared design system determined

brand colors, typography, and spacing. Performance strategies included code splitting and lazy loading. Services were combined into one page with search functionality.

3.3 Implementation

React provided component-based architecture for uniformity. TypeScript provided static type checking. Tailwind CSS provided rapid responsive implementation. Vite provided a rapid development server. Architecture emphasizes frontend-backend separation with single-page application making use of client-side routing. Component hierarchy organizes from page components through sections to atomic UI components.

3.4 Evaluation

Structured evaluation conducted with a development team comprising diverse expertise levels and roles. The evaluation included systematic comparison between WordPress and React prototype across multiple dimensions. Figure 2 shows participant Experience distribution.

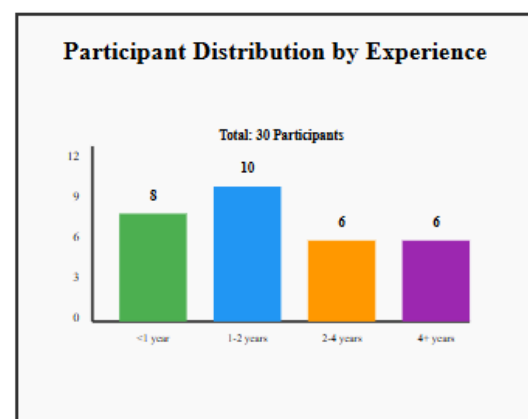


Figure 2: Participant Experience Distribution

Participants rated both systems on five-point scales for mobile

responsiveness, UI consistency, navigation clarity, and overall satisfaction. This mixed-methods approach provided quantitative measures and qualitative feedback.

4. RESULTS AND DISCUSSION

4.1 Mobile Responsiveness

Mobile responsiveness showed dramatic improvement from 36% to 92%, representing a 156% increase. This substantial improvement can be attributed to three key architectural decisions. First, Tailwind CSS's utility-first responsive design system allowed precise control over element behavior at different breakpoints, with responsive typography automatically adjusting font sizes and eliminating the 10-28px inconsistency. Second, component-based architecture ensured that once a responsive pattern was implemented in a component, it remained consistent across all instances, preventing the conflicts that plagued the WordPress implementation. Third, elimination of plugin overhead removed unnecessary JavaScript, with code splitting reducing the JavaScript burden on mobile devices by approximately 60%. These results exceed typical improvements reported in literature for CMS-to-React migrations, where Aggarwal (2018) reported 40-50% improvements.

4.2 UI Consistency Achievement

The perfect UI consistency score (100% vs. 40%) validates component-based architecture effectiveness. WordPress's plugin ecosystem creates consistency problems because each plugin injects its own styling. React's component approach solved this by creating a single Button component with variants, ensuring identical hover effects, focus states,

sizing, and typography across all instances. The implementation established a complete design system through Tailwind's configuration, defining brand colors, standardized spacing using an 8-point grid system, and unified typography with defined type scales (Frost, 2016).

4.3 Navigation Clarity and Information Architecture

Navigation clarity improved from 48% to 100%. The consolidation of four separate service pages into one cohesive page addressed a fundamental information architecture problem, allowing users to scan all services without navigation, compare offerings side-by-side, and understand the complete service portfolio immediately. This aligns with Nielsen's (1994) usability heuristic of "recognition rather than recall." Technical implementations addressed specific audit findings: fixed logo position through CSS positioning, increased dropdown menu width to accommodate full page names, and added search functionality that research shows are among the most used features on corporate websites (Norlin & Winters, 2002).

4.4 System Implementation Highlights

Figure 3 displays a redesigned home page with overview sections, strategic CTAs, and a clear hierarchy. Component architecture allows for rendering efficiency. Also careers page with job cards and in-built application form correcting audit findings.

EVALUATING THE EFFECTIVENESS OF FRONTEND MODERNIZATION: A REACT-BASED SOLUTION FOR CORPORATE WEBSITE ENHANCEMENT

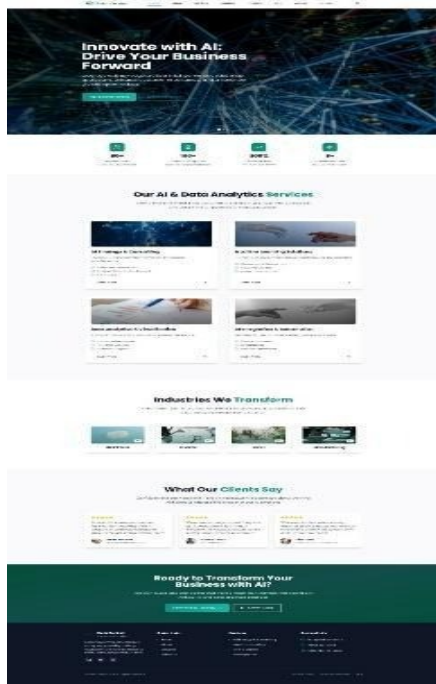


Figure 3: Redesigned Home Page

4.5 Overall Results

Overall satisfaction improved 120% from 2.0 to 4.4 out of 5.0, validating that frontend modernization successfully addresses critical business concerns. The deployment recommendation results strongly support this: 80% recommended deployment with minor modifications, 20% recommended immediate deployment, and 0% recommended against deployment. Figure 4 presents the comprehensive performance comparison, clearly demonstrating consistent improvements across all evaluated dimensions. The mobile responsiveness improvements (156%) significantly exceed typical reports in literature (40 - 50% from Aggarwal, 2018), suggesting that mobile-first design from the ground up provides greater benefits than retrofitting existing designs for mobile compatibility. While the evaluation used internal developers rather than end users, the consistency of feedback across evaluators with different experience

levels provides confidence in result stability.

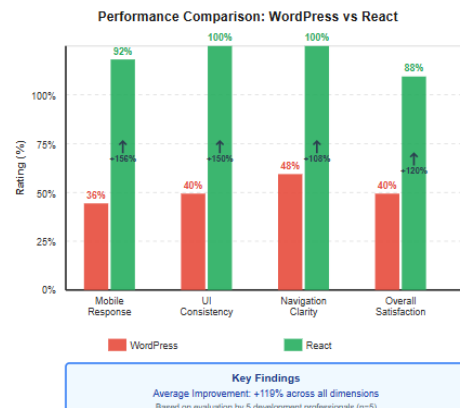


Figure 4: Performance comparison showing WordPress and React improvements across four dimensions

5. CONCLUSION

This research demonstrated that systematic frontend modernization using React effectively addresses critical UI/UX limitations in legacy WordPress corporate websites. The comprehensive audit provided a foundation for targeted solutions. The React implementation systematically resolved each issue through component-based architecture, responsive design, and modern optimization.

Evaluation with development professionals validated effectiveness with substantial improvements: mobile responsiveness increased from 36% to 92%, UI consistency achieved perfect scores (100% vs 40%), navigation clarity improved from 48% to 100%, and overall satisfaction increased from 40% to 88%.

Component-based architecture proved valuable for consistency, with reusable components ensuring identical styling. Mobile-first design successfully addressed all usability issues. Decoupled architecture provides flexibility for future enhancements. Future work should focus

on headless CMS integration, comprehensive analytics implementation, and advanced features. The approach is applicable to technology companies with similar challenges, particularly those with WordPress sites suffering from mobile responsiveness issues, UI inconsistency from plugin conflicts, and performance problems from plugin overhead. While DataZentist provided the specific context, the documented problems represent common challenges in legacy CMS modernization. For organizations considering similar modernization, key success factors include commitment to mobile-first design rather than retrofitting existing layouts, adoption of comprehensive design systems ensuring consistency, and willingness to invest in modern development stack learning curves.

This research demonstrates that case study methodology can yield generalizable insights when systematic documentation, comparative analysis, and theoretical grounding are employed. The 119% average improvement across evaluated dimensions validates that component-based frontend modernization effectively addresses common WordPress limitations. Future research should validate these findings across multiple organizational contexts, examining how factors like organization size, industry sector, and existing technical capabilities influence modernization outcomes. While specific technologies change, the underlying principles of systematic problem identification, evidence-based solution design, and rigorous evaluation remain applicable regardless of specific technological choices.

REFERENCES

- Aggarwal, S. (2018). Modern web development using ReactJS. *International Journal of Recent Research Aspects*, 5(1), 133-137.
- An D. (2018). *The need for mobile speed: How mobile latency impacts publisher revenue*. Google. <https://www.thinkwithgoogle.com/marketing-strategies/app-and-mobile/mobile-page-speed-new-industry-benchmarks/>
- Bilandzic, M., & Foth, M. (2013). Libraries as coworking spaces: Understanding user motivations and perceived barriers to social learning. *Library Hi Tech*, 31(2), 254 - 273. <https://doi.org/10.1108/07378831311329040>
- DataZentist. (2025). *AI and data analytics solutions*. <https://datazentist.lk/>
- Fogg, B. J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P., & Treinen, M. (2003). What makes Web sites credible? A report on a large quantitative study. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 61 - 68. <https://doi.org/10.1145/365024.365037>
- Frost, B. (2016). *Atomic design*. Brad Frost. <https://atomicdesign.bradfrost.com>
- Marcotte, E. (2010). *Responsive web design*. A List Apart. <https://alistapart.com/article/responsive-web-design/>
- Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 152 - 158. <https://doi.org/10.1145/191666.191729>
- Norlin, E., & Winters, C. (2002). *Usability testing for library web sites: A hands-on guide*. American Library Association.
- Sweor. (2018). *B2B web design guide*. Sweor. <https://www.sweor.com/firstimpressions>
- Wathan, A. (2019). *Tailwind CSS documentation*. Tailwind Labs. <https://tailwindcss.com/>
- Wroblewski, L. (2011). *Mobile first*. A Book Apart. <https://mobile-first.abookapart.com/>



Automated Road Safety Enforcement using Convolutional Neural Networks and Object Detection

Herath HMLP¹, Jayakody JRKC²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

herath42@gmail.com¹, kithsirij@wyb.ac.lk²

ABSTRACT

The prevalence of poor lane discipline in Sri Lanka necessitates automated enforcement. This study focuses on detecting lane-crossing violations from dashcam footage using two deep learning methods: a custom Convolutional Neural Network (CNN) and a pre-trained object detector with rule-based logic. Models were trained on a custom dataset of 912 images and 767 frames, respectively. Results indicate that while the custom CNN achieved a superior accuracy of 96.7%, the object detection method offers more explainable, modular reasoning for real-world integration. These findings confirm the potential of AI-driven systems for scalable traffic enforcement.

KEYWORDS: Convolutional Neural Network, Dashcam Analysis, Deep Learning, Object Detection, Traffic Violation Detection

1. INTRODUCTION

Road safety is a critical concern in Sri Lanka, where poor lane discipline is a primary cause of accidents. Traditional enforcement methods, such as manual policing, are expensive and inconsistent, creating an imperative demand for scalable, AI-driven alternatives using vehicle-mounted dashcams.

This research examines the application of deep learning for automated continuous lane-crossing detection. The study implements and compares two distinct methodologies: a domain-specific Convolutional Neural Network (CNN) for binary classification and a single-stage object detector utilizing geometric rules. Both models were trained on a custom dataset reflecting local Sri Lankan driving conditions. Specifically, this research aims to;

1. Develop and compare CNN and object detection-based methods for lane violation detection.

2. Evaluate their suitability for real-world automated enforcement in Sri Lankan traffic environments.

The expected outcomes of this study include improved scalability for wider implementation, low-cost deployment utilizing standard dashcam hardware, and enhanced reliability compared to traditional manual enforcement methods.

2. LITERATURE REVIEW

Automated traffic monitoring has evolved from classical image processing techniques, such as deterministic "trip-wires" (Cucchiara et al., 2000), toward robust deep learning solutions capable of handling complex environments. The field was revolutionized by Convolutional Neural Networks (CNNs), particularly single-stage detectors like the YOLO (You Only Look Once) family, which balance real-time speed with high accuracy (Redmon & Farhadi, 2018).

Recent research has focused on optimizing these models for specific enforcement scenarios. Khanam and

Hussain (2024) highlight YOLOv5's architectural efficiency for constrained edge deployment, while Li et al. (2023) demonstrated its adaptability by integrating attention mechanisms for high-density aerial surveillance. In the context of mobile enforcement, Peng et al. (2022) successfully utilized in-vehicle cameras for real-time illegal parking detection, proving the viability of dynamic, non-stationary monitoring.

Current studies are also exploring newer variants like YOLOv8 to further enhance precision in safety-critical tasks. Modern enforcement often employs a hybrid strategy, combining these powerful perception models with rule-based logic to interpret complex violations. Our research builds upon this hybrid concept by comparing a custom CNN classifier against a rule-based object detection framework for lane discipline enforcement.

3. METHODOLOGY

This section outlines the process for data acquisition and details the architecture and implementation of the two deep learning approaches developed for violation detection.

4.1 Data Collection and Preparation

The foundation of this research is a specially constructed dataset to obtain the characteristics of Sri Lankan traffic environments. Five hours of video were recorded within the Kandy District with a high-definition mobile phone camera attached to a vehicle to mimic a dashcam perspective. The video was shot at a resolution of 1280x720 pixels at 30 FPS. After the raw footage was selected, a Python script from the OpenCV library was used to read out image frames at a rate

of two per second, resulting in a final raw dataset of approximately 1,000 images to be annotated. A sample of these frames appears in Figure 1.



Figure 1: Sample frames from the custom-collected dashcam dataset

4.2 Approach A: Violation detection via Image Classification

This approach treats violation detection as a binary image classification problem. The core idea is to train a model that can classify a tightly-cropped image containing a vehicle as either 'Violation' or 'Non-Violation'.

- **Dataset Generation:**

To create a focused dataset for this task, we used a Python script that was specifically tailored to auto-crop the original full-resolution frames. This operation isolated the primary area of interest of the car and its direct interaction with the road from other irrelevant background elements, such as the sky. These tiny images were tagged manually and curated to form a final dataset of 912 images spread across 'Violation' and 'Non-Violation' classes. Rotation, flip horizontal, and brightness adjustment data augmentation techniques were applied to increase the diversity of the dataset.

Sample final images are shown in Figure 2.



Figure 2: Sample images

Table 1: Summary of the custom CNN architecture

Layer(type)	Output shape	
conv2d (Conv2D)	(None,222,222,32)	896
max_pooling2d (MaxPooling2D)	(None,111,111,32)	0
conv2d_1 (Conv2D)	(None,109,109,64)	18, 496
max_pooling2d_1 (MaxPooling2D)	(None,54,54,64)	0
conv2d_2 (Conv2D)	(None,52,52,128)	73, 856
max_pooling2d_2 (MaxPooling2D)	(None,26,26,128)	0
flatten (Flatten)	(None,86528)	0
dense (Dense)	(None,128)	11, 075, 712
dropout (Dropout)	(None,128)	0
Dense_1(Dense)	(None,1)	129

- **Model Architecture:**

We used a sequential Convolutional Neural Network (CNN) for this task. As shown in Table 1, the design features three Conv2D and MaxPooling2D blocks with filters increasing from 32 to 128.

This is followed by a Flatten layer, a Dense layer (128 units), and a Dropout layer (0.5) for regularization. The final output layer uses a sigmoid activation to produce a probability score. The model was compiled with the Adam optimizer and a binary_crossentropy loss function.

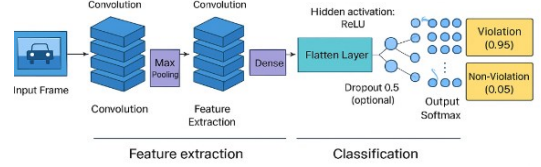


Figure 3: Custom CNN architecture

4.3 Approach B: Violation detection via Object Detection

This method uses a pre-trained one-stage object detector, specifically the YOLOv5s variant, to detect the first objects of interest in the entire frame. The model was trained for 50 epochs with a batch size of 16 and an input resolution of 416x416 pixels. Subsequently, a simple geometric rule is utilized to detect an infringement.

- **Data Annotation:**

To achieve this task, a dataset of 767 frames was manually annotated with bounding boxes using the Roboflow platform. Three primary classes were annotated: 'vehicle', 'single continuous line', and 'dotted line'.

- **Rule-Based Logic:**

The detection logic for this approach is direct and deterministic. After the model processes a frame and generates bounding boxes for all detected objects, a geometric intersection rule is applied. As illustrated in Figure 4, a violation is flagged if the bounding box of an object

classified as 'vehicle' is found to intersect with the bounding box of a 'single continuous line'.



Figure 4: Bounding Box Intersection

4.4 Experimental Environment

All experiments were conducted in a hybrid environment, utilizing both a local workstation and cloud infrastructure. The local setup included an AMD Ryzen 7 4800H CPU and AMD Radeon RX 5500M GPU. Deep learning model training was accelerated using NVIDIA GPUs provided by the Google Colab platform. The implementation relied on Python 3.11.9, TensorFlow 2.13.0, and Ultralytics 8.3.59 frameworks.

4. RESULTS AND DISCUSSION

This section presents and discusses the performance of the two methodologies, evaluated using Accuracy, Precision, Recall, and F1-Score.

4.1 Approach A: Custom CNN Performance

The custom CNN was evaluated on the held-out test set, achieving an overall accuracy of 96.7%. The confusion matrix (Figure 5) provides a detailed breakdown. This performance translates to a Precision of 0.96, a Recall of 0.98, and a balanced

F1-score of 0.97 for the critical 'Violation' class.

The model's success is attributed to its specialized task, training on curated images to learn specific violation features. The CNN achieved high reliability, indicating its suitability for violation detection tasks.

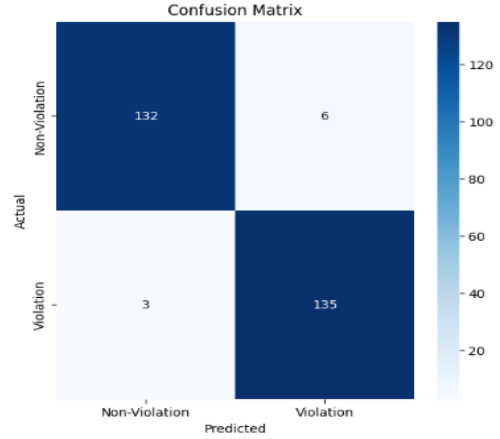


Figure 5: Confusion Matrix for the CNN



Figure 6: Examples of correct CNN classifications

4.2 Approach B: Object Detection Model Performance

The single-stage object detector's performance metrics are shown in Table 2. The model's Precision of 0.613 and recall of 0.784 serve as the input for the rule-based logic. While conceptually straightforward, this approach's performance is limited by the geometric imprecision of bounding boxes, which do not always conform to the exact shape of linear lane markings.

Table 2: Performance Metrics of the Trained YOLOv5 Model

Class	Images	Instances	Precision (P)	Recall (R)	mAP 50	mAP 50-95
All	74	397	0.613	0.784	0.752	0.438
Vehicle	74	262	0.437	0.630	0.567	0.335
Single continuous line	74	113	0.440	0.605	0.574	0.323
Dotted line	74	20	0.578	0.900	0.873	0.547

4.3 Limitations

This study's limitations include a dataset restricted to clear, daytime conditions. Model performance may degrade in challenging scenarios like rain, at night, or with faded lane markings.

5. CONCLUSION

This research validated the use of deep learning for automating lane violation detection from dashcams in Sri Lanka. We successfully implemented two methodologies: a custom CNN for classification and a rule-based object detector. The results confirm that both approaches are viable, with the custom CNN demonstrating high accuracy (96.7%) for the final verification task. This study affirms the potential of AI-driven systems to serve as a scalable tool to augment traditional traffic law enforcement and improve road safety.

Future work will focus on building upon these findings to create a more robust and comprehensive system. The key priorities are:

Improving Perception with Instance Segmentation: To overcome the

geometric imprecision of bounding boxes noted in this study, the next step is to replace the object detector with an instance segmentation model. This will provide precise, pixel-level masks for vehicles and lane lines, enabling more reliable geometric analysis.

Enhancing Model Robustness: The training dataset will be expanded with more challenging environmental conditions, such as nighttime, rain, and faded lane markings, to improve the system's performance in a wider variety of real-world scenarios.

Developing an Integrated LPR Pipeline: A complete, multi-stage pipeline will be developed that integrates a License Plate Recognition (LPR) module. This LPR system will be trained on a specialized dataset of difficult, real-world images and will be triggered only after a violation has been confirmed by the preceding logic stages.

Achieving Real-Time Speed: To make the system practical for deployment, the models will be optimized using techniques like quantization and conversion to an ONNX runtime.

REFERENCES

- Cucchiara, R., Piccardi, M., & Mello, P. (2000). Image analysis and rule-based reasoning for a traffic monitoring system, *IEEE Transactions on Intelligent Transportation Systems*, 1(2), 119 - 130. <https://doi.org/10.1109/6979.880969>
- Khanam, R., & Hussain, M. (2024). *What is YOLOv5: A deep look into the internal features of the popular object detector*. arXiv. <https://doi.org/10.48550/arXiv.2407.20892>
- Li, Z., Fan, B., Xu, Y., & Sun, R. (2023). Improved Yolov5 for Aerial Images Based on Attention Mechanism. *IEEE Access*, 11, 96235 - 96241. <https://doi.org/10.1109/ACCESS.2023.3277931>

- Peng, X., Song, R., Cao, Q., Li, Y., Cui, D., Jia, X., Lin, Z., & Huang, G. B. (2022). Real-Time Illegal Parking Detection Algorithm in Urban Environments. *IEEE Transactions on Intelligent Transportation Systems*, 23(11), 20572 - 20587. <https://doi.org/10.1109/TITS.2022.3180225>
- Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv. <https://doi.org/10.48550/arXiv.1804.02767>



Automated Classification of Mobile App Reviews Using Zero-Shot NLP Model: A Case Study on MYSLT

Jayawardhana KPRK¹, Vidanagama VGTN²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

ruchirakavinda17@gmail.com¹

ABSTRACT

Mobile applications are now central to service delivery, user engagement, and digital transformation. With a growing number of people using apps, user reviews on app stores such as the Google Play Store and Apple App Store are invaluable feedback sources for developers. These reviews, however, are unstructured, multilingual, and voluminous in nature, rendering manual analysis impossible. This study presents an automated model that semantically classifies MYSLT app reviews into three general categories: bug reports, feature requests, and UI/UX complaints. The model uses Natural Language Processing (NLP) and zero-shot classification models, namely the facebook/bart-large-mnli transformer, to perform semantic classification without requiring pre-labeled datasets. The pipeline automatically performs data gathering, preprocessing, translation, classification, and visualization using Python and Streamlit. The evaluation yielded promising results, averaging 0.84 precision, recall of 0.81, and an F1-score of 0.82 across all categories. The outcomes demonstrate that zero-shot models can effectively mine actionable insights from user feedback to support data-driven decisions within mobile application development.

KEYWORDS: App Review Analysis, Feature Request Classification, MYSLT App, Natural Language Processing, Zero-Shot Learning

1. INTRODUCTION

Mobile applications are central to modern digital services, and MYSLT from Sri Lanka Telecom allows users to handle broadband, billing, and customer care on a single platform. As the usage increases, the app receives thousands of reviews that contain very valuable feedback regarding performance, features, and user experience.

However, these comments are unstructured in nature and hence difficult to transform into actionable insights for the developers. Manual review is impracticable, and existing tools support only basic sentiment analysis without domain-specific categorization, making it hard to identify and prioritize key user issues.

This research proposes an automated NLP using zero-shot learning for the classification of MYSLT app reviews into categories like bug reports, feature requests, and UI/UX complaints. It tests whether a zero-shot classifier can effectively analyze feedback without the need for labeled data, thus providing developers with a scalable tool that will speed up insight-driven decision-making.

2. LITERATURE REVIEW

The automatic analysis of mobile application reviews has been growing rapidly as developers try to find efficient solutions for processing voluminous unstructured feedback. Early solutions detected bugs or crashes using keyword- and rule-based techniques which are simple but suffer from synonymy and complex phrasing and clearly lack generalization. Later on, with the

supervised machine learning models like SVMs and Naive Bayes, the accuracy of classification improved significantly but they required extensive manually labeled datasets, which cost and are time-consuming to develop (Huang & Chen, 2020).

Advances in NLP brought with them more robust methods to extract insights from reviews. Techniques of topic modeling, such as LDA, allowed for feature request identification, while sentiment analysis allowed a broad understanding of the level of satisfaction, though failing to yield any underlying causes. More recently, transformer-based models such as BERT (Devlin et al., 2019) and BART (Lewis et al., 2020) substantially improved contextual understanding. Their zero-shot learning capabilities, especially in models like BART-large-MNLI, allow for classification into user-defined categories with absolutely no domain-specific training data, tackling the labeled-data bottleneck directly. These new avenues are used increasingly to detect bugs, identify feature requests, and recognize UI/UX issues much more accurately.

Despite this progress, several gaps remain. For example, there is a lack of multilingual environments, such as Sri Lanka, where reviews often contain both Sinhala and English. Few works provide end-to-end pipelines that combine data collection, preprocessing, classification, and visualization. Additionally, little research uses modern NLP techniques on telecommunication apps in developing regions. This work tries to fill these gaps by combining zero-shot transformers with a real-time visualization dashboard, tuned for MYSLT app reviews.

3. METHODOLOGY

The work is an automated data-driven pipeline shown in Fig. 1 based on study that collects, preprocesses, and classifies application reviews.

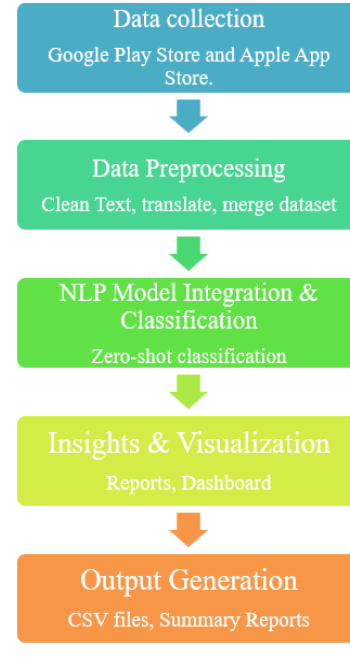


Figure 1: Automated data-driven pipeline

3.1 Data Collection

Table 1: Summary of Preprocessed MYSLT Review Dataset

Attribute	Value
Total Raw Reviews	3,145
Period Collected	All the comments up to present
Platforms	Google Play Store, Apple App Store
Cleaned Final Dataset	2,680

Reviews were collected from Apple App Store and Google Play Store through Python scripts. The google-play-scraper library and iTunes RSS API were used to scrape fields such as username, review,

date, and rating. Reviews were filtered by region ('lk') and language ('en') for localized context.

3.2 Preprocessing

Non-English reviews were translated to English through the deep_translator API. Empty, invalid, or duplicate rows were removed by text cleaning. Review dates were normalized to datetime format.

3.3 Zero-Shot NLP Model Classification

The Facebook / BART-large-MNLI transformer model was utilized in the study because of its high performance and popularity among zero-shot text classification tasks. It was developed by Facebook AI and is founded on the BART architecture, which combines BERT's bidirectional encoding and GPT's autoregressive decoding and yields high ability to understand and generate text. Fine-tuned on the Multi-Genre Natural Language Inference (MNLI) dataset, it can determine complex semantic relations in text pairs. This capacity allows BART-large-MNLI to accurately classify app reviews into labels such as bug reports, feature requests, or UI/UX issues without ever seeing labeled data. As Yin et al. (2019) and Hugging Face (2024) documentation point out, the model generalizes extremely well. The feature requests were also classified into eight subcategories, and confidence scores were recorded for each.

3.4 Subcategory Classification

Feature requests were also categorized into eight subgroups: UI/UX, Payment, Authentication, Notifications, Data Usage, Customer Support, Language

Support, and Other. Confidence scores were recorded per category.

3.5 Automation and Visualization

Automatic updates are executed every 24 hours to fetch new reviews. There is a Streamlit dashboard showing categorized data in bar charts and interactive filters through which developers are able to monitor feature request trends in real time.

4. DATA ANALYSIS

User reviews of the MYSLT mobile application were collected using Python-based automated extraction tools from Google Play and the Apple App Store. The google-play-scraper library together with iTunes RSS API allowed the retrieval of country-specific-lk and English language-en reviews along with review text, rating, timestamp, and username. Over 3,000 reviews were gathered between January and June 2024.

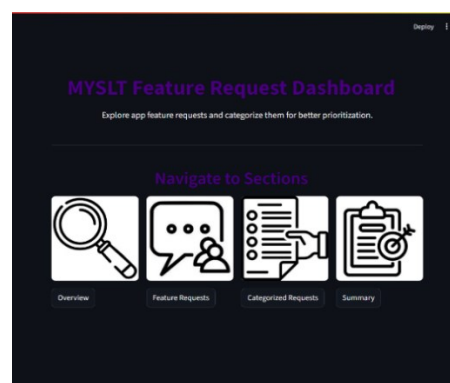


Figure 2: MY SLT Feature Request Dashboard

A strict preprocessing pipeline guaranteed the reliability of data and its analytical validity. This included removing blank and duplicate entries, translating non-English content, normalizing date formats, and cleaning symbols and emojis. After preprocessing, the dataset was refined to 2,680 unique records. In classification, the transformer-

based Facebook/BART-large-MNLI model was used in zero-shot mode to categorize reviews into Bug Reports, Feature Requests, and UI/UX Complaints. Confidence at or above 0.75 was chosen as an optimal threshold for a balance between precision in classification and dataset coverage. For further detailed insight generation, Feature Requests were divided into six thematic subcategories.

For the model's reliability validation, 300 reviews were manually labelled by two independent evaluators. The resulting Cohen's Kappa score of 0.82 indicated a strong inter-rater agreement. Model predictions on those were measured against these gold-standard labels to compute the classification accuracy metrics.

5. RESULTS AND DISCUSSION

Table 2 and Fig. 3 summarizes the classifier's precision, recall, and F1-score for each category, along with the overall average, reflecting its classification performance on a manually validated subset of reviews.

The distribution of average model confidence across primary review categories is illustrated in Fig. 4, highlighting variations in the classifier's certainty when predicting different categories.

Table 2: Table of Classification Performance Matrix

Category	Precision	Recall	F1-Score
Bug Reports	0.85	0.81	0.83
Feature Requests	0.88	0.84	0.86
UI/UX Complaints	0.80	0.77	0.78
Average	0.84	0.81	0.82

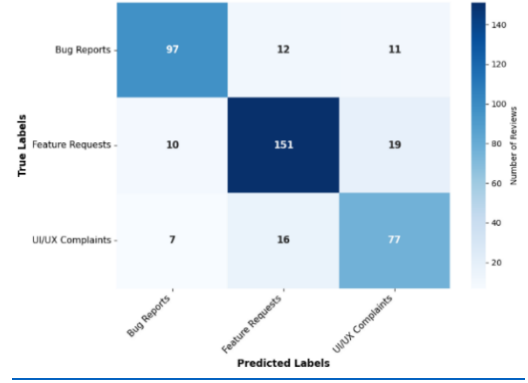


Figure 3: Confusion matrix

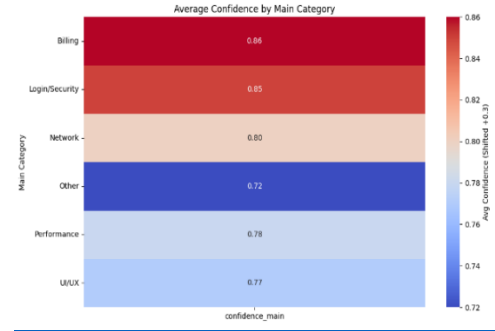


Figure 4: average model confidence across primary review categories

Meanwhile, Fig. 5 depicts the temporal dynamics of categorized user reviews, showing weekly fluctuations in review frequency across main categories over time.

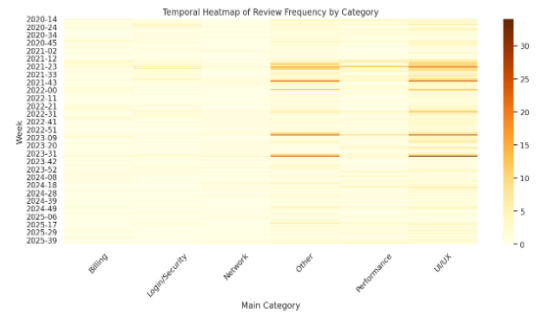


Figure 5: Temporary heatmap of review frequency by category

Together, these figures provide insights into both the reliability of the classification model and the evolving trends in user feedback across different functional aspects of the application.

Automation within the system. The results indicate that the zero-shot classification approach can effectively classify app reviews without training on domain-specific labeled data. High precision and recall for feature request detection confirm its readiness for practical applications.

Visualizations revealed feature requests as the most prevalent category, with 46% of total reviews, followed by bug reports (34%) and UI/UX complaints (20%). Some frequent feature requests included dark mode, biometric login, and real-time bill notifications.

Enables developers to monitor user needs around the clock, enabling quicker updates and improved customer satisfaction. Nevertheless, challenges, such as, processing mixed-language (Sinhala-English) and sarcastic reviews are areas for additional improvement.

6. CONCLUSION

This study shows that zero-shot NLP models are able to effectively classify unstructured mobile app reviews into semantically meaningful categories without labeled data. The pipeline configuration of the MYSLT mobile application successfully mined user feedback patterns and visualized them in interactive dashboards, giving rise to a data-driven understanding of user requirements and app performance. The findings contribute to software maintenance automation by offering a language-agnostic, scale-out approach to extracting app improvements to be prioritized using transformer-based models.

Future studies could focus on model fine-tuning such as BART, RoBERTa, or

DistilBERT over domain-specific datasets to increase contextual accuracy. Multilingual representations such as XLM-RoBERTa or mBERT could be incorporated to further support regional languages, and sentiment and aspect-based analysis would be used to increase contextual interpretation for better product development and user experience.

REFERENCES

- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019)*, 4171 - 4186.
- Huang, Y. F., & Chen, P. H. (2020). Fake News Detection Using an Ensemble Learning Model Based on Self-adaptive Harmony Search Algorithms. *Expert Systems with Applications*, 159, 113584. <https://doi.org/10.1016/j.eswa.2020.113584>
- Hugging Face. (2024). *facebook/bart-large-mnli – Zero-shot classification model*. <https://huggingface.co/facebook/bart-large-mnli>
- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2020). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 7871 - 7880. <https://doi.org/10.18653/v1/2020.acl-main.703>
- Yin, W., Hay, J., & Roth, D. (2019). Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, China*, 3914 - 3923.



A Comparative Study of the Usability of Different Software Tools in Auditing

Karunarathne KTT¹, Munasinghe B²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

thamodhikarunaratna@gmail.com¹, bhagya@wyb.ac.lk²

ABSTRACT

This study evaluates the performance of three audit analytics tools (i.e., Microsoft Excel, Alteryx, and ACL Analytics) used by an auditing company, providing insights from a software engineering perspective to improve the auditing process. Using a two-phase approach combining secondary data analysis and a task-based practitioner survey, the first phase compared the tools across functionality, usability, scalability, integration, and impact on audit quality. The second phase linked eight analytical audit tasks to overall audit performance. Findings indicate that Alteryx excelled for large, complex audits due to automation and integration, ACL Analytics was best for audit-specific and compliance tasks with high precision and control, and Excel remained most accessible for smaller, structured audits. The study shows that no single tool is universally superior; effectiveness depends on task type, data complexity, and audit objectives. Overall, it provides a practical framework for selecting audit analytics tools to enhance efficiency, accuracy, and audit coverage.

KEYWORDS: ACL Analytics, Alteryx, Audit Analytics, Data-driven analytics

1. INTRODUCTION

Auditing plays a critical role in ensuring the accuracy and reliability of financial reporting, but traditional methods are often inefficient, error-prone, and difficult to scale for large or complex datasets. Specialized audit analytics tools, such as ACL Analytics and Alteryx, offer advanced functionalities including automated testing, fraud detection, workflow automation, and integration with multiple data sources which can improve audit efficiency, accuracy, and quality. Despite their growing adoption, there is limited research comparing these tools systematically in terms of usability, scalability, integration, and impact on audit outcomes. This study addresses this gap by evaluating Excel, ACL Analytics, and Alteryx, aiming to provide practical guidance for auditors and organizations in selecting the most suitable tool to enhance audit effectiveness. These three tools were

selected because they represent the main audit tool types: Excel (traditional), ACL Analytics (audit-focused), and Alteryx (modern automated analytics).

2. LITERATURE REVIEW

The auditing profession is changing quickly as technology improves data analysis, interpretation, and validation. Automation and data analytics tools make audits more efficient, accurate, and scalable, reducing the need for manual work. (Appelbaum et al., 2017). Traditional audit methods, dependent on sampling and spreadsheets, struggle with the volume and complexity of modern financial data (Janvrin et al., 2008). Computer-Assisted Audit Techniques such as ACL Analytics and Alteryx have therefore become essential in enhancing audit quality and supporting evidence-based decision-making. Microsoft Excel remains widely used for tasks like trial balance reconciliation, journal testing,

and variance analysis due to its accessibility, flexibility, and low cost (Janvrin et al., 2009; Dowling & Leech, 2014). However, Excel is error-prone, inefficient for large datasets, and lacks automation, limiting its suitability for complex audits (Panko & Aurigemma, 2010; Alles, 2015). ACL Analytics, designed specifically for auditing, allows automation of key procedures such as control testing and duplicate detection, enhancing accuracy and traceability (Alles, 2015; Warren et al., 2015). Its steep learning curve and limited integration with cloud-based systems remain challenges (Kokina & Davenport, 2017).

Alteryx, a newer platform, emphasizes automation, usability, and integration, enabling auditors to build visual workflows, connect to multiple data sources, and perform predictive analyses without extensive programming (Sun & Vasarhelyi, 2021; Henderson et al., 2020). While highly scalable and suitable for large datasets, its cost and training requirements may limit adoption (Cao et al., 2015). Tool adoption also depends on usability, organizational support, and data integration capabilities (Bierstaker et al., 2014).

Although research on analytics in auditing is increasing, few studies directly compare Excel, ACL, and Alteryx under similar conditions (Appelbaum et al., 2017; Alles, 2015). This study fills that gap by assessing the tools' functionality, usability, scalability, integration, and impact on audit quality, offering guidance for auditors in selecting appropriate tools for modern audit practices.

3. METHODOLOGY

3.1. Research Approach

The study adopted a two-phase approach combining secondary data analysis and primary data collection to evaluate Excel, Alteryx, and ACL Analytics in auditing. Phase One involved secondary research to understand the theoretical and functional aspects of each tool, while Phase Two collected practical insights from audit professionals through a structured survey. This approach ensured both conceptual understanding and real-world application was captured.

3.2. Methodology: Phase 01

3.2.1. Data Collection

Information was gathered from literature, vendor documentation, professional resources, and case studies to find out Special features relevant to auditing, Capabilities supporting audit processes, Practical applications, Limitations and challenges of Excel, Alteryx and ACL Analytics.

3.2.2. Comparative Analysis

This stage aimed to develop a comprehensive understanding of how each tool functions in an audit environment. The findings from this phase were planned to be structured into a comparative framework based on five key evaluation dimensions:

- 1) Functionality: type of tasks supported and level of automation.,
- 2) Usability: ease of learning and ease of use.,
- 3) Scalability: ability to handle increasing data size.,
- 4) Integration with other systems: ability to connect with databases, ERPs, and other systems,
- 5) Impact on audit quality: effect on accuracy, coverage, and error reduction.

A COMPARATIVE STUDY OF THE USABILITY OF DIFFERENT SOFTWARE TOOLS IN AUDITING

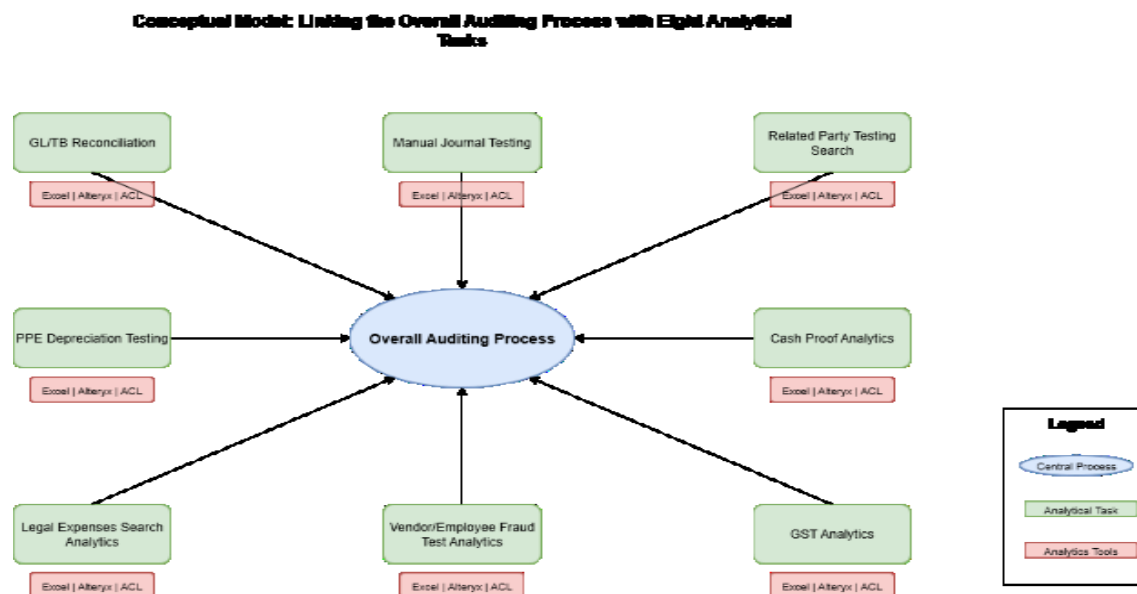


Figure 1: Conceptual model of the audit tasks

3.3 Methodology: Phase 02

3.3.1. Survey Design & Data Collection

A structured survey was designed to capture auditors' practical experiences with Microsoft Excel, Alteryx, and ACL Analytics across eight key audit tasks: GL to TB Reconciliation, Manual Journal Testing, Related Party Testing, Cash Proof Analytics, GST Analytics, Vendor/Employee Fraud Testing, Legal Expenses Analytics, and PPE Depreciation.

The survey was distributed to the Data Analytics team of the GRCFDA (Governance, Risk, Compliance, Forensic & Data Analytics) department, which plays a central role in applying data-driven techniques for auditing and assurance tasks. Participants represented different experience levels, including junior auditors, analysts, and senior audit professionals, ensuring a balanced range of perspectives on tool usage and effectiveness. The participant distribution is shown in Figure 2. The questionnaire consisted of a mix of closed-ended and

rating-scale questions, focusing on the following aspects:

- Experience level and tools used
- Task-specific tool usage and effectiveness (1 – 5 scale)
- Benefits and challenges encountered
- Recommended tool for each task

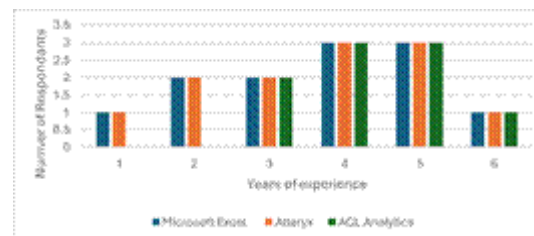


Figure 2: Number of respondents on each tool

3.3.2. Conceptual Model of the Audit Analytic tasks

The eight analytical tasks represent key components of the auditing process, covering objectives from data reconciliation and risk assessment to fraud detection and financial reporting. The model shown in Figure 1 illustrates eight independent audit workflows, each

performed using Excel, Alteryx, or ACL Analytics.

3.3.3. Data Analysis

Survey responses were analyzed using descriptive statistics, including frequency distributions, mean effectiveness ratings, and percentage comparisons across the tools. Responses for each audit task were summarized using average scores on the 1 – 5 scale to compare Excel, Alteryx, and ACL Analytics. Qualitative comments were grouped into themes relating to benefits and challenges. Inferential statistical tests were considered. The analysis was mainly descriptive due to the small sample size. Inferential tests were considered but not used, as the sample was too small for reliable results. Therefore, tool comparisons were based on descriptive trends.

4. RESULTS AND DISCUSSION

4.1. Results and Findings: Phase 01

Table 1 compares Excel, Alteryx, and ACL Analytics across functionality, usability, scalability, integration, and audit quality. Excel is highly accessible and effective for small to medium-sized datasets and routine tasks but lacks automation and advanced analytics.

Alteryx handles large (above 200,000), complex datasets with automated workflows, predictive analytics, and strong integration, though it requires moderate training. ACL Analytics, designed for auditing, offers high accuracy, robust automation, and standardized reporting, but has a steeper learning curve. In summary, Excel suits simple tasks, Alteryx excels in scalable, data-intensive audits, and ACL Analytics is ideal for specialized audit testing requiring precision.

4.2 Results and Findings: Phase 02

The survey involved professionals from the GRCFDA department with different levels of auditing experience. A total of 16 questionnaires were distributed, and 12 responses were received, resulting in a response rate of 75%. 12 participants were familiar with Excel and Alteryx, while only 8 had used ACL Analytics, indicating its more specialized use in enterprises.

Alteryx was the preferred tool for tasks involving large datasets (e.g., GL/TB Reconciliation, Cash Proof, and Fraud Tests), ACL Analytics excelled in audit-specific procedures, and Excel remained suitable for smaller, routine tasks.

Table 1: Comparison Analysis between Microsoft Excel, Alteryx and ACL Analytics

Dimension	Excel	Alteryx	ACL Analytics
Functionality	Small-medium datasets (approximately 100,000–200,000); basic reporting; limited analytics	Large datasets (above 200,000; automated workflows; predictive/ML analytics	Audit-focused; transaction testing; fraud detection; strong automation
Usability	Familiar interface; easy for routine tasks; VBA for advanced	Workflow interface; moderate learning curve; training recommended	Command-driven; steep learning curve; auditor-focused
Scalability	Limited; performance issues with many users; macros	Handles large datasets; reusable workflows; multi-user	Enterprise audits; built-in automation; controlled access

A COMPARATIVE STUDY OF THE USABILITY OF DIFFERENT SOFTWARE TOOLS IN AUDITING

Integration	SQL/Access; manual ERP import; limited API	Extensive DB & ERP connectors; BI & cloud integration	Audit-specific DB/ERP support; exports to BI; workflow APIs
Audit Quality	Limited coverage; basic fraud detection; weak audit trail	Broad coverage; automated error reduction; ML-assisted fraud detection	High coverage; standardized detection; compliance-focused; strong audit trail

4.2.1. Average Effectiveness of Analytical Tools

Respondents rated the effectiveness of tools on a five-point scale (1 = Very Low, 5 = Very High). Overall, Average effectiveness scores indicated that using data analytics tools in auditing was more effective than manual methods across most tasks (Table 2).

Table 2: Average Effectiveness of Analytical Tools

Audit Context	Average Effectiveness Rating (1-5)
Related party transaction search	4.50
Legal expenses Search Analytics	4.50
Manual Journal testing	4.33
Vendor Employee Fraud test analysis	4.33
GL TB Reconciliation	4.17
Cash proof Analysis	4.17
PPE Depreciation analysis	4.08
GST Analytics	3.92

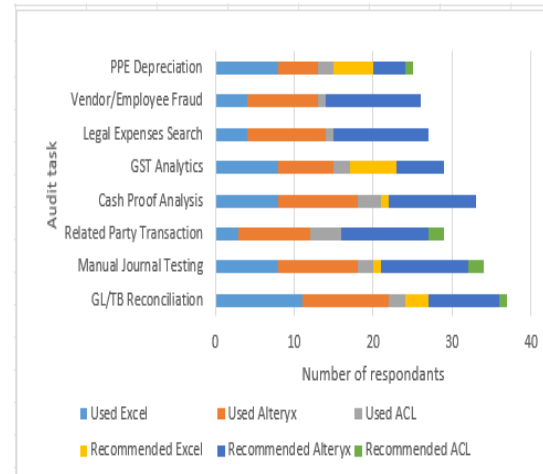


Figure 3: Tool usage vs tool suitability

4.2.2 Comparative Tool usage vs. tool suitability analysis

Respondents were asked which tools they currently use and which they consider most suitable for each audit task.

Figure 3 presents a comparison of actual usage versus perceived suitability. Alteryx emerged as the top choice for tasks requiring large-scale (above 200,000), data manipulation. ACL Analytics was rated highest for audit-focused procedures such as compliance testing and risk assessment, whereas Excel remained adequate for smaller, routine tasks.

4.2.3 Benefits and Challenges

Practitioners reported several benefits and challenges when using audit tools. In terms of efficiency, tools reduce time-consuming tasks by speeding up GL-TB matching and quickly scanning legal-

related expenses, though they struggle with messy data, complex mapping, and require manual updates. For reliability, automated matching enables analysis of large volumes and testing of 100% of assets, but standardization is difficult, formulas can reduce transparency, and complex journal entries may still need manual judgment. Regarding accuracy, tools provide high precision, easy transaction selection, and rapid detection of mismatches or duplicates; however, outputs depend on pre-defined rules, GST inputs/outputs are hard to identify, and manual review remains necessary.

The successful use of audit analytics tools depends on several practical challenges. Data quality issues, such as incomplete or inconsistent records, can limit the accuracy of automated tests. Most tools require proper training for auditors to use advanced functionalities effectively, especially in Alteryx and ACL Analytics. Manual review is still necessary for complex transactions, exceptions, and judgment-based areas. Additionally, organizational support, such as access to licensed tools, clear workflows, and management encouragement, plays a critical role in ensuring effective adoption of these tools in practice.

5. CONCLUSION

This study evaluated and compared three audit analytics tools; Microsoft Excel, Alteryx, and ACL Analytics using a two-phase approach that combined secondary analysis and a practitioner survey. Guided by a conceptual model linking eight analytical tasks; GL/TB Reconciliation, Manual Journal Testing, Related Party Testing Search, Cash Proof

Analytics, GST Analytics, Vendor / Employee Fraud Test Analytics, Legal Expenses Search Analytics, and PPE Depreciation Testing to the auditing process, the study examined each tool's capabilities across five evaluation dimensions: functionality, usability, scalability, integration, and impact on audit quality.

Findings revealed that Excel remains popular for its simplicity and flexibility in smaller tasks but struggles with large datasets (above 200,000) and lacks advanced automation. Alteryx emerged as the most powerful and efficient tool, offering robust automation, scalability, and integration for data-intensive audits. ACL Analytics excelled in audit-specific and compliance tasks, delivering high accuracy and control, though it requires greater technical expertise and investment. The results highlight that tool effectiveness depends on task type rather than the tool itself. The conceptual model provides a practical framework to match audit tasks with the most suitable tool. Integrating analytics tools enhances audit efficiency, reduces manual effort, and improves audit quality, although successful implementation requires proper training, standardized processes, and organizational support.

Despite limitations such as a small sample and focus on only three tools, the study underscores the critical role of audit analytics in modern auditing. It offers actionable guidance for auditors to leverage Excel, Alteryx, and ACL Analytics strategically, promoting more accurate, efficient, and reliable audit practices while advancing data-driven, technology-enabled auditing.

REFERENCES

- Alles, M. G. (2015). Drivers of the use and facilitators and obstacles of the evolution of Big Data by the audit profession. *Accounting Horizons*, 29(2), 439 - 449. <https://doi.org/10.2308/acch-511067>
- Appelbaum, D., Kogan, A., & Vasarhelyi, M. A. (2017). Big Data and analytics in the modern audit engagement: Research needs. *Auditing: A Journal of Practice & Theory*, 36(4), 1 - 27. <https://doi.org/10.2308/ajpt-51684>
- Bierstaker, J., Janvrin, D., & Lowe, D. J. (2014). What factors influence auditors' use of computer-assisted audit tools? *Advances in Accounting*, 30(1), 67 - 74. <https://doi.org/10.1016/j.adiac.2013.12.005>
- Cao, M., Chychyla, R., & Stewart, T. (2015). Big Data analytics in financial statement audits. *Accounting Horizons*, 29(2), 423 - 429. <https://doi.org/10.2308/acch-51068>
- Dowling, C., & Leech, S. A. (2014). A Big 4 firm's use of information technology to control the audit process: How an audit support system is changing audit practice. *Contemporary Accounting Research*, 31(1), 230 - 252.
- Henderson, D., Park, J., & Kim, J. (2020). Automating the audit process through workflow analytics: A case study of Alteryx. *International Journal of Accounting Information Systems*, 37(1), 1 - 12.
- Janvrin, D., Bierstaker, J., & Lowe, D. (2008). An examination of audit information technology use and perceived importance. *Accounting Horizons*, 22(1), 1 - 21.
- Janvrin, D., Bierstaker, J., & Lowe, D. (2009). An examination of auditing professionals' technology adoption and use. *International Journal of Accounting Information Systems*, 10(4), 170-194.
- Kokina, J., & Davenport, T. H. (2017). The emergence of artificial intelligence: How automation is changing auditing. *Journal of Emerging Technologies in Accounting*, 14(1), 115 - 122.
- Panko, R. R., & Aurigemma, S. (2010). Revisiting the Panko-Halverson taxonomy of spreadsheet errors. *Decision Support Systems*, 49(2), 235 - 244.
- Sun, T., & Vasarhelyi, M. A. (2021). Transforming auditing through robotic process automation and analytics platforms. *International Journal of Digital Accounting Research*, 21, 1 - 25.
- Warren, J., Moffitt, K., & Byrnes, P. (2015). How Big Data will change accounting. *Accounting Horizons*, 29(2), 397 - 407.



A Contextual Analysis of Process Effectiveness and Industry Alignment in a Software Industry Workplace, Including a Benchmarking System Prototype

Fernando NTS¹, Munasinghe B²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

tharushifernando953@gmail.com¹, bhagya@wyb.ac.lk²

ABSTRACT

This study analyzes software development practices in a small to medium-scale Sri Lankan Software Company, assessing their effectiveness and alignment with industry standards. Using a pragmatic approach based on qualitative methods through observations, surveys, interviews, and document review, it identifies the company's strengths, weaknesses, and improvement areas. A benchmarking system prototype was developed to compare company practices against frameworks like Software Development Life Cycle (SDLC), Agile, and DevOps, offering transparent scoring and recommendations. Results show partial adoption of modern methods such as Agile and Continuous Integration / Continuous Deployment (CI/CD), but highlight gaps in documentation, continuous monitoring, and integration. The prototype offers practical insights for aligning processes with standards. Overall, the research provides both evidence and a tool for continuous improvement, stressing the need for regular evaluation, adaptability, and benchmarking to achieve process maturity and competitiveness in the software industry.

KEYWORDS: Agile, Benchmarking System, Industry Alignment, Process Effectiveness

1. INTRODUCTION

Software development organizations face ongoing challenges in ensuring that their processes are effective, efficient, and aligned with evolving industry standards. Variations in the adoption of different methodologies, such as Agile, DevOps, and SDLC, often lead to inconsistent performance and suboptimal outcomes across teams and projects. Understanding these gaps is essential for organizations aiming to enhance productivity, quality, and competitiveness.

During and after the COVID-19 pandemic, the global software industry experienced major transformations, marked by rising demand for digital solutions and expanded opportunities for small and medium-sized enterprises (Ghimire et al., 2023; Surahman et al.,

2023). This shift also spurred the growth of small to medium-scale software development firms in Sri Lanka's IT sector, reflecting trends in digital entrepreneurship and market diversification (Indika et al., 2020; Pandithasekara & Hewage, 2022). However, the adoption of systematic software engineering practices by these emerging companies remains understudied.

This study, conducted during an industrial training period of six months, intends to understand the software engineering (SE) practices in small to medium-scale software companies in Sri Lanka through investigating software development practices in an industry workplace, analyzing its business process to investigate the effectiveness and

alignment of their practices with SE best practices. A benchmarking system prototype was also developed and used in the study to systematically compare organizational practices against established standards, providing actionable insights.

2. LITERATURE REVIEW

A range of SE approaches adopted by organizations to balance efficiency, adaptability, and alignment with industry standards are found in literature. Omonije, (2024). One of the studies emphasizes balancing competing principles to influence teams toward either Agile or structured approaches focusing on underlying principles rather than rigid practices. Such practices enable flexibility, better decision-making, and the evolution of more adaptive methodologies (Dalla, 2020). Modern software development increasingly integrates SDLC principles with DevOps practices, emphasizing continuous integration, delivery, and automation to create flexible, efficient, and user-focused software. This shift highlights the importance of adopting adaptable, feedback-driven workflows that balance structured processes with iterative, automated practices in contemporary software engineering (Quiroz-Vasquez, 2024).

The choice between Agile and Waterfall methodologies depends on project-specific factors, emphasizing the need to align development approaches with organizational goals, team capabilities, and the complexity of project requirements (Pargaonkar, 2023). Reviews on the Agile approach highlights its emphasis on flexibility, iterative development, and customer

responsiveness, positioning it as highly effective in dynamic and uncertain environments (Anifa et al., 2024)

Offerman & Blinde (2022) identifies 14 core DevOps practices and finds that higher adoption and maturity levels generally improve software delivery and organizational performance yet indicate potential drawbacks, such as decreased predictability and reduced job satisfaction, highlighting the need for balanced, context-aware practices. Integrating DevOps into SDLC enhances efficiency and project performance through technical, measurement, and agreement practices, but further empirical research is needed due to reliance on secondary sources and rapidly evolving technologies (Amaro et al., 2025).

3. METHODOLOGY

This study adopts a pragmatic approach based on qualitative observation-oriented research integrated with a case study method to evaluate the software development practices within a selected software industry workplace (Figure 1). Data for this research was collected over a six-months during the first author's industrial training, employing multiple methods to gain a comprehensive understanding of the organization's software development practices. These included direct observations of team workflows and meetings, structured surveys of employees, informal interviews with project staff, and review of project documents. Data collection emphasized identifying communication patterns, supervision quality, workload balance, and adherence to frameworks such as Agile, DevOps, and SDLC, allowing for

meaningful comparisons with recognized industry standards.

The data was then analyzed qualitatively using comparative analysis, thematic coding, and basic statistical summaries, to understand workflows, team coordination, and management practices. The approach facilitated in-depth exploration of how projects are planned, executed, and monitored in real-world conditions. Qualitative insights from interviews and observations were coded into recurring themes such as task clarity, documentation gaps, and workflow efficiency. Quantitative survey data were summarized through descriptive statistics and visualized using bar and pie charts.

Additionally, an independent benchmarking system prototype was developed as a supplementary component to compare organizational practices against standardized performance criteria. The combined analytical approach ensured that the findings were both evidence-based and actionable, supporting recommendations for improving process effectiveness and industry alignment.

4. RESULTS AND DISCUSSION

The analysis revealed a paradox between high technical competency and operational effectiveness. Although 100% of staff reported Agile adoption, expected benefits like flexibility and predictability were not fully realized, highlighting those modern practices alone do not ensure efficiency without proper alignment in structure, communication, and planning.

1. *Central Paradox: Competency vs. Operational Strain*

Despite high technical competency and full Agile adoption, the organization struggles with process inefficiency and unpredictable outcomes. While teams demonstrate strong tool proficiency and Agile literacy, gaps in workflow governance and coordination undermine consistency. The benchmarking system showed a moderate alignment of 68% with industry best practices, underscoring the gap between nominal compliance and actual effectiveness.

2. *Core Structural Weakness: Planning Deficit*

A significant contributor to inefficiency is the lack of structured pre-development planning. Only 37.5% of staff confirmed the use of documentation such as Test-Driven Development and Architectural Decision Records (TDD/ADR) before development, while most remained neutral. The absence of technical blueprints leads to frequent rework, extended timelines, and heavier workloads, reflecting a reactive rather than proactive approach. Staff consistently requested “more planning and structured techniques,” confirming the need for stronger upstream governance.

3. *Operational and Contextual Friction*

Three major friction points emerged affecting workflow and collaboration;

- **Cross-Functional Communication Bottlenecks:**

Collaboration among quality assurance, human resource management and development teams remains inconsistent, with 7 of 13 participants opting to be neutral about commenting on

communication quality, leading to coordination overhead.

- **Remote Parity Failure:**

About 80% of remote staff reported unclear goals, showing Agile ceremonies fail to align distributed teams, causing miscommunication and delivery delays.

- **Technical Maturity Gaps:**

While 11 of 13 participants confirmed strong tool usage, DevOps practices like CI/CD and automated testing remain partial, with testing thoroughness at **75%**, contributing to technical debt and slower feedback cycles.

Strategic Imperatives and Benchmarking Insights

To overcome these issues, the research proposes a roadmap focusing on lightweight governance, automation, and standardized collaboration frameworks. The benchmarking system, developed as a complementary tool, enables organizations to assess process maturity, compare against industry standards, and obtain actionable insights. Key recommendations include:

- **Mandate Upstream Governance:**

Adopt lightweight TDD/ADR documentation before development to reduce rework.

- **Achieve Remote Parity:**

Standardize Agile ceremonies across distributed teams for clarity and accountability.

- **CI/CD Gap:**

Expand automation in testing and deployment to lower technical debt.

- **Formalize Interdepartmental Protocols:**

Establish Service Level Agreements (SLAs) to improve cross-functional coordination.

Overall, the results indicate that, in a small to medium-scale company, despite the strong technical capability, true process effectiveness requires tighter integration between governance, communication, and automation. The results further validate the benchmarking system's value in promoting data-driven decision-making and alignment with evolving software industry standards.

5. CONCLUSION

The study shows that despite strong technical competency, operational inefficiencies persist within the company's business process due to gaps in planning, governance, and automation. Issues like cross-functional bottlenecks, remote misalignment, and partial DevOps adoption reduce workflow efficiency. The benchmarking system proposed in this study offers a practical framework to assess process maturity, compare performance with industry standards, and identify improvements. Implementing the benchmarking framework can guide targeted interventions, improve team coordination, and enhance overall organizational performance, with potential applicability to other Small and Medium-sized Enterprises (SMEs) in similar software development contexts.

These findings highlight that effective software development requires combining technical expertise with structured planning, standardized collaboration, and continuous evaluation to address operational challenges and align with best practices.

REFERENCES

- Amaro, R., Pereira, R., & De Silva, M. M. (2025). Mapping DevOps capabilities to the software life cycle: A systematic literature review. *Information and Software Technology*, 17, 107583. <https://doi.org/10.1016/j.infsof.2024.107583>
- Anifa, M., Ramakrishnan, S., Joghee, S., & Kabiraj, S. (2024). Systematic Review of Literature on Agile Approach. *Sage Journals*, 32, 84 - 105.
- Dalla, L. O. F. B. (2020). Lean software development practices and principles in terms of observations and evolution methods to increase work environment productivity. *International Journal of Engineering and Modern Technology*, 6, 23 - 32.
- Ghimire, A., Jaiswal, B., Shrestha, J., Shrestha, K., & Shrestha, S. (2023). COVID-19 and Digitization of SMEs: The Impact and Way Forward. New Perspective: *Journal of Business and Economics*, 6(1), 83 - 90.
- Indika, M. G. N., Perera, H. A. P. K., & Abey Siriwardena, N. K. (2020). How does COVID-19 Influence the Micro Small and Medium Enterprise Sector in Sri Lanka? *Sri Lanka Journal of Advanced Social Studies*, 10(1), 121 - 142. <https://doi.org/10.4038/sljass.v10i1.7160>
- Offerman, T., & Blinde, R. (2022). A Study of Adoption and Effects of DevOps Practices. *ResearchGate*. <https://www.researchgate.net/publication/365486977>
- Omonije, A. (2024). Agile Methodology: A Comprehensive Impact on Modern Business Operations. *International Journal of Science and Research (IJSR)*, 13.
- Pandithasekara, D. N., & Hewage, H. M. (2022). Sri Lankan SMEs amidst the COVID-19 Crisis: Assessing Impact and Combating Strategies. *International Journal of Formal Sciences: Current and Future Research Trends*, 15(1), 85 - 100.
- Pargaonkar, S. (2023). A Comprehensive Research Analysis of Software Development Life Cycle (SDLC) Agile & Waterfall Model Advantages, Disadvantages, and Application Suitability in Software Quality Engineering. *International Journal of Scientific and Research Publications*, 13(8), 120 - 123.
- Quiroz-Vasquez, C. (2024). What is software development? *IBM* <https://www.ibm.com/think/topics/software-development>.
- Surahman, Z. F., Shee, H., Adi, A. S., & Yudaruddin, R. (2023). The effect of digital transformation and innovation on SMEs' performance in times of COVID-19. *Problems and Perspectives in Management*, 21(4), 84 - 100.



A Multi-model Deep Learning System for Intelligent Vehicle Recognition and Market Valuation

Galagamage RA¹, Gunasekara RPTH²

^{1,2}Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka

galagamagerandika@gmail.com¹

ABSTRACT

The used vehicle market often suffers from information gaps and subjective judgments, leading to inaccurate evaluations. This study presents an Artificial Intelligence (AI)-based system that automates vehicle assessment through four specialized intelligent agents: Brand Recognition, Model Recognition, Damage Detection, and Price Prediction, coordinated by a central agent managing workflow and decision-making. Recognition and detection agents use ResNet50 Convolutional Neural Networks (CNNs) with transfer learning to extract image features, while the Price Prediction Agent combines these features with user-provided data (year, fuel type, engine capacity, transmission, mileage) in a deep neural network optimized with Huber loss, batch normalization, and dropout. The Coordinator Agent integrates all outputs to ensure consistency. Users simply upload vehicle images and details, and the system identifies the brand and model, detects damages, and predicts market value. Experimental results manifest high effectiveness: price prediction achieved $R^2 = 0.97$, MAPE = 5.77%, MAE = 397,676 LKR; damage detection resulted an Area Under Curve of 0.94 and 84% accuracy; brand recognition 69.91%; and model recognition achieved an average training accuracy of 61%. These results prove that the proposed multi-agent architecture provides an effective and scalable framework for automating used-vehicle evaluation in real-world market environments.

KEYWORDS: Artificial Intelligence, Deep Learning, Intelligent Agents, Market Value

1. INTRODUCTION

Buying or selling a used vehicle is often uncertain, as pricing usually depends on surface-level checks or personal opinions. This reliance on human judgment causes inconsistency and inaccurate valuations, leading to information gaps in the used vehicle market. With advances in artificial intelligence (AI), especially in deep learning and computer vision, machines can now identify car brands, models, and damage from images. When combined with structured data like engine capacity, fuel type, and mileage, these technologies enable more accurate price predictions. However, most existing systems handle these tasks separately. To overcome this, our research presents an Integrated Multi-

Agent AI System that unites specialized agents for brand recognition, model identification, damage detection, and price prediction under a central Coordinator Agent. This design replicates expert decision-making in vehicle evaluations.

2. LITERATURE REVIEW

The prediction of used car prices has gained attention with advances in AI and machine learning, but most studies focus on separate tasks rather than an integrated vehicle analysis system.

2.1 Car Model Recognition and Brand Identification

Car model recognition is crucial for exact price prediction. Yang et al. (2018)

developed the Artificial Intelligence Blue Book model using transfer learning with a Visual Geometry Group 16-layer network (VGG-16) and Squeeze Net, along with saliency mapping and gradient-weighted Class Activation Mapping (Grad-CAM) to highlight price-relevant image features, but it lacked model classification and structured data integration. Yadav et al. (2021) introduced the Used Car Price Analysis System, combining You Only Look Once (YOLO) and Faster Region-Based Convolutional Neural Network (Faster R-CNN) for vehicle recognition with Random Forest and K-Nearest Neighbors (KNN) for price prediction. Though effective in merging visual and structured data, it omitted damage detection, a key factor in evaluating vehicle condition.

2.2 Vehicle Damage Detection

Accurate damage detection is vital for condition-based price variations. Totare et al. (2023) used Mask Region-Based CNN (Mask R-CNN) to segment and categorize damage areas, visualizing severity with heatmaps, but relied on a small dataset and lacked model recognition or year estimation. Shaikh et al. (2022) applied CNNs with Rectified Linear Unit (ReLU) and pooling layers for damage classification but depended on textual data for brand and model identification. Catruna et al. (2022) combined CNN-based visual detections with structured data like fuel type and mileage, improving predictive accuracy.

2.3 Structured Data-Based Price Prediction

Many studies use structured data for price prediction. Pillai et al. (2022) trained an artificial neural network (ANN) on 18

variables, modeling interactions but excluding images. Bukvic et al. (2022) focused on structured features like mileage and fuel consumption, while Carvalho et al. (2023) combined structured and textual data using co-attention mechanisms, still without visual analysis. Overall, most studies remain limited to isolated data types, overlooking integrated visual and structured approaches for holistic price prediction.

3. METHODOLOGY

3.1 Data Collection and Description

Data for this research had collected from riyasewana.com, a Sri Lankan used vehicle trading platform, using web scraping techniques. The structured dataset contained 4,073 records with 9 attributes: brand, model, edition, year, transmission, fuel type, engine capacity (cc), mileage (km), and price (LKR). Three image datasets were created for computer vision tasks: a brand recognition dataset with 3,277 images from five brands (Honda, Micro, Nissan, Suzuki, Toyota), a model recognition dataset with 3,764 images across six models (Alto, Panda, Vezel, Vitz, WagonR, Xtrail), and a damage classification dataset with 2,603 images labeled as damaged or undamaged. The structured and image datasets were split 80/20 for training and testing, with 10% of the training data used for validation where structured dataset applied five-fold cross-validation on the training set to estimate confidence intervals for performance metrics.

3.2 Data Preprocessing and Feature Engineering

To ensure data quality, records missing price, engine capacity, or mileage

had removed, while missing edition values had replaced with “Standard.” Outliers in price and mileage had handled using the Interquartile Range (IQR) method. Two engineered features, vehicle age and mileage per year, had introduced to strengthen predictive performance. Categorical variables were label-encoded, and numerical attributes had standardized using Standard Scaler.

Images had preprocessed using Image Data Generator for resizing (224×224 pixels), normalization (scaling pixel values to [0,1]), and augmentation (random shear, zoom, and horizontal flips). These transformations improved generalization and reduced over-fitting.

3.3 Multi-Agent System Architecture

The system implements a coordinated multi-agent architecture with four specialized agents managed through a centralized Agent Manager using standardized Agent Request/ Agent Response protocols. The Car Price Prediction Agent employs a neural network with Label Encoders and Standard Scaler, processing structured features with engineered variables. Car Brand Recognition Agent and Car Model Recognition Agent utilize ResNet50-based CNNs for 5-class brand and 6-class model classification respectively, accepting normalized images. Car Damage Detection Agent performs binary classification with confidence scoring. Information flows sequentially: the Manager validated inputs, routes images to vision agents for parallel processing, aggregates predictions using consensus logic (majority voting for brand/model, logical OR for damage), and forwards combined outputs with structured features

to the Price Prediction Agent, enabling modular fault-tolerant operation.

3.4 Model Training

Forward propagation computes weighted linear combinations followed by activation functions, while backpropagation determines weight updates using the Adam optimizer with adaptive learning rate (0.001). Advanced training strategies include early stopping (patience=15), learning rate reduction using Reduce LROnPlateau (factor=0.7).

3.5 Neural Network Architectures

The three computer vision models employ ResNet50-based CNNs using transfer learning from ImageNet pre-trained weights. The input layer processes 224×224×3 RGB images after normalization. Convolution layers with 3×3 filters capture low-level features such as edges and textures. Activation functions use ReLU to introduce non-linearity, enabling complex pattern recognition through $\max(0, z)$ operations. Pooling layers employ max pooling to reduce spatial dimensions while preserving dominant features and minimizing computational complexity. Deeper layers progressively extract higher-level automotive details like headlights, grills, and brand logos. The fully connected layers map these learned features to outputs, where softmax handles multi-classification and sigmoid supports binary assessment.

The Price Prediction Agent uses a feed-forward DNN with ReLU activations, batch normalization, dropout (20% and 10%), and a linear output layer has improved by using Huber loss optimization for outlier robustness.

4. RESULTS AND DISCUSSION

4.1 Computer Vision Models Performance

The Damage Detection Agent achieved balanced performance, reporting 83.5% accuracy and, for the damaged class specifically, a precision of 80%, a recall of 89%, and an F1-score of 84%. Its conservative bias toward detecting damage is particularly beneficial in insurance applications, where the cost of false negatives outweighs that of false positives.

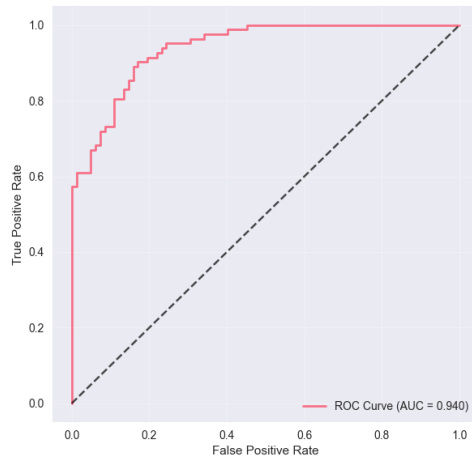


Figure 1: ROC Curve for Damage Detection Model

Fig. 1 shows the Receiver Operating Characteristic (ROC) curve of the Damage Detection Agent, with an Area Under Curve (AUC) of 0.94, indicates excellent distinction between damaged and undamaged vehicles, high accuracy, and minimal false positives. The Brand Recognition Agent recorded an overall accuracy of 69.91%, performing exceptionally well on the dominant Suzuki class, with 81% precision and 85% recall. Meanwhile, the Model Recognition Agent achieved 60.95% training accuracy in fine-grained car model classification. Models like Alto, Panda, Vezel, Vitz, and WagonR performed better due to more balanced

datasets. The weighted precision (0.1426), recall (0.1529), and F1-score (0.1287) reflect class imbalance. In particular, the Nissan X-Trail, which is representing a minority vehicle category in the Sri Lankan used-car market, had significantly fewer advertisement samples. This scarcity of training data limited the model's ability to learn distinctive visual and structural features, reducing prediction reliability for this class. Increasing dataset diversity and balance is essential to improve recognition accuracy and overall system performance.

4.2 Price Prediction Model Excellence

The Price Prediction Agent demonstrated exceptional performance with $R^2 = 0.9717$, while denote 97% variance explanation in price predictions. The Mean Absolute Percentage Error (MAPE) of 5.77%, Mean Absolute Error (MAE) of 397,676 LKR and Root Mean Squared Error (RMSE) of 675,812 LKR maintain high practical accuracy. The model effectively combines automated vision-based features with user inputs to produce reliable market valuations.

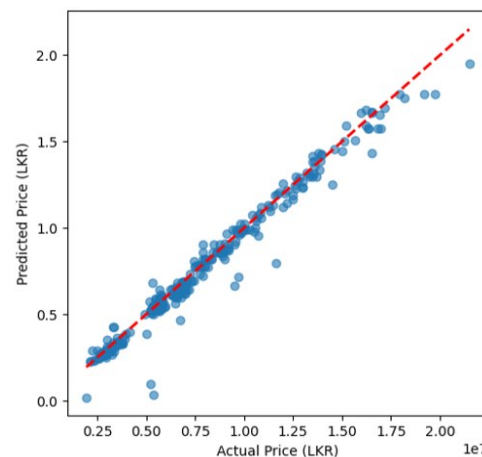


Figure 2: Predication Vs Actual Values

The scatter plot in Fig. 2 shows a strong linear relationship between actual and predicted vehicle prices, with most points close to the perfect prediction line, confirming the model's accuracy.

For comparison, a classical machine learning linear regression model has evaluated, achieving moderate performance with R^2 of 0.54, MAE of 1.98 million LKR, a RMSE of 2.75 million LKR, and a MAPE of 29.4%. To improve accuracy, then the deep learning-based regression model has applied significantly outperforming the classical method. The results show that combining visual and structured features in a deep learning model greatly improves prediction accuracy over classical methods. High R^2 and low errors confirm strong generalization, supporting its use in automated vehicle valuation.

5. CONCLUSION

This research successfully demonstrates the viability of integrated artificial intelligence systems with agent-based coordination for comprehensive vehicle analysis, achieving breakthrough performance in automated price prediction (97% accuracy) through sophisticated neural network architectures. The combination of CNN-based computer vision models with deep neural network regression, orchestrated through intelligent AI agents, creates a powerful framework for objective vehicle assessment. Future work should expand the training dataset for better classification, adopt hierarchical methods, improve damage assessment, and strengthen coordination between AI agents. The established neural network framework offers a robust foundation

while providing immediate value through its key components.

REFERENCES

- Bukvić, L., Škrinjar, J. P., Fratrović, T., and Abramović, B. (2022). Price Prediction and Classification of Used-Vehicles Using Supervised Machine Learning. *Sustainability*, 14(24), 17034. <https://doi.org/10.3390/su142417034>
- Carvalho, G., Viana, D., Barbosa, L., and Ren, T. I. (2023). Combining structured and unstructured data using co-attention for car price prediction, *Procedia Computer Science*, 222(C), 646–655. <https://doi.org/10.1016/j.procs.2023.08.202>
- Catruna, A., Betiu, P., Tertes, E., Ghita, V., Radoi, E., Mocanu, I., and Dascalu, M. (2022). Car Full View Dataset: Fine-grained car orientation prediction. *Electronics*, 12(24), 4947. <https://doi.org/10.3390/electronics12244947>
- Pillai, A. S. (2022). A Deep learning approach for used car price prediction. *Journal of Science and Technology*, 3(3), 31 – 50.
- Shaikh, M. K., Zaki, H., Tahir, M., Khan, M. A., Siddiqui, O. A., and Rahim, I. U. (2022). Framework for car price prediction and damage detection. *Pakistan Journal of Engineering and Technology*, 5(4), 52 – 59. <https://doi.org/10.51846/vol5iss4pp52-59>
- Totare, R., Bhalsing, V., Lende, M., Maramwar, T., and Naikwadi, C. (2023). Car damage detection and price prediction using CNN. *International Research Journal of Engineering and Technology (IRJET)*, 10(1), 458-461.
- Yadav, A., Kumar, E., & Yadav, P. K. (2021). Object detection and used car price predicting analysis system (UCPAS) using machine learning technique. *Linguistics and Culture Review*, 5(2), 1131-1147. <https://doi.org/10.21744/lingcure.v5nS2.1660>
- Yang, R. R., Chen, S., and Chou, E. (2018). *AI Blue Book: Vehicle price prediction using visual features*. ArXiv <https://doi.org/10.48550/arXiv.1803.11227>



Adapting to Remote Working: A Case Study on a Multinational Workplace

Sewwandi PHAJ¹, Munasinghe B²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

jayanisewwandi99@gmail.com¹, bhagya@wyb.ac.lk²

ABSTRACT

This case study examines how an Australia-based multinational software development company adapts its operations to support effective remote work across teams in Australia, Singapore, and Sri Lanka. Using a qualitative approach involving six months of participant observation, formal documentation, and interviews with four employees, the study investigates how adjustments to working hours, coordination strategies, communication systems, and task management processes shape productivity and employee experience. Thematic analysis identified four key themes: remote working practices, adaptation to advanced time zone schedules, communication and coordination systems, and productivity management. Overall, the findings show that structured communication, disciplined workflow management, and clear procedures enable productive collaboration across time zones, although early-hour work schedules present notable adaptation challenges.

KEYWORDS: Communication, Coordination, Productivity, Remote Work, Time Zone Adaptation

1. INTRODUCTION

The COVID-19 pandemic accelerated the global shift towards remote working environments, leading organizations to redesign traditional work structures to maintain efficiency and collaboration. Remote working in the Information Technology (IT) and Software Engineering (SE) sectors has not only reshaped how tasks are performed but also transformed team composition, enabling highly multicultural and geographically distributed teams. This global reach offers significant advantages, such as diverse problem-solving approaches, broader innovation, and access to a worldwide talent pool (Sarker & Sahay, 2004).

The company subjected to this study was headquartered in Victoria, Australia, is a Managed Service Provider offering IT support, web development, cybersecurity, and automation services. It has branches in Singapore and Sri Lanka. This case

study presents the experience of employees who are physically located in Sri Lanka, who are 4.5 hours behind the base company time zone. The company operates under a fully remote working model in its cross-border collaboration with the Sri Lankan branch.

This case study examines the company's remote work system, specifically its scheduling practices, communication structures, and performance management mechanisms affect Sri Lankan employees. Unlike prior studies focusing broadly on remote work, this paper provides insight into time-zone adaptation within an entry-level workforce in a multinational IT environment.

This case study contributes to the limited empirical focused area of how extreme time-zone adaptation affects junior-level employees in a multinational IT context.

2. LITERATURE REVIEW

Remote working has become a transformative aspect of the Information Technology and Software Engineering-related industries around the globe, offering both flexibility and productivity gains (Ford et al., 2021). Employees benefit from reduced commuting time and improved work-life balance, while organizations leverage cost savings and access to a global talent pool. Research indicates that remote software teams can maintain or even enhance productivity when supported by effective communication tools and agile methodologies (Ford et al., 2021). Distributed teams also provide access to diverse skill sets and international markets (Sarker & Sahay, 2004).

The challenges of working remotely persist, including collaboration difficulties, time zone discrepancies, and the risk of employee isolation, decreased team cohesion, and lower engagement without structured communication channels (Ralph et al., 2020). Multicultural teams operating remotely face unique challenges, particularly in communication, cultural expectations, and coordination across time zones. Cultural differences can affect collaboration styles, decision-making processes, and even perceptions of hierarchy and deadlines (Hinds, Liu, & Lyon, 2011). Asynchronous communication caused by time zone gaps can delay feedback loops and responses, and require additional planning demands that complicate agile practices such as daily stand-ups, and require deliberate planning to maintain productivity and team cohesion (Herbsleb & Mockus, 2003).

The rapid adoption of remote work, accelerated by the COVID-19 pandemic, has prompted companies to invest in digital infrastructure and redefine team dynamics to sustain performance in distributed environments. Moreover, it intensified remote work adoption, prompting organizations to reevaluate workflows and invest in collaboration tools (Bao et al., 2020). Overall, while remote working introduces new complexities, it aligns well with the inherently digital nature of IT and SE roles.

3. METHODOLOGY

3.1 Research Design

A qualitative case study approach was used to understand employees' real-world experiences within a natural organizational setting.

This case study adopts a pragmatic approach based on the researcher's firsthand experience during a six-month undergraduate training at the company. Data were collected through: 1) formal reports including daily task records and workflow guidelines, 2) personal observation notes by the researcher during her training period, and 3) semi-structured online interviews with four participants of junior-level and trainee staff members.

The interviews were conducted in English and were transcribed manually. Participation was voluntary, interviewees provided verbal informed consent, no identifying information was disclosed, and all data were stored securely and used only for academic purposes.

The data were qualitatively analyzed using thematic analysis, which involved:

1. Generating preliminary themes.
2. Merging, refining, or removing codes through iterative comparison.

3. Validating themes using cross-checking between interviews and documented observations.

This process ensured reliability and consistency in theme development.

4. RESULTS AND DISCUSSION

Complying well with the findings in literature, the themes that emerged from this case study include mode of working, adapting to schedules, communication, managing performance, and productivity in a remote setting. The four major themes that emerged from the analysis are summarized in Table 1.

Table 1: Summary of Emergent Themes

Theme	Description
Mode of Working	Fully remote setup supported by collaboration software; focus on independent task execution.
Adaptation to Schedules	Sri Lankan employees follow a 4:00 a.m.–1:00 p.m. schedule aligned with Australia.
Communication & Coordination	Structured daily meetings, monthly CEO reviews, and shared Excel-based task tracking.
Productivity Management	Use of Trello, Teams, daily reports, and attendance monitoring to ensure workflow consistency.

The following subtopics discuss the findings related to the case of interest in this study.

4.1 Mode of Working

ITS Gippsland operates entirely remotely, with employees working from their private locations using software

tools such as Microsoft Teams, Trello, and Google Workspace. Employees manage tasks, review progress, and plan collaboratively, ensuring accountability and team coordination. The model offers flexibility, reduces commuting stress, and supports focused, independent work, though no additional allowances (e.g., internet reimbursement, equipment support) are provided.

Daily routines follow a structured pattern: employees report previous-day tasks, discuss blockers, and receive instructions for the current day. This enhances accountability and synchronization while allowing employees to benefit from flexible, home-based work environments.

4.2 Adaptation to Schedules

The standard working hours practices adapted by the company, which were 8.30 a.m. to 4.30 p.m. local time, have been aligned to normal Sri Lankan working hours. This was changed to align with the Australian time zone to enhance real-time collaboration with clients and the main office during the researcher's period of tenure, allowing her to experience the effects firsthand. The new work schedules shifted to 4.00 a.m. to 1.00 p.m. local time, a notable 4.5-hour advancement to very early hours in the local setting.

This shift improved real-time communication but introduced several challenges. It intends to ensure real-time communication with the main branch during their office hours, synchronized task progress, especially for client support and website updates, and improved productivity in overlapping time frames. The adaptation to the early working hours was observed in two perspectives; 1) personal and 2) organizational.

Personal Perspective

Both interview and observation data indicated four aspects in personal perspective: 1) disrupted sleep cycles, 2) adjustment to daily routines and family life, 3) Extended working hours due to unfinished tasks (i.e. previous day tasks reporting and next day task identification and prioritizing), and 4) potential risk of burnout and reduced work–life balance.

The first three aspects demand a physical need with the potential to raise adverse health effects. The latter aspects have indirect effects on employees' daily routines. Having the official work hours ending closer to mid-day has the potential for the employee to continue working unless they force themselves, which seldom happens. This creates problems in maintaining work-life balance in a local setting. Moreover, such efforts would not be considered in the employee's payment structure. The extreme time zone balances of this nature can therefore have the potential to demand serious compromising of an employee's physical and mental conditions.

Organizational Perspective

The organizational perspective indicated that the company adopted measures such as: 1) updated workflow timelines, 2) flexibility for emergencies, 3) automated attendance and productivity tracking, and 4) increased supervisor engagement. These adaptations supported business processes in planning, HR, and task management, though long-term financial impacts remain unobserved.

The company also had to emphasize the well-being and communication to ensure team members could adapt comfortably to the new time structure

without burnout. These aspects directly influenced the business process in key areas such as planning, human resource management, and systems administration. Given the timeline of this case study, the effect of the financial process could not be recorded or observed.

4.3 Communication and Coordination System

Effective communication was critical in this cross-border environment. The need and dependency of business collaboration tools are very high in a remote working environment. Gippsland uses Microsoft Teams as the collaboration software for routine daily meetings of approximately 15-20 minutes with lower-level employees, led by their immediate supervisor. This routine ensured task clarification, early detection of issues, and shared responsibility and transparency.

A meeting with the CEO of ITS Gippsland is held at the beginning of every month. The purpose of this session is to 1) review monthly performance and goals, 2) discuss strategic improvements, and 3) encourage open communication between management and employees. These monthly meetings provide an opportunity for the lower-level employees to be exposed to business matters beyond their routine work, as well as make them feel inclusive in a business setting which is otherwise quite impersonal and individual. This structured communication also ensures alignment between operational tasks and the company's strategic vision.

For task and attendance management, the company used a shared Excel workbook. Employees recorded daily tasks, pending items, and sign-in/out times. This simple system ensured

transparency without the complexity of advanced tracking platforms. Leave procedures require emailing HR, the supervisor, and the CEO five days in advance, except in emergencies handled by direct supervisor communication.

4.4 Managing Productivity in Remote Settings

Managing productivity is seen as another theme that emerged in the study. The company employs several mechanisms to ensure consistent productivity, namely: 1) clear task allocation during daily meetings, 2) use of collaborative tools (e.g. Trello for progress tracking and Google Workspace for document sharing), 3) daily progress reporting and Excel-based attendance monitoring, and 4) an effective communication culture emphasizing accountability.

This structured approach fosters a high level of accountability among team members and creates a well-organized framework for collaboration, thereby enabling the remote working environment to function with a level of efficiency, coordination, and productivity comparable to, if not exceeding, that of a traditional in-office setting.

The study found ITS Gippsland's remote work model strengths include strong communication, high supervisor-employee trust, efficient time use from early scheduling, and clear coordination. Challenges include initial difficulty in adapting to the Australian time zone, disrupted sleep, and limited social interaction due to virtual-only collaborations. Employees reported gradual adaptation supported by team familiarity and routine reinforcement.

5. CONCLUSION

Although remote work offers flexibility and technological support, fully remote operations remain challenging. Cross-border teams must address significant challenges, particularly when working across wide time zones. Success relies on clear communication, adaptable schedules, and cultural sensitivity to foster collaboration and inclusivity across locations.

ITS Gippsland's experience shows that structured communication, disciplined workflow practices, and effective coordination can sustain productivity in a global virtual environment. Structured procedures for meetings, attendance, and leave ensure accountability. The 4.5-hour advanced schedule enhances collaboration with the Australian head office. However, Sri Lankan employees reported physical and mental strain. While current impacts are visible in planning, HR operations, and administrative adjustments, long-term outcomes related to employee well-being and organizational sustainability are yet to be fully observed. This case study contributes practical insights into how multinational companies can adapt remote working systems to maintain productivity while safeguarding employee well-being in geographically dispersed settings.

REFERENCES

- Bao, L., Li, T., Xia, X., Zhu, K., & Zhang, L. (2020). How does working from home affect developer productivity? A case study of Baidu during COVID-19. *Empirical Software Engineering*, 25(6), 4927 - 4961.
- Ford, D., Storey, M.-A., Zimmermann, T., Bird, C., Jaffe, S., & Butler, J. (2021). A Tale of Two Cities: Software Developers Working from Home During the COVID-19 Pandemic. *ACM Transactions on*

Software Engineering and Methodology (TOSEM), 31(2), 1 - 37.

- Herbsleb, J. D., & Mockus, A. (2003). An empirical study of speed and communication in globally distributed software development. *IEEE Transactions on Software Engineering*, 29(6), 481 - 494. <https://doi.org/10.1109/TSE.2003.1205177>
- Hinds, P. J., Liu, L., & Lyon, J. (2011). Putting the global in global work: An intercultural lens on the practice of cross-national collaboration. *Academy of Management Annals*, 5(1), 135 - 188. <https://doi.org/10.5465/19416520.2011.586108>
- Ralph, P., Baltes, S., Adisaputri, G., Torkar, R., Kovalenko, V., Kalinowski, M., et al. (2020). Pandemic Programming: How COVID-19 Affects Software Developers and How Their Organizations Can Help. *Empirical Software Engineering*, 25, 4927 - 4961. <https://doi.org/10.1007/s10664-020-09875-y>
- Sarker, S., & Sahay, S. (2004). Implications of space and time for distributed work: An interpretive study of US–Norwegian systems development teams. *European Journal of Information Systems*, 13(1), 3 - 20. <https://doi.org/10.1057/palgrave.ejis.3000485>



AI Fish Identification System

Perera SRV¹, Jayakody JRKC²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

shalinirvithanage@gmail.com¹, kithsirij@wyb.ac.lk²

ABSTRACT

This research examines fish species using machine learning and image processing techniques. The main reason for this study is that many existing systems cannot identify fish correctly, when the image background, light condition, or fish position is different. Most previous works do not give stable or consistent results for local fish species, especially when the datasets are small or the images are taken under natural conditions. To solve this problem, this research develops a Convolutional Neural Network (CNN) model as a practical solution. Total 1222 original fish images were collected by using a mobile camera, and the dataset was doubled to 2444 images using augmentation methods. The specific augmentation methods were: Horizontal Flip, which flips randomly 50% of images for adding variation; Rotation, where each image was rotated randomly between -10° and $+10^\circ$ for reducing orientation sensitivity; Brightness Adjustment, where pixel values were multiplied by random factors from 0.8 to 1.2 for handling different light conditions; and Gaussian Blur, applied with random sigma values from 0 to 1.0 for increasing resistance to noise and focus problems. These augmentation steps helped the model to be more stable with different image resolutions (4:3, 16:9, and square). Three custom CNN variations were designed and tested to identify certain fish species, and the best accuracy, 95.03%, was obtained from the model which has two convolutional layers. The results confirm that a well-designed CNN structure can classify fish species with high accuracy and reliability.

KEYWORDS: Convolutional neural network, Data augmentation, Fish species identification, Image processing, Machine learning

1. INTRODUCTION

Fish identification has a very important role in many areas such as biodiversity study, fisheries management, and food quality checking. In many developing countries, including Sri Lanka, identifying fish species is still mostly done by human observation, which depends on expert knowledge and visual judgment. This manual process takes much time and often makes mistakes when fish have similar shape or color. So, there is a growing need for an automatic and reliable method to identify fish species correctly. Recent progress in machine learning and image processing gives new possibilities to solve these

classification problems. Traditional image processing methods depend much on manual feature extraction, which is not always good because image conditions can change, like light, rotation, or background. Many previous research also use small datasets or study foreign fish species, so they are less useful for local or practical cases. Because of this, these models cannot get high accuracy or work well when new images are given. To fix these problems, this study suggests a deep learning method using Convolutional Neural Networks (CNNs) for identifying fish species. CNNs can learn and extract important image features automatically, reducing need for manual work. Fish identification has a very important role in

many areas such as biodiversity study, fisheries management, and food quality checking. In many developing countries, including Sri Lanka, identifying fish species is still mostly done by human observation, which depends on expert knowledge and visual judgment. This manual process takes much time and often makes mistakes when fish have similar color. So, there is a growing need for automatic and reliable methods to identify fish species correctly.

Recent progress in machine learning and image processing gives new possibilities to solve these classification problems. Traditional image processing methods depend much on manual feature extraction, which is not always good because image conditions can change, like light, rotation, or background. Many previous research also use small datasets or study foreign fish species, so they are less useful for local or practical cases. Because of this, these models cannot get high accuracy or work well when new images are given.

To fix these problems, this study suggests a deep learning method using Convolutional Neural Networks (CNNs) for identifying fish species. CNNs can learn and extract important image features automatically, reducing need for manual work. In this research, a dataset of 1222 original fish images was collected with a mobile camera, and it was made bigger by several augmentation methods: horizontal flip, rotation, brightness adjustment, and Gaussian blur to double the dataset to 2444 images. These augmentations help the model to be more stable with different light, camera resolution, and image angle. Three custom CNN models were designed and trained to test their performance in fish species classification.

The model with two convolution layers reached highest accuracy 95.03%, mobile, or lightweight solutions for fish identification to support real-world use. Priyankan and colleagues (2019) made a YOLO-based mobile application for 16 Sri Lankan species and showed that on-device species shape or showing that well-optimized CNN can classify fish very well even with a small dataset. This study shows deep learning potential to automate fish identification and supports making efficient, low-cost recognition systems for real-world use. Research, a dataset of 1222 original fish images were collected with a mobile camera, and it was made bigger by several augmentation methods: horizontal flip, rotation, brightness adjustment, and Gaussian blur to double the dataset to 2444 images. These augmentations help the model to be more stable with different light, camera resolution, and image angle.

Three custom CNN models were designed and trained to test their performance in fish species classification. The model with two convolution layers reached highest accuracy 95.03%, showing that well-optimized CNN can classify fish very well even with a small dataset. This study shows deep learning potential to automate fish identification and supports making efficient, low-cost recognition systems for real-world use.

2. LITERATURE REVIEW

Automated fish recognition using deep learning first became possible when researchers applied convolutional neural networks (CNNs) on curated and well-labelled image sets and reported promising classification results (Rauf, 2019). These early works showed that CNNs can learn important visual patterns

without manual feature design, but most experiments used controlled images and quite large datasets, so it is still unclear how well such models work with mobile-camera photos taken in different field conditions (Rauf, 2019).

After that, several studies tried practical, mobile, or lightweight solutions for fish identification to support real-world use. Priyankan and colleagues (2019) made a YOLO-based mobile application for 16 Sri Lankan species and showed that on-device species detection is possible, but their work focused more on detection and app features than systematically testing small classification networks trained only on mobile images (Priyankan et al., 2019). Similarly, other region-focused projects showed that local datasets and mobile apps can be made, but they usually depend on large annotated sets or detection frameworks instead of small classifiers for low-resource settings (Priyankan et al., 2019; Shah, 2023).

Comparisons of deep transfer models and lightweight networks show a trade-off between depth and practicality. Transfer learning with deep backbones (like VGG, ResNet, MobileNet) often reaches high accuracy, but fine-tuning deep networks can cause overfitting on small datasets and requires more computational resources than shallow custom models (Abu et al., 2022; Sturekova, 2025). This suggests that for small, local image collections captured with mobile phones, well-designed shallow CNNs with careful regularization and augmentation can compete with heavy transfer models (Abu et al., 2022).

Reviews of CNN design and best practices emphasize that architecture choices (number of convolution layers, pooling, batch normalization, dropout)

strongly influence generalization when training data are small (Alzubaidi et al., 2021). These surveys support trying multiple compact CNN designs instead of assuming one deep model works well for all datasets (Alzubaidi et al., 2021).

3. METHODOLOGY

3.1 Data Collection and Dataset Characteristics

A dataset was created to automatically identify two local fish types found in Sri Lanka, called “Gal Malu” and “Lenapara”. A total of 1,222 images were collected, where 578 belong to “Gal Malu” and 644 to “Lenapara”. Photos were taken using different aspect ratios such as 4:3, 16:9, and 1:1 to make some variation in light and image framing. After collecting, the dataset was divided into 1,955 images for training and 489 for validation, and both fish types were kept in balance for better model learning.

3.2 Image Augmentation Techniques

We applied several augmentation methods to increase data diversity and reduce overfitting. Horizontal flip was used with 50% chance; rotation ranged from -10° to $+10^\circ$; brightness was scaled between 0.8 and 1.2; Gaussian blur (σ 0–1.0), zoom, and shear were also applied. After augmentation the dataset size doubled to 2,444 images. This augmentation helped the model to learn despite differences in orientation, lighting, camera distance, and perspective.

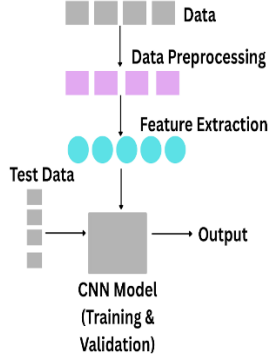


Figure 1: Overview of Model Development

3.3 Color and Contrast Preprocessing

All images were resized to 224×224 pixels and normalized from 0 to 1 for training purposes. Before starting the training, I checked the images visually to see some color and contrast differences between “Gal Malu” and “Lenapara”. These steps made the dataset more uniform and suitable to use in the CNN model.

3.4 Model Development

3.4.1 CNN Architecture

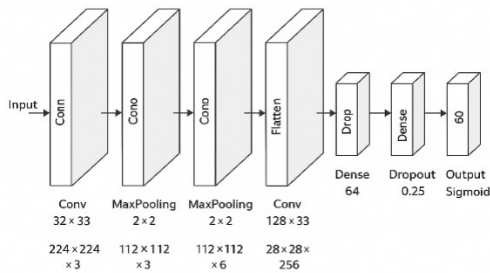


Figure 2: CNN Architecture

In this research, a custom CNN model was designed to classify the two fish types. The network has four convolution blocks with 32, 64, 128, and 256 filters, and each block uses 3×3 kernels with ReLU activation. MaxPooling is added

after every block to reduce the image size step by step. A 25% dropout layer is also used to control overfitting because the dataset is small. After the convolution layers, the features are flattened and passed to a dense layer with 64 neurons, and finally a sigmoid layer is used to give the output. Compared to a normal CNN, this model is lighter and uses fewer layers and regularization to perform better on mobile-captured fish images. The whole model has about 11.46 million parameters.

3.4.2 Training and Validation

The CNN model trained using Adam optimizer with binary cross-entropy loss. Training used batch size 32. Early stopping with patience 5 and model checkpointing applied to avoid overfitting and keep best weights. Augmented images were used for training, but validation images only normalized to evaluate generalization.

4. RESULTS

Each CNN variant, comprising 1, 2, and 3 convolutional layers, was evaluated separately using the same dataset. Performance metrics including validation accuracy, precision, recall, F1-score, and confusion matrices were computed for each model. The 1-layer CNN achieved a validation accuracy of 88.45%, showing lower performance due to limited feature extraction. The 2-layer CNN performed significantly better, reaching 92.86% validation accuracy with balanced precision and recall (F1-score of 0.93 for both classes), representing the best trade-off between accuracy and model complexity. The 3-layer CNN slightly improved validation accuracy to 93.20% but required higher computational

resources, while achieving a comparable F1-score (0.93–0.94).

Table 1: Overall Performance Matrix

No of Convolutional Layers	Validation Accuracy (%)
1	88.45
2	92.86
3	93.20

Overall, the two-layer CNN provides an optimal balance of efficiency and performance for this dataset.

Table 2: Precision / Recall / F1-Score

No. of Conv Layers	Precision (galmalu/lenapara)	Recall (galmalu/lenapara)	F1-score (galmalu/lenapara)
1	0.89 / 0.88	0.88 / 0.89	0.88 / 0.89
2	0.93 / 0.93	0.93 / 0.93	0.93 / 0.93
3	0.93 / 0.94	0.93 / 0.94	0.93 / 0.94

5. CONCLUSION

This study shows that a customized CNN can identify the two local Sri Lankan fish types, “Gal Malu” and “Lenapara”, with good accuracy even when the dataset is small and images are taken under natural mobile-camera conditions. By using targeted augmentation methods such as rotation, brightness changes, blur, and flipping, the model became more stable against common real-world variations like lighting, background, and image angle. Three CNN models with different numbers of convolution layers were tested, and the model with three layers achieved the highest final validation accuracy of 93.20%, while the two-layer model also performed well during early training.

These findings suggest that a shallow and carefully optimized CNN can work effectively without needing very deep or complex architectures. The results also highlight those simple models, combined with proper preprocessing and augmentation, can be used to build low-cost and practical fish identification tools for local applications. However, the study is limited to only two fish species, so future work should include more species and compare different model families to improve generalization and performance.

REFERENCES

- Abu, R., Hamid, A., & Kamarudin, R. (2022). Comparative study of deep and shallow CNNs for small-scale fish image classification. *Journal of Computer Vision and Applications*, 15(3), 145 - 160.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, Article 53. <https://doi.org/10.1186/s40537-021-00444-8>
- Priyankan, D., Perera, L., & Fernando, R. (2019). YOLO-based mobile application for on-device fish species detection in Sri Lanka. *Proceedings of the International Conference on Machine Learning Applications*, 102 - 110.
- Rauf, H. (2019). Convolutional neural networks for automated fish species classification. *Journal of Marine Informatics*, 7(2), 55 - 67.
- Shah, P. (2023). Regional datasets and mobile fish recognition: Challenges and opportunities. *International Journal of Applied AI Research*, 10(1), 25 - 38.
- Sturekova, I. (2025). Transfer learning and lightweight CNNs for mobile image classification. *Computers & Electrical Engineering*, 105, 108601.



Automated Plastic Waste Classification and Recommendation

Malsinghe MPPN¹, Gunasekara RPTH²

^{1,2}*Department of computing & Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

pradee.malsinghe@gmail.com¹

ABSTRACT

The escalating global plastic waste crisis demands innovative waste management solutions to address inefficiencies in recycling processes. A key challenge lies in accurately identifying diverse plastic types and their contamination levels to provide context-aware recycling guidance. Image recognition is utilized because images capture rich visual details of plastic waste, such as type, cleanliness, and structure, that are essential for precise classification. Compared to manual sorting or chemical analysis, image-based classification using Convolutional Neural Networks (CNNs) offers a scalable, non-destructive, and faster method for handling large volumes of waste with improved accuracy. The novelty of this approach lies in integrating CNN-based visual classification with an ontology-driven knowledge base that leverages logical inference to generate nuanced recycling recommendations, including “Recycle”, “Wash and Recycle”, “Not Recycle”, and “Reuse”. This combination enables context-sensitive decision-making that adapts to varying plastic conditions and recycling requirements, an advancement over traditional rule-based or purely data-driven methods. The system achieved an 85% classification accuracy, outperforming many existing techniques, and demonstrated that ontology-driven reasoning significantly enhances adaptability and decision making. This AI-powered solution offers notable improvements in recycling efficiency and plastic waste reduction by delivering intelligent, precise, and actionable recycling guidance.

KEYWORDS: Convolutional Neural Network, Ontology, Plastic Waste Management, Recycling Recommendation

1. INTRODUCTION

The global plastic waste crisis has become one of the most urgent environmental challenges of the 21st century, threatening ecosystems, wildlife, and human health. Although efficient waste management is crucial, existing sorting and recycling processes remain largely manual, error-prone, and inefficient, resulting in low recycling rates and high contamination levels.

Recent advances in artificial intelligence have enabled image-based plastic classification using Convolutional Neural Networks (CNNs). However, most AI-driven systems focus solely on visual identification and fail to account for contextual factors, such as contamination, item condition, or local recycling policies.

This limitation leads to generic and often impractical recycling recommendations. To address these issues, this study proposes a hybrid waste management system that integrates CNN-based visual classification with an ontology-driven knowledge base employing rule-based logical inference. This combination enables accurate identification of plastic type and condition while generating context-aware recycling recommendations. The framework bridges data-driven perception with semantic reasoning, offering a scalable, intelligent, and adaptable solution to enhance recycling efficiency and support global waste reduction efforts.

2. LITERATURE REVIEW

Early research in smart waste management, such as Sinha and Couderc (2013) and Glouche and Couderc (2013), introduced the concept of “smart bins” that could interact with waste items containing embedded information, laying the groundwork for data-driven disposal systems.

With the emergence of computer vision, recent studies by Soundarya et al., Al-Mamun et al. (2023), and Kruthika et al. (2024) have shown that Convolutional Neural Networks (CNNs) are highly effective for automated waste classification, achieving strong accuracy across multiple waste types.

Meanwhile, Dimitrijevic (2022) highlighted the role of ontologies in structuring knowledge about waste materials and their relationships, enabling logical reasoning and interoperability in decision-making systems.

Despite these advances, most existing models focus only on identifying waste types without providing contextual recycling guidance. This study addresses that gap by combining CNN-based visual classification with a Protégé ontology to provide intelligent, rule-based recycling recommendations tailored to specific plastic conditions.

3. METHODOLOGY

The system consists of two core components: a CNN classifier and a Protégé knowledge base. The CNN classifies plastic waste by type (e.g., PET, HDPE), cleanliness, and shape structure based on a custom-labeled dataset. The classification output feeds into the Protégé knowledge base, which uses SWRL rules to infer recycling actions. Rule examples include recommending

“Reuse” for highly clean PET with good shape integrity and “WashAndRecycle” for moderately clean items. The integrated application enables users to upload images and receive personalized treatment recommendations.

3.1 CNN Model for Plastic Classification

Dataset and Training: A specialized dataset was curated, incorporating various lighting conditions, degrees of degradation, and common contaminants (e.g., residue). The dataset contained images of five primary plastic types: PET (Polyethylene Terephthalate), HDPE (High-Density Polyethylene), LDPE (Low-Density Polyethylene), PP (Polypropylene), and PS (Polystyrene). The dataset was split into 80% for training, 10% for validation, and 10% for testing. Transfer learning was used, with the model based on the VGG16 architecture was initialized with weights pre-trained on ImageNet and fine-tuned using the plastic waste image dataset. Training involved standard augmentation techniques (rotation, scaling, brightness adjustment) to enhance model robustness. The model was trained using the Adam optimizer and a categorical cross-entropy loss function over 50 epochs.

The classification task is a multi-class problem with 45 unique class labels, derived from the cross-product of the 5 Plastic Types, 3 Clean Levels (High, Medium, Low), and 3 Shape Structures (High, Medium, Low) to enable context-aware reasoning by the ontology.

3.2 Ontology SWRL Rules for Treatment Recommendation

The knowledge base incorporates SWRL (Semantic Web Rule Language) rules to encode expert decision logic for

recommending treatments based on combinations of plastic type, cleanliness, and shape structure. This constitutes the system's semantic reasoning component. The ontology, developed using Protégé, defines the domain using several key classes:

- Core Item: PlasticItem (the object instance).
- Feature Classes: PlasticType (PET, HDPE, LDPE, PP, PS), CleanLevel (High, Medium, Low), and ShapeStructure (High, Medium, Low).
- Action Class: Treatment (Reuse, Recycle, WashAndRecycle, NotRecycle).

These classes are linked by Object Properties such as hasType, hasCleanLevel, and hasShapeStructure, which receive the CNN's classification output as facts. The rule for reusing highly clean PET with high shape integrity is formally expressed as:

```
PlasticItem(?p) ^hasType(?p,
PET)^hasCleanLevel(?p,HighCleanLevel
) ^
hasShapeStructure(?p,HighShapeStructur
e)->hasTreatment(?p, Reuse)
```

PET Rules

- If PET plastic is highly clean and has high shape integrity, then treatment is “Reuse”.
- If PET plastic is highly clean and has medium or low shape integrity, then treatment is “Recycle”.
- If PET plastic is medium clean and has high or medium shape integrity, then treatment is “Wash and Recycle”.
- If PET plastic is medium clean and has low shape integrity, then treatment is “Not Recycle”.
- If PET plastic is low clean (any shape), then treatment is “Not Recycle”.

```
# PET Rules (Existing)
Imp().set_as_rule(
f"PET(?p) ^ hasCleanLevel(?p,
{high_clean_instance.iri}) ^
hasShapeStructure(?p,
{high_shape_instance.iri}) ->
hasTreatment(?p, {reuse_instance.iri})"
)
```

Similar SWRL rules have been defined for other common plastic types such as HDPE, LDPE, PP, and PS, using the same treatment logic based on cleanliness and shape-integrity criteria.

3.3 System Integration and Workflow

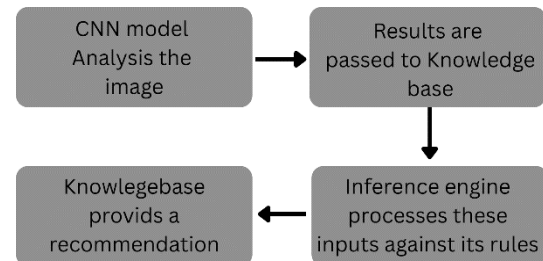


Figure 1: System Integration

The user-facing application initiates the workflow when an image is uploaded. The image is first processed by the CNN model. The CNN's final softmax layer produces a probability distribution across the 45 unique classification classes. The class with the highest probability is selected to determine the classification results: Plastic Type (e.g., PET), Clean Level, and Shape Structure.

These classification outputs are then passed as facts to the Protégé knowledge base, whose inference engine processes them against its set of logical rules to deduce a single, contextually appropriate, and actionable recommendation (e.g., "Recycle" or "Reject"), which is then presented directly to the user.

This workflow highlights a powerful design pattern in AI systems, known as

hybrid AI, where a deep learning model (the CNN) handles complex perception tasks, and a symbolic AI system (the Protégé knowledge base/ontology) performs logical reasoning and rule-based decision-making. This combination ensures both high accuracy in image recognition and explainability in the final recommendation.

The following figure illustrates the final system output presented to the user, where the model classifies the plastic and recommends the appropriate treatment in this case, “Wash and Recycle”.

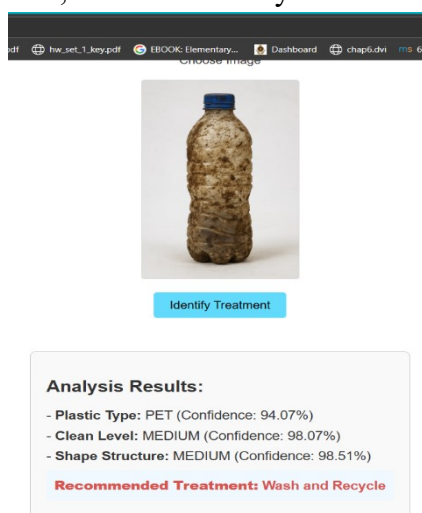


Figure 2: Final result for a PET bottle

4. EVALUATION

The performance of the proposed hybrid plastic waste classification and recommendation system was evaluated using the dataset. The Convolutional Neural Network (CNN) model was assessed with standard metrics Precision, Recall, F1-Score, and Accuracy, which collectively measure its ability to correctly identify and classify different plastic types under varying conditions accurately.

For the evaluation, the model produced the following confusion matrix results: True Positives (TP) = 389, False

Positives (FP) = 53, False Negatives (FN) = 97, and True Negatives (TN) = 461.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \text{ -----(1)}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \text{ -----(2)}$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \text{ -----(3)}$$

Table 1: Evaluation of CNN model

Evaluation Metric	Percentage
Precision	88%
Recall	80%
F1-Score	84%
Accuracy	85%

Overall, the evaluation validates that the hybrid approach effectively combines deep learning perception with symbolic reasoning, offering high accuracy, adaptability, and interpretability. Future work will focus on expanding the dataset, conducting cross-domain validation, and deploying in real-world settings to further enhance scalability and generalization.

5. CHALLENGES

Challenges encountered include designing a comprehensive ontology capable of representing diverse waste types and the difficulty in visually distinguishing plastic types solely by shape and cleanliness. These limitations suggest future work should incorporate multimodal data to improve classification accuracy and ontology comprehensiveness.

6. CONCLUSION

This study demonstrates the viability of integrating CNNs with an ontology-based knowledge base to improve plastic

waste management. Future work includes expanding the dataset, refining rule-based recommendations with local policies, integrating additional sensor modalities, and conducting field trials to evaluate behavioral impacts. Despite the high performance, certain limitations and avenues for future research exist;

Expansion of Material Types: The current model is focused primarily on common plastics. Future work could expand the classification to include paper, glass, metals, and electronic waste.

Fusion of Sensor Data: Integrating additional sensory data, such as near-infrared (NIR) spectroscopy or weight, could further resolve ambiguities between visually similar plastic types, potentially pushing accuracy to near 100%.

In: Balandin, S., Andreev, S., Koucheryavy, Y. (Eds.), *Internet of Things, Smart Spaces, and Next Generation Networking. ruSMART NEW2AN 2013 Lecture Notes in Computer Science: Vol 8121*. Springer. https://doi.org/10.1007/978-3-642-40316-3_7

REFERENCES

- Al-Mamun, A. M., Hossain, R., Sharmin, M., Kabir, E., & Iqbal, A. (2023). Garbage classification using convolutional neural networks (CNNs). *Material Science & Engineering International Journal*, 7(3), 140 - 144. <https://doi.org/10.15406/mseij.2023.07.00217>
- Dimitrijevic, K. (2023). Visualising Plastic Ocean Pollution: Designing Waste Ontologies. [Doctoral dissertation, University of London]. <https://research.gold.ac.uk/id/eprint/33131/>
- Glouche, Y., & Couderc, P. (2013). A Smart Waste Management with Self-Describing objects. *Proceedings of the Second International Conference on Smart Systems, Devices and Technologies (SMART'13), Italy*, 63 - 70.
- Kruthika, M., Rajadevi, R., Sathya, D., Varshini, S. S., Sowbharanika, J. J.S., & Suresh, B. K. (2024). Garbage Classification: A Deep Learning Perspective. *International Research Journal on Advanced Engineering Hub (IRJAEH)*, 2(12), 2774 - 2780. <https://doi.org/10.47392/irjaeh.2024.0384>
- Sinha, A., & Couderc, P. (2013). A Framework for Interacting Smart Objects.



Consumer Trends in Vegan and Plant-Based Products: A Reddit-Based Sentiment Analysis

Karunaratne KTT¹, Gunasekara RPTH²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

thamodhikarunaratna@gmail.com¹

ABSTRACT

Over the past decade, veganism and plant-based lifestyles have evolved from niche movements into mainstream consumer preferences, fueled by growing health awareness, environmental sustainability goals, and ethical considerations. Gaining a deep understanding of consumer attitudes has become essential for guiding product innovation, marketing strategies, and business development in this expanding market. This study explores consumer preferences and sentiments toward vegan and plant-based products using Reddit discussions, offering an alternative to traditional survey methods that often lack scale and timeliness. Two datasets were analyzed: a Kaggle dataset with 37,056 comments and a self-collected dataset of 7,489 comments from 2020 to 2024. The findings of this research, obtained from the available data, indicate that 65% of sentiments were positive, 22% were neutral, and 13% were negative. Meat alternatives emerged as the most popular and frequently discussed product category. Additionally, the study identified the best-reviewed, most-discussed, and most-criticized products across categories. These findings demonstrate Reddit's value as a dynamic source for understanding consumer behavior in the vegan and plant-based market and provide insights for product development, marketing, and consumer engagement strategies.

KEYWORDS: Consumer behavior, Plant-based products, Reddit data, Sentiment analysis, Veganism

1. INTRODUCTION

Over the past decade, veganism and plant-based lifestyles have evolved from niche trends to mainstream consumer choices, driven by concerns for health, the environment, and ethics. Understanding consumer attitudes is vital for product development, marketing, and innovation.

Traditional surveys and interviews are limited by small samples and cannot capture rapidly evolving consumer discussions. Social media platforms, especially Reddit, provide large-scale, dynamic data that reflects real consumer experiences. Subreddits dedicated to veganism offer valuable insights into the motivations, challenges, and perceptions of plant-based products.

This study utilizes Reddit data, combined with sentiment analysis, natural

language processing (NLP), and deep learning, to examine consumer trends and preferences. It examines sentiment, identifies popular product categories, and interprets perceptions, providing actionable insights to guide businesses and policymakers in promoting sustainable, consumer-aligned growth in the plant-based industry.

2. LITERATURE REVIEW

The growth of social media platforms, such as, Twitter and Reddit have enabled researchers to analyze user-generated content, offering valuable insights into public opinion, consumer preferences, and lifestyle trends. A key focus area has been veganism and plant-based diets, which have gained popularity for health, environmental, and ethical reasons. Sentiment analysis, combining machine

learning (ML), natural language processing (NLP), and deep learning, has become essential for interpreting public discourse and emotions at scale.

2.1 *Advances in Sentiment Analysis*

Research in sentiment analysis has evolved from classical ML to deep learning and transformer-based architecture. Papadimitriou et al. (2025) found that CNNs outperform Naïve Bayes, Decision Trees, and SVMs by leveraging contextual learning. Hanić et al. (2024) demonstrated that transformer-based models, such as BERT, achieve higher accuracy but require more computational resources. Similarly, Zhang et al. (2023) demonstrated that RoBERTa outperforms BERT and BiLSTM in Reddit sentiment analysis, though rare emotions remain difficult to classify. Shamoï et al. (2022) found that optimized ML pipelines with feature selection continue to deliver efficient, competitive results on large datasets. Collectively, these studies highlight progress toward more accurate and scalable sentiment models.

2.2 *Social media and Veganism*

Sentiment analysis has also been applied to study online trends related to veganism. Drescher et al. (2023) found that #Veganuary discussions emphasize health, novelty, and ethics. In contrast, Aleixo et al. (2020) showed that “vegan” is associated with activism and identity, while “plant-based” conveys a neutral, health-focused tone. These findings underscore the influence of linguistic framing on consumer perception.

2.3 *Research Gaps*

Despite methodological progress, most studies focus on basic sentiment

polarity rather than motivations such as health, ethics, or sustainability. Moreover, Twitter dominates as a data source, while Reddit, which offers richer and in-depth discussions, remains underexplored. Future research should adopt multi-platform, motivation-centered, and mixed-method approaches to capture the full complexity of consumer discourse on veganism.

3. METHODOLOGY

The methodology focuses on analyzing consumer sentiments and trends toward vegan and plant-based products using social media data. It includes data collection, preprocessing, sentiment labeling, and deep learning-based model training to ensure accurate and reliable results. Each stage of the process is described in the following subsections.

3.1 *Data Collection*

Kaggle Dataset: The primary dataset was sourced from Kaggle and contains labeled vegan-related social media comments. Initially, there were 37,056 comments; after cleaning duplicates and irrelevant entries, 37,029 comments remained for analysis and model training.

Reddit Product Comment Dataset: Using Python Reddit API Wrapper (PRAW), comments were scraped from five vegan-focused subreddits (r/vegan, r/plantbased, r/VeganFood, r/veganrecipes, and r/PlantBasedDiet) covering January 2020 to December 2024. The initial 7,488 comments were cleaned to 7,478. Applying a filter of 127 product-related keywords ensured the dataset focused on relevant vegan and plant-based product discussions.

3.2 Data Pre-Processing

The Kaggle dataset was pre-labeled into positive, neutral, and negative categories using the *cardiffnlp/twitter-roberta-base-sentiment* model, while Reddit comments were processed similarly for consistency. Sentiment labels were encoded numerically and split into training and validation sets (80:20) to ensure balanced learning. Text data was then tokenized with the *Keras Tokenizer*, converting comments into standardized 100-token sequences for model input.

3.3 Model Architecture

The sentiment classification model used in this study is a Bidirectional Long Short-Term Memory (BiLSTM) neural network designed to capture contextual meaning in text. The architecture consists of:

- **Embedding Layer:** Converts tokens into dense semantic vectors.
- **BiLSTM Layers:** Two stacked layers process input sequences in both directions
 - forward and backward directions to capture the full context.
- **Dropout Layers:** Applied for regularization to reduce overfitting.
- **Dense Output Layer:** Uses *softmax* activation function to classify sentiment as *positive*, *neutral*, or *negative*.

This configuration strikes a strong balance between model complexity and generalization, enabling it to detect sentiment nuances in Reddit discussions accurately.

3.4 Model Training

The BiLSTM model was trained under the following (Table 1) configuration:

Table 1: Model Training Components

Component	Description
Optimizer	Adam (lr = 0.001)
Loss	Categorical Cross-Entropy
Batch Size	128
Validation Split	20%
Environment	Python 3 + Keras (VS Code / Jupyter)

This configuration provided efficient learning, stable convergence, and reliable sentiment classification performance for vegan and plant-based product discussions.

4. RESULTS AND DISCUSSION

4.1 Model Performance Metrics

The effectiveness of the Bidirectional Long Short-Term Memory (BiLSTM) model developed for sentiment classification was evaluated using four key metrics. (Accuracy, Precision, Recall, and F1 score.) The results summarized in Table 2 indicate a high degree of reliability in capturing consumer sentiment expressed in social media comments.

Table 2: Model Performance Metrics

Metric	Value
Accuracy	83%
Precision	82%
Recall	83%
F1 Score	82%

Table 3: Product Category Rankings

Product category	Average Score	Mention Count	Normalization Score	Normalization Mentions	Final Score	Rank
Meat Alternatives	0.486905	4204	0.952237	1.000000	0.971342	1
Vegan & plant-based condiments	0.490978	1871	1.000000	0.438643	0.775457	2
Vegan & plant-based snacks	0.475392	1926	0.817198	0.451877	0.671070	3
Vegan supplements	0.458346	48	0.617279	0.000000	0.370367	4
Dairy alternatives	0.405716	1627	0.000000	0.379933	0.151973	5

4.2 Comparative Analysis

To evaluate the performance of different sentiment classification approaches, several machine learning and deep learning models were trained and tested on the collected dataset.

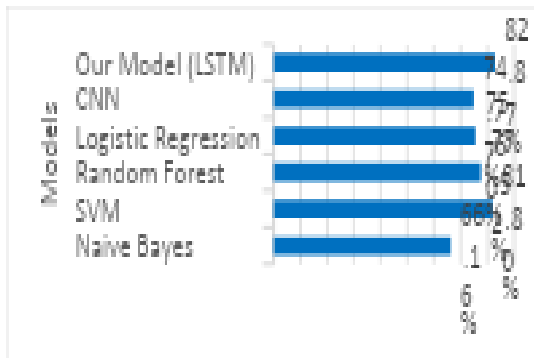


Figure 1: Comparative Analysis between other models

As shown in Figure 1, LSTM achieved the highest accuracy (82.87%), outperforming all other models, followed by SVM (81.80%). The strong performance of the BiLSTM model in this study is consistent with prior findings showing that LSTM-based approaches outperform traditional machine learning methods for social-media sentiment

analysis. Studies such as Zhang et al. (2024) and Papadimitriou et al. (2025) highlight this advantage, especially when working with informal, user-generated text. Since Reddit comments are longer and more context-rich than Twitter posts, BiLSTM models are better able to capture nuanced sentiment, which supports the effectiveness observed in this study.

4.3 Sentiment Distribution

The sentiment predictions across the dataset revealed that the majority of consumer comments were positive (65%), suggesting generally favorable attitudes towards vegan and plant-based products. (Figure 2)

CONSUMER TRENDS IN VEGAN AND PLANT-BASED PRODUCTS: A REDDIT-BASED SENTIMENT ANALYSIS

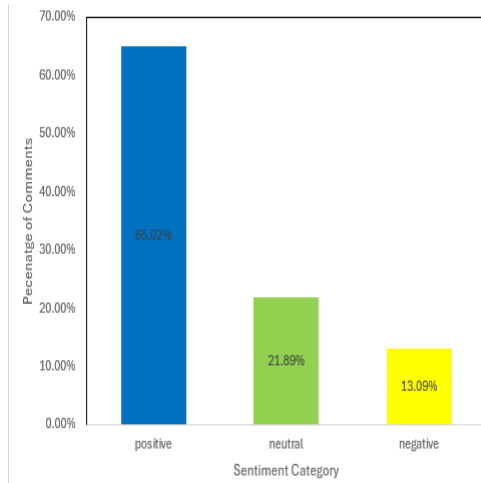


Figure 2: Sentiment Distribution

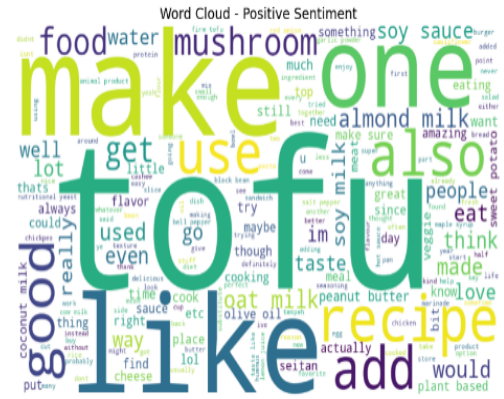


Figure 3: Word cloud for Positive sentiment category

Table 4: Product Category perspectives

Product category	Best Reviewed Product	Most Discussed product	Most Complained product	Most criticized product
Meat Alternatives	Almond cream	Oat milk	Oat milk	Rice cheese
Vegan & plant-based condiments	Lentil patties	Tofu	tofu	tofu
Vegan & plant-based snacks	Oatmeal cookies	Almond	almond	Vegan muffins
Vegan supplements	harissa	Peanut butter	Peanut butter	chimichurri
Dairy alternatives	Rice protein powder	Spirulina	spirulina	Vegan vitamin D

4.4 Key Insights

To evaluate which vegan and plant-based product categories are most prominent in Reddit discussions, both sentiment scores and mention frequencies have been analyzed. The final ranking was then derived by combining sentiment and frequency measures. The results are summarized in Table 3.

To capture detailed consumer perceptions, individual products within each category were analyzed based on four perspectives;

- **Best Reviewed Product:** Received the highest proportion of positive sentiment in consumer discussions.

- **Most Discussed Product:** Appeared most frequently in Reddit comments, regardless of sentiment.

- **Most Complained-Product:** Most associated with negative sentiment or user complaints.

- **Most Criticized Product:** Most frequently mentioned in strongly negative contexts

Table 4 summarizes these findings, revealing which products gain the most attention and how consumers distinguish between positive and negative experiences on Reddit.

5. CONCLUSION

This study provides valuable insights into consumer perceptions of vegan and plant-based products using social media sentiment analysis. The results revealed that 65% of comments were positive, reflecting growing acceptance of veganism driven by health, sustainability, and ethical awareness, while 22% were neutral and 13% negative, indicating some lingering skepticism. The BiLSTM model achieved the highest accuracy (82.87%), outperforming traditional models such as Logistic Regression and Random Forest, confirming the advantage of deep learning in capturing contextual nuances from social media text. Product category analysis showed that meat alternatives were the most positively discussed, whereas dairy alternatives attracted more criticism regarding taste and texture. Further, products such as oat milk and vegan muffins frequently appeared in negative discussions, highlighting areas for improvement. Word cloud analysis revealed that positive comments focused on health and enjoyment, while negative comments emphasized taste and price. Overall, the study demonstrates how social media analytics can uncover consumer attitudes and help businesses refine products and strategies to meet evolving expectations.

Although this study offers valuable insights into consumer attitudes toward vegan and plant-based products. It is limited by its reliance on Reddit data, which may not represent wider or non-English-speaking consumers, and by automated sentiment labels that may miss sarcasm or cultural nuances. Future work could include multiple social media platforms, multilingual datasets, and more

advanced transformer models. Exploring motivations such as health, ethics, and sustainability through topic modeling would further deepen the analysis.

REFERENCES

- Drescher, L. S., Grebitus, C., & Roosen, J. (2023). Exploring Food Consumption Trends on Twitter with Social Media Analytics: The Example of #Veganuary. *EuroChoices*, 22(2), 45 - 52. <https://doi.org/10.1111/1746-692X.12403>
- Hanić, S., Bagić Babac, M., Gledec, G., & Horvat, M. (2024). Comparing Machine Learning Models for Sentiment Analysis and Rating Prediction of Vegan and Vegetarian Restaurant Reviews. *Computers*, 13(10), 248. <https://doi.org/10.3390/computers13100248>
- Aleixo, M. G. B., Sass, C. A. B., Leal, R. M., Dantas, T. M., Pagani, M. M., Pimentel, T. C., Freitas, M. Q., Cruz, A. G., Azeredo, D. R. P., & Esmerino, E. A. (2020). Using Twitter® as source of information for dietary market research: a study on veganism and plant-based diets. *International Journal of Food Science and Technology*, 56(1), 61 – 68. <https://doi.org/10.1111/ijfs.14743>
- Papadimitriou, O., Kanavos, A., Maragoudakis, M., & Gerogiannis, V. C. (2025). A Comparative Study of Sentiment Analysis on Twitter and Reddit Using Deep Learning Techniques. *Proceedings of the 2024 8th International Conference on Algorithms, Computing and Systems, China*, 50 - 55. <https://doi.org/10.1145/3708597.3708605>
- Shamo, E., Turdybay, A., Shamo, P., Akhmetov, I., Jaxylykova, A., & Pak, A. (2022). Sentiment analysis of vegan related tweets using mutual information for feature selection. *PeerJ Computer Science*, 8, e1149. <https://doi.org/10.7717/peerj-cs1149>
- Zhang, X., Qi, X., & Teng, Z. (2024). Performance evaluation of Reddit Comments using Machine Learning and Natural Language Processing methods in Sentiment Analysis. ArXiv. <https://arxiv.org/abs/2405.16810>



Enhancing Small Business Management Through Customizable Mobile Applications in Sri Lanka

Vidana Pathirana YK¹, Munasinghe B²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

yashodakawindi@gmail.com¹, bhagya@wyb.ac.lk²

ABSTRACT

This study investigates how Sri Lankan small business owners (SBOs) use mobile applications and addresses the lack of suitable tools for their business needs. Using a two-phase approach (i.e., a pilot study followed by the design and development of a prototype application), the research followed a mixed-method approach, involving a survey with 100 participants in Phase 1 and a prototype evaluation with 10 beta testers in Phase 2. The research found a significant gap between existing mobile application features and SBOs' operational requirements. In response, a unified, customizable mobile application was developed and tested. Participants reported notable increases in confidence in mobile application usage in their businesses and in task efficiency after using the prototype, particularly in operational tasks. Feedback results from beta testers showed improvements in their initial views on efficiency, particularly in financial record-keeping, expense tracking, and communication, along with increased confidence in using digital tools. The findings suggest that customizable mobile solutions that cater to local small business needs can drive digital transformation in the small business sector. The study offers insights for mobile application developers and policymakers and recommends future research on expanding features, such as analytics and cloud integration for greater scalability in applications of this nature.

KEYWORDS: Mobile Applications, Small Business Management, Small Business Owners

1. INTRODUCTION

Mobile applications are essential for modern business efficiency, automation, and accessibility. In Sri Lanka, small business owners (SBOs) who are key contributors to employment and local economic growth, increasingly use mobile applications for tasks like sales tracking, inventory management, customer communication, and financial record-keeping due to growing smartphone and mobile internet penetration. Despite increasing mobile technology use, most SBOs rely on generic applications not tailored to their specific operational, financial, and strategic needs. Limited technical knowledge, lack of training, low awareness of suitable apps, and low digital literacy further impede effective

adoption. (Jayarathne et al., 2023). As a result, many SBOs either underutilize available mobile tools or depend on multiple applications, which can lead to inefficiency and data inconsistency. The findings of the pilot study identified a significant gap between user needs and the functionality of existing solutions, highlighting that SBOs required a unified, user-friendly, and affordable mobile solution matching their business context (Kawindi and Munasinghe, 2025).

Accordingly, this research proposes a unified, customizable mobile application and demonstrates how a user-centered mobile solution can bridge the gap between awareness and practical adoption of digital tools among small business owners. Furthermore, it provides empirical evidence of the benefits such

solutions offer in improving efficiency, decision-making, and overall digital engagement within Sri Lanka's small business sector. This study aims to identify the challenges faced by Sri Lankan small business owners in adopting mobile applications and to design and evaluate a unified, customizable mobile application tailored to their needs.

2. LITERATURE REVIEW

While mobile applications are increasingly important for small business management, their usability and adoption depend on users' literacy, digital skills, and perceived usefulness. Research on mobile technology adoption in Sri Lanka, specifically among SBOs, remains limited.

Previous research shows that Sri Lankan users, especially in rural and semi-urban areas, face usability and adoption challenges with mobile applications due to limited technical knowledge, inconsistent technology exposure, and language barriers (Raveendran & Rajeshwaran, 2016; Sahina & Vilani, 2021). Raveendran and Rajeshwaran (2016) have emphasized the necessity of implementing strategic initiatives that align with the rapid growth of mobile device usage and internet adoption in Sri Lanka. However, these studies primarily focused on marketing and consumer behavior rather than operational or managerial use within small enterprises.

Research on mobile applications in small business management highlights that user-centered design is crucial for adoption and satisfaction, as apps tailored to users' workflows, constraints, and preferences greatly enhance business efficiency and data-driven decision-

making (Venkatesh, 2003). Despite the benefits of user-centered mobile applications, Sri Lanka has limited availability of localized, purpose-built tools, forcing small businesses to rely on generic apps that do not meet their specific needs. Silva and Kariyapperuma (2021) have studied Sri Lankan managers' attitudes toward mobile devices but did not examine how business owners use mobile applications for operational tasks, highlighting the limited empirical evidence on SBOs' adoption of mobile apps. Although post-adoption studies show user satisfaction, ease of use, adaptability, and positive impacts on digital transformation, no such research exists for the Sri Lankan context.

3. METHODOLOGY

The study was conducted in two phases: Phase 1) a survey-based pilot study to understand the mobile application usage and awareness of SBOs, and Phase 2) development and evaluation of a unified, customizable mobile application prototype for small business management employing an incremental software development approach (Figure 1). The prototype development was followed by a feedback analysis that compared the findings of the pilot study with post-adoption feedback to assess the effectiveness, usability, and impact of the developed application.

The study targeted Sri Lankan small business owners (SBOs) managing small-scale, self-run businesses (e.g., grocery shops, beauty salons, medical practices, retail stores, and tutoring services) typically employing fewer than 10 people and relying on owner-led management.

3.1 Phase 1 – Pilot Study Details

The objectives of the pilot study were;

1. Exploring the nature of mobile applications usage of Sri Lankan SBOs
2. Identify limitations and challenges faced by SBOs when using existing mobile applications, with a focus on feature gaps, usability issues, and operational constraints.
3. Determining the demand and preferences for a mobile application tailored to small business needs.

Table 1: Categories of respondents

Categories of SBOs	No of Participants
Do not use any mobile application	37
Use multiple mobile applications for different tasks	58
Interested in adopting a unified mobile solution	5

A total of 100 respondents completed a structured questionnaire distributed via social media, WhatsApp business groups, and email. Participants were grouped by current mobile app usage (Table 1). The responses were analyzed quantitatively and qualitatively, with themes identified by coding descriptive answers alongside supporting quantitative data.

3.2 Phase 2 – System Development and User Feedback Survey

Based on Phase 1 findings, a custom mobile application prototype was developed with features tailored for small business management, following an incremental software development approach (i.e., adding features by feature into the prototype, assessing the performance with every increment). Five

core features, namely 1) finance management, 2) inventory management, 3) customer management, 4) appointments, and 5) staff scheduling. Each core feature consists of sub-features (i.e., functions) specific to their area. The users can choose one or more of the main features and customize their sub-features according to the needs of their business.

The main interfaces and the sub-feature interfaces were created to minimize the problems indicated by the SBOs in the pilot study, ensuring the user friendliness while offering improved performance presented in a simple set of interfaces.

Ten SBOs from the pilot survey participants were selected to beta-test the application. A structured questionnaire was used to gather feedback on usability, ease of navigation, satisfaction, perceived impact on business operations, and suggestions for improvement. The data were analyzed quantitatively and qualitatively to evaluate the application's effectiveness and identify areas of enhancement.

4. RESULTS AND DISCUSSION

4.1 Pilot Study

The findings of the pilot study have been published by the authors in Kawindi and Munasinghe (2025). The study found that only 63% of participants used mobile applications for business tasks such as finance management, customer handling, and scheduling, while 37% did not use any apps, revealing a significant gap in digital adoption that underscored the need for increased awareness and provisions to access affordable, user-friendly mobile solutions. Furthermore, 92.1% of the app users rely on using multiple applications for different operations, emphasizing the

absence of a suitable integrated, all-in-one mobile solution.

Small business owners found HRM and inventory management apps the least user-friendly, while finance and scheduling apps received mixed feedback, suggesting that many existing tools are too complex for users with limited technical skills. The most desired features in a tailored app include financial management, customer notifications, performance reporting, and appointment scheduling. Overall, 63% of users reported usability challenges, such as complex interfaces, slow response times, and data synchronization issues. Although many businesses mobile tools exist, usability remains a major issue for small business owners. About 63% reported difficulties such as complex interfaces, poor response times, data management problems, and syncing issues, with even basic features posing challenges for those with limited digital experience.

Non-users of business applications cited the lack of suitable apps, high costs, limited knowledge, and security concerns as key barriers. However, most showed willingness to adopt an affordable, unified mobile solution, revealing strong market demand for such business management tools.

4.2 Preliminary Usability Evaluation of the Developed Application

The preliminary usability evaluation in the form of a beta-test was conducted with 10 selected SBOs from the original survey group to assess the perceived suitability and effectiveness of the newly developed mobile solution.

Overall Suitability for Business Management

Participants rated the prototype's suitability for managing business activities on a least-to-most five-point scale, with most rating it as "Suitable (4)" or "Very Suitable (5)," indicating that the proposed solution effectively addresses previously identified relevance and usability issues.

Evaluation of Core Features

Participants evaluated the prototype's five core features (i.e., finance management, inventory management, customer management, appointments, and staff scheduling) on usefulness and effectiveness. Finance and inventory management scored highest, showing that the user-centered design reduced complexity. Customer management and appointments received positive feedback, while staff scheduling, though rated slightly lower, remained valuable for certain business types like salons and clinics.

4.5.1 Summary of Qualitative Feedback

Qualitative feedback reinforced the quantitative findings, with participants describing the prototype as "intuitive," "easy to use," and comprehensive. Users praised features, such as inventory management and invoice generation, while suggesting enhancements such as advanced reporting and customizable notifications. Overall, the evaluation confirmed that the user-centered design effectively addressed key usability gaps for Sri Lankan small business owners.

4.3 Discussion

Overall, the findings indicate that a unified, customizable mobile application can greatly improve operational efficiency and user satisfaction among Sri Lankan small business owners, effectively bridging the gap between available technology and real-world usability.

5. CONCLUSION

The study examined mobile app usage among Sri Lankan SBOs and tested a newly developed unified, customizable business management application. The prototype addressed issues identified in a pilot study, with evaluation showing high user satisfaction, particularly in finance and inventory management. Participants valued the app's simplicity, usability, and integrated processes. The findings suggest that user-centered mobile applications can significantly enhance operational efficiency and satisfaction, with future work recommended to improve the features further, to add real-time analytics, cloud integration, and based on large-scale testing.

REFERENCES

- Jayarathne, P. G. S. A., Chathuranga, B. T. K., Dewasiri, N. J., & Rana, S. (2023). Motives of mobile payment adoption during COVID-19 pandemic in Sri Lanka: A holistic approach of both customers' and retailers' perspectives. *South Asian Journal of Marketing*, 4(1), 51 - 73.
- Kawindi, Y., & Munasinghe, B. (2025, September). Mobile app usage and awareness among Sri Lankan small business owners. *Proceedings of the 4th SLIIT International Conference on Engineering and Technology (SICET 2025)*, 112 - 118.
- Raveendran, N., & Rajesvaran, A. (2016). Are businesses succeeding in Sri Lanka through online marketing? *Proceedings of the International Conference on Contemporary Management (ICCM 2015)*, University of Jaffna.
- Shahina, F., & Vilani, S. (2021). Mobile marketing initiatives on consumer attitudes and buying behaviour of young consumers in Sri Lanka. *Asian Journal of Advanced Research and Reports*, 15(1), 54 - 67.
- Silva, A. A. S. D., & Kariyapperuma, S. (2021). Conspicuous consumption of mobile phones among Sri Lankan managers. *Proceedings of the International Conference on Business Management*, 18, Faculty of Management Sciences, University of Sri Jayewardenepura.
- Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425 - 478.



Evaluating the Effectiveness of Standard Operating Procedures in Website Development Projects: A Mixed-Methods Analysis

Amarasekara PEM¹, Arudchelvam T²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

poojaerandathi1999@gmail.com¹, arul@wyb.ac.lk²

ABSTRACT

Standard Operating Procedures (SOPs) are widely recognized as essential tools for ensuring consistency and efficiency in structured industries such as healthcare and manufacturing. However, their application in creative and dynamic fields, such as website development, remains underexplored. This study evaluates the effectiveness of SOPs in website development projects through a mixed-methods approach, examining both quantitative performance metrics and qualitative stakeholder experiences. The research analyzes multiple projects within an organization that employs a comprehensive SOP framework covering all phases from requirement gathering to post-launch support. Key Performance Indicators, including project completion time, error rates, rework frequency, and client satisfaction, were measured alongside interviews with developers, designers, and project managers. Results indicate that SOPs significantly improve quality assurance, accountability, and client satisfaction, reducing error rates and enhancing workflow consistency. However, challenges including rigidity, increased documentation workload, and constraints on creative flexibility were identified, particularly in smaller projects. The study provides practical recommendations for developing adaptive SOP frameworks that balance standardization with the flexibility required in digital project environments.

KEYWORDS: Project Management, Quality Assurance, Software Development, Standard Operating Procedures, Website Development

1. INTRODUCTION

Standard Operating Procedures (SOPs) have long been established as fundamental instruments for maintaining consistency, efficiency, and accountability in structured industries. In healthcare, aviation, and manufacturing sectors, SOPs provide step-by-step guidelines that minimize errors, ensure compliance, and improve overall quality (International Organization for Standardization, 2015). However, the role of SOPs in creative and dynamic fields such as website development present a unique challenge, as these industries demand both systematic organization and innovative thinking.

Website development projects typically require a delicate balance

between technical accuracy and creative design, where teams must respond rapidly to evolving client requirements, technological advancements, and market demands. Unlike traditional industries where processes are highly standardized, digital projects often prioritize flexibility and adaptability. This creates tension between the structured nature of SOPs and the dynamic requirements of creative work (Sommerville, 2015).

While many organizations have begun adopting SOPs in their digital projects to achieve greater efficiency and client satisfaction, there remains limited empirical evidence evaluating their actual impact. Most existing research focuses on traditional industries, leaving a significant knowledge gap regarding SOP

effectiveness in website development contexts. This study addresses this gap by critically evaluating both the benefits and limitations of SOP implementation in digital project environments.

The research adopts a mixed-methods approach, combining quantitative analysis of project performance metrics with qualitative insights from team members and clients. By examining projects that follow structured SOP frameworks alongside those employing flexible approaches, this study provides balanced insights into how SOPs influence efficiency, quality, collaboration, and innovation in website development.

2. LITERATURE REVIEW

The application of SOPs has been extensively studied in structured industries where standardization is critical. In healthcare, SOPs ensure patient safety through consistent medical practices and infection control protocols (International Organization for Standardization, 2015). Aviation relies heavily on SOP-driven checklists for flight operations and maintenance, minimizing human error in high-stakes environments. Manufacturing sectors use SOPs to maintain quality standards, streamline production processes, and ensure compliance with international certifications.

Project management frameworks such as PRINCE2 and PMBOK incorporate SOP-like elements through documented procedures for planning, execution, and monitoring (Project Management Institute, 2021). These frameworks demonstrate that procedural approaches can enhance predictability and accountability. However, Agile methodologies like Scrum, emphasize

lightweight procedures that support adaptability and continuous improvement, suggesting that structure and flexibility can coexist (Schwaber & Sutherland, 2020).

In creative industries, research on SOPs remains limited. Studies suggest that while standardization can improve workflow consistency, overly rigid procedures may constrain creativity and limit problem-solving opportunities (Nonaka & Takeuchi, 1995). The software engineering literature highlights the importance of balancing process discipline with innovation, particularly in environments where client requirements frequently change (Sommerville, 2015).

Knowledge management research indicates that SOPs serve as repositories of organizational learning, facilitating knowledge transfer and reducing dependency on individual expertise (Davenport & Prusak, 1998). However, excessive reliance on documented procedures may discourage critical thinking and adaptability, particularly in rapidly evolving digital environments. This dual nature of SOPs—supporting both standardization and potential rigidity—forms the foundation for investigating their effectiveness in website development projects.

Recent studies have explored SOP adoption within modern software and digital service environments. Zhang & Li (2021) highlight how Agile teams use lightweight SOPs to maintain consistency while preserving adaptability. Olanrewaju et al. (2022) report that SOPs in creative digital workflows improve quality and handover efficiency but must remain flexible to avoid constraining innovation. Similarly, Rahman (2019) and Patel & Desai (2023) emphasize hybrid SOP

models in web development and UI/UX teams, where standardized documentation supports knowledge transfer without limiting creativity. These recent findings align with the current study and highlight the growing relevance of SOPs in digital and fast-paced project environments.

3. METHODOLOGY

This study employed a mixed-methods research design to evaluate SOP effectiveness in website development projects. The research was conducted within an organization that maintains a comprehensive SOP framework covering all phases of the project lifecycle, from requirement gathering through post-launch maintenance.

3.1 Research Design

The study adopted a case-based comparative strategy, analyzing multiple website development projects with varying degrees of SOP adherence. Projects were selected based on scale, complexity, and workflow structure, enabling meaningful comparison between SOP-driven and flexible approaches.

3.2 Data Collection

Quantitative data was extracted from project documentation, including completion times, error logs, rework records, milestone tracking sheets, and client satisfaction surveys. Key Performance Indicators (KPIs) measured included project delivery timelines, defect rates, adherence to milestones, and resource utilization efficiency. Qualitative data were gathered through semi-structured interviews with developers, designers, project managers, and clients. Interview questions explored perceptions of SOP usefulness, impacts on creativity

and collaboration, and experiences with workflow flexibility. Direct observations during design reviews, quality control sessions, and client interactions provided additional contextual insights. Figure 1 illustrates the structured workflow implemented in SOP-driven projects, demonstrating the sequential phases and quality checkpoints embedded throughout the development process.

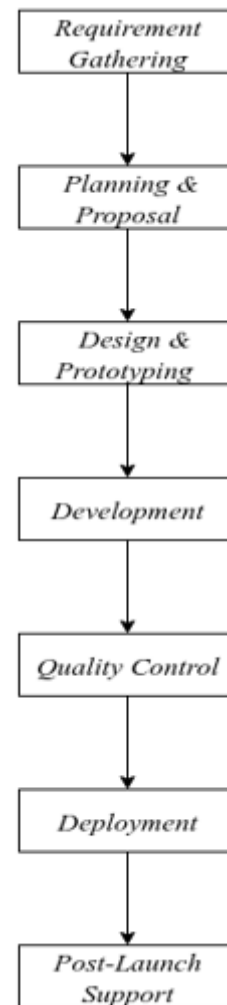


Figure 1: Complete project lifecycle workflow

3.3 Analysis Methods

Quantitative data were analyzed using descriptive statistics and comparative analysis to identify performance patterns. Qualitative data underwent thematic

coding to identify recurring themes related to efficiency, creativity, collaboration, and satisfaction. Data triangulation was employed to cross-validate findings from multiple sources, enhancing reliability and reducing bias.

4. RESULTS AND DISCUSSION

4.1 Quantitative Findings

The respondents represented a balanced distribution of organizational roles, including the CEO, Tech Lead, Senior Web Developer, Web Developer, Associate Web Developer, and Web Developer Intern, with each role contributing 16.7% of the total responses. This equal participation ensured that feedback reflected perspectives from both leadership and operational levels.

Table 1: Summary of Quantitative Evaluation of SOP Implementation Across Key Performance Areas

Evaluation Criterion	Evaluation Criterion	Dominant Response (% Rated)
Work process clarity and organization	4.83	83.3%
Team efficiency and time savings	5.00	100%
Creativity in work process	4.50	50%
Achievement of targets	4.83	83.3%
Overall Average	4.79	79.2%

Table 1 presents the quantitative summary of the SOP evaluation, capturing how the procedure influenced clarity, efficiency, creativity, and target achievement across the team.

4.2 Qualitative Insights

Open-ended feedback supported the quantitative findings. Participants

described the SOP as a “clear guide that keeps everyone aligned” and a “time-saving framework that minimizes repeated mistakes”.

Team members highlighted that the standardized steps enhanced communication between departments and reduced task duplication.

However, a few respondents suggested maintaining flexibility for creative work. One participant remarked that while the SOP streamlined workflow, it sometimes “felt restrictive for design-related decisions,” implying that creativity can benefit from controlled adaptability within standard guidelines.

Overall, both quantitative ratings and qualitative insights confirm that SOP implementation significantly improved workflow clarity, efficiency, and goal alignment, while promoting a balance between structure and creativity.

4.3 Challenges Identified

Three primary challenges emerged from the analysis.

First, the SOP framework introduced rigidity that conflicted with the need for rapid adjustments in creative projects. Minor client-requested changes required formal approval cycles, creating delays that frustrated both teams and clients.

Second, smaller projects experienced disproportionate overhead from documentation and repeated quality checks. While these processes added value in complex projects, they sometimes reduced efficiency in simpler cases, suggesting the need for scalable SOP models.

Third, scope creep remained problematic despite structured requirement gathering. Clients frequently introduced new features mid-project, and

the SOP's sequential structure struggled to accommodate unexpected changes without disrupting timelines.

5. CONCLUSION

This study demonstrates that SOPs significantly improve project performance in website development through enhanced quality assurance, greater accountability, and improved client satisfaction. The structured workflow reduces errors, streamlines coordination, and provides transparency that builds client trust. However, the research also reveals important limitations, particularly regarding rigidity, creative constraints, and scalability to smaller projects.

The findings suggest that effective SOP frameworks for digital projects must be adaptive rather than prescriptive. Organizations should develop tiered or modular SOP models that can be scaled according to project complexity, with lightweight versions for small projects and comprehensive protocols for larger initiatives. Incorporating flexible revision cycles and automation tools can help maintain structure while reducing administrative overhead.

This study contributes to both academic knowledge and industry practice by providing evidence-based insights into SOP effectiveness in website development. The recommendations for adaptive frameworks offer practical guidance for organizations seeking to optimize their project management practices while maintaining the creativity

and responsiveness essential for success in dynamic digital markets.

REFERENCES

- Davenport, T. H., & Prusak, L. (1998). *Working Knowledge: How Organizations Manage What They Know*. Harvard Business School Press.
- International Organization for Standardization. (2015). *Quality Management Systems – Requirements* (ISO Standard No. 9001:2015). <https://www.iso.org/standard/62085.html>
- Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press.
- Olanrewaju, A. et al. (2022). Evaluating SOP Adoption in Creative IT Workflows. *Information & Management*.
- Patel, R., & Desai, P. (2023). Hybrid SOP Models in UI/UX and Web Development Agencies. *ACM Digital Library*.
- Project Management Institute. (2021). *PMBOK Guide: A Guide to the Project Management Body of Knowledge*. <https://www.pmi.org/standards/pmbok>
- Rahman, K. (2019). SOP Optimization for Small & Medium IT Companies. *IEEE Access*.
- Schwaber, K., & Sutherland, J. (2020). *The Scrum Guide*. <https://scrumguides.org>
- Sommerville, I. (2015). *Software Engineering* (10th Ed.). Pearson Education Limited.
- Zhang, S., & Li, H. (2021). Standard Operating Procedures in Agile Software Teams: Balancing Structure and Flexibility. *Journal of Systems & Software*.



Factors Influencing Cloud ERP Adoption among SMEs in Developing Countries: A Systematic Review

Mariyam F¹, Wickramarachchi R²

^{1,2}*Department of Industrial Management, Faculty of Science, University of Kelaniya*
fathima-im20026@stu.kln.ac.lk¹, Ruwan@kln.ac.lk²

ABSTRACT

This systematic literature review examines key success factors influencing the adoption of Cloud Enterprise Resource Planning (ERP) among small and medium-sized enterprises (SMEs) in developing countries. Using the PRISMA framework, 24 peer-reviewed studies from 2020 to 2025 were thoroughly analyzed within the Technology–Organization–Environment (TOE) framework. Fifteen critical factors across technological, organizational, and environmental domains were identified, including technological compatibility, data security, top management support, organizational readiness, and regulatory environment. Beyond listing, this review critically analyzes regional variations and contradictions; for example, top management support is more emphasized in Southeast Asia compared to Africa, where infrastructure constraints often outweigh security concerns. Cost acts paradoxically as both a facilitator and a barrier, depending on the economic context. Interdependencies among factors reveal a nuanced adoption dynamic. Significant research gaps include scarce longitudinal studies, underexplored roles of emerging technologies like AI and big data, and limited cross-sectoral comparisons. These insights provide actionable guidance for practitioners, policymakers, and researchers aiming to foster successful digital transformation in resource-constrained SME environments.

KEYWORDS: Adoption, Cloud ERP, Critical Success Factors, Developing Countries, Technology-Organization-Environment (TOE) Framework

1. INTRODUCTION

Enterprise Resource Planning (ERP) systems have evolved significantly with cloud computing, shifting focus from costly on-premise infrastructure to scalable, flexible cloud solutions. For Small and Medium-sized Enterprises (SMEs), Cloud ERP offers advantages such as reduced upfront costs, simplified maintenance, and real-time accessibility, making it vital for competitiveness in digitizing markets (Thompson, 2025).

However, Cloud ERP adoption among SMEs in developing countries remains limited due to challenges including inadequate digital infrastructure, data security concerns, insufficient top management support, and low

organizational readiness (Mohammed et al., 2023). These barriers indicate that successful adoption requires synchronized organizational capabilities and supportive environmental factors.

The Technology–Organization–Environment (TOE) framework provides an analytical lens considering technology features, organizational resources, and external pressures affecting adoption (Alsharari, 2022). This systematic literature review synthesizes 24 peer-reviewed studies to identify, categorize, and critically analyze factors influencing Cloud ERP adoption among SMEs in developing regions, addressing gaps in empirical synthesis and highlighting key geographic and sectoral contradictions.

2. METHODOLOGY

2.1 Search Strategy and Database Selection

A systematic search was conducted across IEEE Xplore, ACM Digital Library, and Google Scholar to gather diverse perspectives on Cloud ERP adoption. Three Boolean search strings targeted technological, organizational, and environmental dimensions:

1. ("Cloud ERP" OR "On-Premise ERP") AND ("adoption" OR "deployment" OR "decision-making") AND ("framework" OR "model" OR "factors")
2. ("Cloud ERP" OR "On-Premise ERP") AND ("Sri Lanka" OR "developing countries") AND ("adoption" OR "implementation") AND ("success factors" OR "barriers")
3. ("ERP systems") AND ("decision framework" OR "selection model") AND ("Cloud" OR "On-Premise")

These combinations facilitated the identification of studies addressing adoption factors across various contexts.

2.2 Inclusion /Exclusion Criteria and Study Selection

Included were peer-reviewed studies published between 2020 and 2025, written in English, focusing on ERP adoption with accessible full texts. Excluded were grey literature, preprints, studies unrelated to ERP, or lacking methodological transparency. From 72 initial records, 24 relevant studies were selected after screening titles, abstracts, and full texts.

2.3 Data Extraction and Quality Assessment

Data extraction used a standardized template capturing author/year, region, study design, sample, critical factors, methodology, and findings. Quality assessment considered research objectives, methodology, sampling, and peer review status. Although studies were generally moderate to high quality, some lacked detailed reporting on effect sizes and longitudinal outcomes.

2.4 Data Synthesis and Analytical Framework

Synthesized data were categorized using the Technology–Organization–Environment (TOE) framework. Qualitative content analysis identified common patterns, contradictions, regional variations, and emerging themes, with special focus on geographic and sectoral differences influencing Cloud ERP adoption factors.

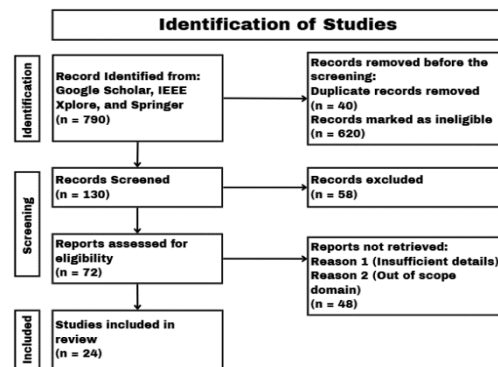


Figure 1: PRISMA workflow

3. RESULTS AND CRITICAL ANALYSIS

3.1 Critical Success Factors Identified

The systematic analysis identified 15 key CSFs categorized within the TOE framework;

Technological: Compatibility, Data Security & Privacy, Relative Advantage, Complexity, Scalability, Customization Limitations

Organizational: Top Management Support, Technical Knowledge & Skills, Firm Readiness, Change Management, and Organizational Culture

Environmental: External Pressure, External Vendor Support, Government Regulation Support, and Infrastructure Availability.

3.2 Comparative Regional Analysis and Critical Contradictions

Finding 1: Differential Importance of Top Management Support

Top management support is consistently identified as crucial in Southeast Asian contexts (for example, Malaysia: Razzaq et al., 2021). Conversely, African studies highlight the influence of organizational culture and peer-learning networks which are reflecting differences in hierarchical versus distributed decision-making structures across regions.

Finding 2: Security Concerns Versus Infrastructure Constraints

Data security is the main barrier in middle-income countries (Jordan: Zamzeer et al., 2020; UAE: Alsharari et al., 2020), whereas infrastructure availability is prioritized in low-income regions (Ethiopia, Sub-Saharan Africa). This suggests CSF hierarchies depend on context; security concerns presuppose adequate infrastructure, while infrastructure scarcity poses a primary barrier.

Finding 3: Cost Sensitivity and Adoption Decision-Making

However, contradictory evidence exists regarding cost: for instance, middle-income SMEs perceive cost as a relative advantage driving adoption (Thompson, 2025), while those in the lowest-income countries see cost as a barrier despite potential savings (Mohammed et al., 2023). This reflects varying financial capacities and risk perceptions.

3.3 Contextual Interdependencies and Systematic Relationships

The analysis reveals interdependencies among factors. Strong top management support enhances organizational readiness and mitigates perceived technological complexity. Regulatory support strengthens compliance capacity, influencing adoption decisions. External vendor support is crucial for low-capacity organizations, offsetting internal capability gaps.

4. DISCUSSION

4.1 Synthesis and Knowledge Contribution

This review moves beyond merely listing critical success factors (CSFs) by critically synthesizing evidence and uncovering systematic contradictions and regional variations that have not been previously integrated. The varying importance of CSFs across developing regions challenges universal adoption models and highlights the need for context-specific approaches. Interdependency analysis shows that isolated CSF interventions are inadequate; successful Cloud ERP implementation requires ecosystem-level alignment spanning technological fit, organizational capability, and environmental support. In

developing countries, environmental constraints such as limited infrastructure, regulatory uncertainty, and sparse vendor support compound the challenges beyond those in developed markets.

4.2 Implications and Research Gaps

For practitioners: Adoption strategies must prioritize CSFs according to context. Low-income regions should focus on foundational infrastructure before addressing security; middle-income countries need strong leadership and regulatory clarity; higher-income developing nations benefit from enhancing organizational capabilities and change management.

For researchers: Key gaps include a lack of longitudinal studies tracking adoption sustainability, insufficient research on emerging technologies like AI and big data in Cloud ERP, limited cross-sectoral comparative analyses, and underexplored reasons for geographic variations in CSF importance.

5. CONCLUSION

This systematic review synthesized evidence from 24 studies to identify 15 critical success factors influencing Cloud ERP adoption among SMEs in developing countries. Its novel contribution lies in critically analyzing regional hierarchies of these factors, contextual contradictions, and their interdependencies. While technological compatibility and security remain universally important, emphasis on organizational readiness and environmental constraints varies significantly by regional development level. Successful Cloud ERP adoption requires contextualized implementation strategies that recognize these regional differences. Policymakers should focus

on strengthening ICT infrastructure and regulatory frameworks; practitioners need phased adoption plans that account for local constraints; and researchers must pursue longitudinal studies and explore integration of emerging technologies. This evidence-based framework supports better navigation of the complex and context-dependent Cloud ERP adoption landscape in developing economies.

REFERENCES

- Alsharari, N. M. (2022). Cloud computing and ERP assimilation in the public sector: Institutional perspectives. *Transforming Government: People, Process and Policy*, 16(1), 97–110.
- Alsharari, N. M., Al-Shboul, M., & Alteneiji, S. (2020). Implementation of cloud ERP in the SME: Evidence from the UAE. *Journal of Small Business and Enterprise Development*, 27(2), 299–327.
- Mohammed, G. J., Burhanuddin, M. A., Dawood, F. A. A., Alyousif, S., Alkhayyat, A., Ali, M. H., Malik, R. Q., & Jaber, M. M. (2023). An empirical study on the affecting factors of cloud-based ERP system adoption in Iraqi SMEs. *International Journal of Advanced Computer Science and Applications*, 14(1), 430–441.
- Razzaq, A., Asmai, S. A., Zainal Abidin, Z., Talib, M. S., Ali, M. F., & Mohammed, A. A. (2021). Propose a conceptual framework for the cloud ERP adoption among Malaysian SMEs. *Journal of Engineering Science and Technology*, 16(4), 3387–3406.
- Thompson, L. (2025). Cloud-based management information systems: A paradigm shift in enterprise resource planning. *OTS Canadian Journal*, 4(6), 62–73.
- Zamzeer, M., Alshamaile, Y., Alsawalqah, H. I., Al-Hassan, M., Jalal, E., & Almubideen, S. (2020). Determinants of cloud ERP adoption in Jordan: An exploratory study. *International Journal of Business Information Systems*, 34(2), 204–228.



Factors Influencing the Transition of Non-IT Graduates into the IT Industry: A Systematic Review

Thavarasa Y¹, Wickramarachchi R²

^{1,2}*Department of Industrial Management, Faculty of Science, University of Kelaniya*
thavarasayuvitha@gmail.com¹, ruwan@kln.ac.lk²

ABSTRACT

The increasing demand for technological expertise across industries has led to a significant rise in individuals from non-IT backgrounds transitioning into the IT sector. This phenomenon reflects broader transformations in global career pathways, driven by rapidly evolving labour markets, the availability of accessible online learning platforms, and the growing need for digital competencies in diverse professional domains. Using the PRISMA framework to ensure methodological rigour, this study conducts a systematic literature review to explore the factors influencing these career transitions, the enabling pathways and resources followed by non-IT graduates during this process, and the impacts of this trend on the IT industry. The review synthesizes findings from recent scholarly work to identify key motivations such as higher salary prospects, job stability, personal interest in technology, and opportunities for career growth. Furthermore, it examines the implications of this trend for employers, educational institutions, and policymakers, emphasizing the importance of inclusive and adaptable talent development ecosystems. By combining insights from multiple contexts, this study contributes to the ongoing discourse on cross-disciplinary career mobility, the shaping of hybrid skill sets, and the redefinition of professional identity in the digital era. Ultimately, it provides a foundation for future research and practical frameworks supporting smoother transitions into IT careers.

KEYWORDS: Career transition, IT industry, Non-IT graduates, Tech careers, Workforce entry

1. INTRODUCTION

The rapid development of the digital economy and the widespread use of technology across all fields have significantly increased the demand for information technology professionals. The IT industry is now open to a broader talent pool, particularly as employers try and reduce skill shortages (Prommegger et al., 2020). These jobs often offer competitive wages, job security, and various career opportunities, rendering the IT sector more attractive to non-technical individuals (Assaf et al., 2025). This role, which tends to be dynamic and creative, combined with the existence of leading areas such as data science, cloud computing, cybersecurity, and artificial intelligence in IT, is also reinforcing this trend (Assaf et al., 2025). Non-IT

graduates are moving to more data-focused fields, such as those of data science or machine learning, even if their educational background is in business (Zhu, 2023).

2. METHODOLOGY

This systematic literature review followed the (PRISMA 2020) guidelines.

2.1 Eligibility Criteria

Studies were included if they focused on individuals transitioning into the IT field from non-technical or unrelated educational and professional backgrounds, including nontraditional entry routes such as coding bootcamps, self-learning, conversion programs, or requalification pathways. The search, conducted between June and July 2025,

used Google Scholar and ACM Digital Library, selected for their broad coverage of interdisciplinary and computing-focused research. Other databases (e.g., IEEE Xplore, Scopus) were considered, but most relevant studies were already indexed in the chosen databases, making additional sources redundant.

Combinations of following Boolean search query was applied: ("career shift" OR "career transition" OR "career change") AND ("non-technical graduates" OR "non-IT" OR "arts" OR "social science" OR "humanities") AND ("tech careers" OR "IT industry" OR "ICT jobs") AND ("implication" OR "impact" OR "outcome") Filters included English language and publication years 2019 - 2025. The initial search yielded 291 results, with duplicates removed using manual Excel filtering, resulting in 288 unique studies. Studies were screened in two stages: first, titles and abstracts were reviewed for relevance, and second, full-text articles were assessed against eligibility criteria. 23 studies meeting all inclusion criteria were included in the final synthesis. The PRISMA flow diagram (Figure 1) illustrates the screening and selection process.

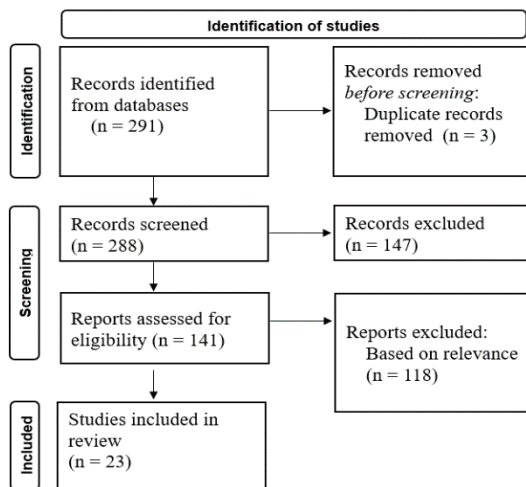


Figure 1: PRISMA flow diagram

3. RESULTS AND DISCUSSION

This section summarizes the key inferences based on the systematic literature review and interprets their implications.

3.1 Motivations for Transition

Career changes are often driven by dissatisfaction with previous jobs or fields, including low job satisfaction, burnout, employment insecurity, or limited advancement opportunities (Assaf et al., 2025; Barnes et al., 2024). Many individuals are also motivated by an intrinsic passion for IT, programming, or technology, described as a personal “calling” or the discovery of hidden aptitude (Gomes et al., 2025; Preoteasa, 2021). Personal growth, empowerment, and coping with structural uncertainties further shape these transitions. Building self-efficacy through workshops or hands-on experience can also inspire career shifts to IT (Lyon & Green, 2021), as can the desire to apply existing skills creatively (Gomes et al., 2025).

The IT field’s higher salaries and economic prospects remain powerful motivators (Assaf et al., 2025). IT careers also offer stability compared to limited opportunities in non-STEM fields (Cunha et al., 2021; Ong et al., 2022), with AI skills adding wage premiums (Ehlinger & Stephany, 2023). Social networks, friends, relatives, and mentors in IT strongly influence such career decisions (Preoteasa, 2021), whereas the absence of early career counseling may lead to later regret (Maharjan, 2025; Preoteasa, 2021).

Flexibility, remote work options, and improved work-life balance attract many to IT (Hyrnsalmi, 2024). Broader developments such as digitization, the COVID-19 pandemic, and changing labor markets also encourage these shifts (Preoteasa, 2021). The prestige of IT

careers and gender inclusion initiatives such as women in Tech further motivate, especially female professionals (Gribanova, 2021; Hyrynsalmi, 2024; Hyrynsalmi & Sutinen, 2019).

3.2 Enabling Pathways and Resources

Various routes facilitate these transitions. Self-study via online platforms, YouTube, and books is common (Preoteasa, 2021; Zhu et al., 2022). Bootcamps, training programs, and professional certifications provide industry-specific skills (Lyon & Green, 2020; Zhu et al., 2022). Formal education options include open university courses, postgraduate conversion programs, and second degrees for non-computer science graduates (Krenz et al., 2021; Molla et al., 2025; Zhu, 2023). Early exposure to IT work in non-IT roles or research assistantships also fosters interest (Gomes et al., 2025; Preoteasa, 2021). Supportive communities and networks, particularly for women, offer guidance and networking opportunities (Hyrynsalmi, 2024; Hyrynsalmi & Sutinen, 2019; Preoteasa, 2021).

Organizational support through internal training or employer partnerships further eases transitions (Prommegger et al., 2020), while practical factors like time and financial resources influence career decisions (Hamilton, 2022).

3.3 Impact on the IT industry

The IT sector is experiencing a more diverse workforce, valuing non-technical skills such as business knowledge, project management, and social skills alongside technical expertise (Hyrynsalmi & Sutinen, 2019; Prommegger et al., 2020). This trend enhances interdisciplinarity and the boundless nature of the profession (Gomes et al., 2025; Prommegger et al., 2020). Recruiting non-IT women supports

societal transformation and expands female participation in IT (Heinle et al., 2023; Hyrynsalmi & Sutinen, 2019), while diversifying the sector's skills pool (Cunha et al., 2021; S. Hyrynsalmi & Hyrynsalmi, 2019b). However, the influx of requalified professionals may increase competition among traditional IT employees (Cunha et al., 2021). AI introduces new opportunities and risks, requiring ethical attention to prevent bias and protect diversity (Hyrynsalmi et al., 2025). To address evolving demands, skill-focused educational models are necessary.

3.4 Discussion

The findings reveal that the transition of non-IT graduates into the IT industry is shaped by a blend of personal motivations, accessible learning opportunities, and ongoing industry transformations. To contextualize these patterns, the search strategy itself helped clarify how widely this topic is addressed across domains. Searches combining career-shift terms with general non-IT descriptors yielded the highest records, while adding IT-specific terms progressively narrowed the results. The full Boolean string produced only a small set of highly focused studies, and after screening and removing duplicates, just a limited number met the criteria for synthesis, reflecting how concentrated yet diverse the literature on this transition remains.

Across these studies, extrinsic factors such as salary, stability, and remote work typically trigger the initial shift, whereas intrinsic motivations such as problem-solving interest, creativity, and personal growth are more influential for long-term commitment. Learning pathways operate as a connected system in which self-study, bootcamps, certificates, and formal

conversion programs collectively support diverse entrants.

Industry-level findings show that non-IT professionals bring communication, business awareness, and interdisciplinary strengths, reflecting IT's move toward a more hybrid profession. However, these benefits coexist with tensions, including increased competition with IT graduates and questions about technical depth, challenges intensified by AI-driven skill demands.

Overall, successful transitions occur when personal motivation aligns with accessible learning routes and inclusive organizational practices, illustrating how digitalization and labour-market shifts shape entry into IT and long-term progression.

4. CONCLUSION

This review explored the growing movement of non-IT graduates entering the IT industry. Motivated by higher income, job security, flexibility, and career independence, these individuals pursue self-directed learning through coding bootcamps, certifications, and online platforms to bridge skill gaps. The trend highlights how evolving labour markets and accessible learning pathways reshape workforce composition. It also emphasizes the need for inclusive, adaptable training models and supportive organizational practices to sustain cross-disciplinary mobility. Future studies should examine contextual differences and underrepresented groups to strengthen policies promoting equitable participation in the digital economy.

REFERENCES

- Assaf, A., Abdellatif, H., & Shahrouy, F. (2025). Exploring career shifts to IT: Motivations and mechanisms. *World Transactions on Engineering and Technology Education*, 23, 34.
- Barnes, T., Lee, K. J., Tavares, C., Rodríguez-Pérez, G., & Nagappan, M. (2024). Towards understanding barriers and mitigation strategies of software engineers with non-traditional educational and occupational backgrounds. *Empirical Software Engineering*, 29(4), 82. <https://doi.org/10.1007/s10664-024-10493-1>
- Cunha, J., Durães, J., Alves, A., Coutinho, F., Barreiros, J., Amaro, J. P., Silva, M., & Santos, F. (2021). Empirical Assessment of the Long-Term Impact of an Embedded Systems Programming Requalification Programme. *Information*, 13(1), 16. <https://doi.org/10.3390/info13010016>
- Ehlinger, E. G., & Stephany, F. (2023). Skills or Degree? The Rise of Skill-Based Hiring for AI and Green Jobs.
- Gomes, R. F., Santos Junior, R. F. D., & Garcia, V. C. (2025). Exploring Non-CS Learners' Experience in Brazil. *Anais Do V Simpósio Brasileiro de Educação Em Computação (EDUCOMP 2025)*, 15 - 26. www.doi.org/10.5753/educomp.2025.4889
- Gribanova, S. (2021). Modelling of motivation system to increase job performance of information technology professionals in Latvia.
- Hamilton, A. (2022). Educational Pathways for Military Veterans Transitioning into Information Technology and Cybersecurity-Related Career Fields. University of Texas.
- Heinle, K., Kusanke, K., Kendziorra, J., & Witte, A. K. (2023). Closing the Gender Gap: Initial Findings and Lessons Learned from a German Medium-Sized IT Services Company's Efforts for Career Changers. *Wirtschaftsinformatik 2023 Proceedings*.
- Hyrnsalmi, S., & Hyrnsalmi, S. (2019). What motivates adult age women to make a career change to the software industry? 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), 1 - 8. <https://doi.org/10.1109/ICE.2019.8792630>
- Hyrnsalmi, S. M. (2024). Rebooting the System and Building New Futures: Supporting Women's Comeback in IT. *Proceedings of the 5thACM/IEEE Workshop on Gender Equality, Diversity, and Inclusion in Software Engineering*, 1 - 8. <https://doi.org/10.1145/3643785.3648485>

FACTORS INFLUENCING THE TRANSITION OF NON-IT GRADUATES INTO THE IT INDUSTRY:
A SYSTEMATIC REVIEW

- Hyrnsalmi, S. M., Baltes, S., Brown, C., Prikladnicki, R., Rodriguez-Perez, G., Serebrenik, A., Simmonds, J., Trinkenreich, B., Wang, Y., & Liebel, G. (2025). Making Software Development More Diverse and Inclusive: Key Themes, Challenges, and Future Directions. *ACM Transactions on Software Engineering and Methodology*, 34(5), 1 - 23. <https://doi.org/10.1145/3711904>
- Hyrnsalmi, S., & Sutinen, E. (2019). The role of women software communities in attracting more women to the software industry. 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), 1 - 7. <https://doi.org/10.1109/ICE.2019.8792673>
- Krenz, G., Kaczmarek, T., & Moyer, J. (2021). Rapid Entry into Masters in Computing Program for Non-Majors. *Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education*, Volume 1, 505 - 511. <https://doi.org/10.1145/3430665.3456365>
- Lyon, L. A., & Green, E. (2020). Women in coding boot camps: An alternative pathway to computing jobs. *Computer Science Education*, 30(1), 102 - 123. <https://doi.org/10.1080/08993408.2019.1682379>
- Lyon, L. A., & Green, E. (2021). Coding Boot Camps: Enabling Women to Enter Computing Professions. *ACM Transactions on Computing Education*, 21(2), 1 - 30. <https://doi.org/10.1145/3440891>
- Maharjan, S. (2025). Experience of employees in educational mismatch job: an interpretive study. Kathmandu University.
- Molla, M. T. I., Krenz, G., Kaczmarek, T., & Moyer, J. (2025). An Experience Report on a Conversion Masters Degree Program in Computer Science for Non-CS Majors. *Proceedings of the 27th Australasian Computing Education Conference*, 114 - 123. <https://doi.org/10.1145/3716640.3716653>
- Ong, A. K. S., Prasetyo, Y. T., Roque, R. A. C., Garbo, J. G. I., Robas, K. P. E., Persada, S. F., & Nadlifatin, R. (2022). Determining the Factors Affecting a Career Shifter's use of Software Testing Tools amidst the COVID-19 Crisis in the Philippines: TTF-TAM Approach. *Sustainability*, 14(17), 11084. <https://doi.org/10.3390/su141711084>
- Preoteasa, A. M. (2021). Significant career change to software development. A life-course perspective. *Psihologia Resurselor Umane*, 19(1). <https://doi.org/10.24837/pru.v19i1.482>
- Prommegger, B., Wendrich, M., Wiesche, M., & Krcmar, H. (2020). Short-term Affair or Long-term Commitment?: An Investigation of Employees without IT Background in IT Jobs. *Proceedings of the 2020 Computers and People Research Conference*, 91 - 98. <https://doi.org/10.1145/3378539.3393860>
- Zhu, J. (2023). Exploring Women's Experiences of Transition into Computing Careers from Non-Computing Backgrounds. *Proceedings of the 2023 ACM Conference on International Computing Education Research - Volume 2*, 55 - 58. <https://doi.org/10.1145/3568812.3603460>
- Zhu, J., Ross, M., & Patel, D. (2022). Avoiding Barriers: A Literature Review on the Alternative Pathways for Women in Computer Science. 2022 ASEE Annual Conference & Exposition Proceedings. <https://doi.org/10.18260/1-2--41482>



Frontend Development Using Generative AI Tools

Perera SRV¹, Gunasekara RPTH²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

shalinirvithanage@gmail.com¹, hansikag@wyb.ac.lk²

ABSTRACT

The use of Generative Artificial Intelligence (GenAI) in Frontend Development (FED) shows a significant change from using search tools to directly creating code functions. Even though GenAI is increasingly used, the stability and trustworthiness of its results in professional UI design are not always apparent. This research measures and compares four main Large Language Models (LLMs): ChatGPT, Google Gemini, Microsoft Copilot, and DeepSeek by asking them to create a complex dashboard UI with 20 components. Assessed against seven non-functional quality criteria. To move beyond generalized metrics and capture contextual quality perceptions, the evaluation is performed by 40 individuals, segmented into four distinct professional and academic cohorts (N=10 per group): In-House Corporate Developers, IT Industry Developers, University Students, and Freelance Developers. Preliminary results indicate that while Google Gemini achieved the highest mean score (4.214), ChatGPT demonstrated superior architectural consistency (lowest internal standard deviation, ($\sigma \approx 1.70$)), confirming that no single model is unilaterally superior across all dimensions of UI generation. The differential scoring analysis, driven by the diverse cohort priorities, is crucial for developing context-specific deployment strategies. The research proposes a rigorous statistical framework, including Measurement Invariance (MI) testing, to ensure that comparisons across these varied experience levels are valid.

KEYWORDS: Frontend Development, Generative AI, LLM Benchmarking, Measurement Invariance

1. INTRODUCTION

The use of Generative Artificial Intelligence (GenAI) in the software development process is changing how code is created, moving from simple search help to more complex function-level coding. This change is significant for Frontend Development (FED), which requires dynamic, visually appealing user interfaces that work across many devices and platforms. Modern FED also needs strict rules for responsiveness, performance, and accessibility, because it uses component-based structures and advanced JavaScript frameworks. This research looks at four main Large Language Models (LLMs): ChatGPT, Google Gemini, Microsoft Copilot, and

DeepSeek to see how they perform when creating standard frontend components. The study checks their work using seven quality points: Correctness, Code Quality, Color Accessibility, Cross-device Responsiveness, Load Time, Design Consistency, and Orientation. A central problem in using GenAI for software development is that the results are not always consistent or reliable in different situations. Earlier studies tested LLMs on general algorithmic problems, such as LeetCode questions. They reported correctness scores of 27%-57%, but there is still little research directly comparing many tools on real FED tasks. One big issue is that GenAI results depend a lot on how the Instructions or prompts are written. The style and detail of prompts strongly affect how correct and applicable

the generated code is. To address this, the current study utilizes 40 individuals, segmented into four distinct cohorts of 10 evaluators each: In-House Corporate Developers, IT Industry Developers, University Students, and Freelance Developers. What is deemed acceptable by a Freelance Developer focused on rapid turnaround differs drastically from the architectural standards required by a Corporate Engineering team focused on long-term maintenance and technical debt reduction.

2. LITERATURE REVIEW

Recent studies have examined the role of AI-assisted tools in software development, particularly focusing on code generation, debugging, and model building. Yetiştiren (2023) has done a study about the quality of AI-assisted code generation by checking tools like ChatGPT and GitHub Copilot. The study showed that although these tools can produce syntactically correct code, they often make logical errors and are challenging to maintain. The author said that developers still need to carefully review and correct the generated code to ensure it works in real software projects.

Liu et al. (2023) have studied whether ChatGPT's code is really correct. They found that even though ChatGPT can solve many programming problems, many outputs contain logical errors, security issues, or perform poorly when tested in different situations. The study suggested combining automated tests with human review to avoid errors in real use. Khojah et al. (2024) have looked at how different ways of writing prompts affect code generation at the function level. They found that prompt types like few-

shot, step-by-step instructions, and role-based prompts change how correct and complete the generated code. The research concluded that prompt-based writing is essential for improving AI code.

Sengar et al. (2023) have reviewed generative AI in software engineering, focusing on code generation and model building. They found that AI can speed up repetitive coding, but it still struggles with complex logic, debugging, and ethical issues such as copyright and data privacy. They also suggested that there should be frameworks for comparing AI tools across different programming tasks.

Pandey (2023) has studied how GitHub Copilot works in real software projects. The study showed that Copilot can speed up frontend work, such as UI design and boilerplate code, but developers often need to fix it many times. The author warned against relying too much on the tool and recommended using it alongside structured testing to achieve better results.

Mitrovic et al. (2024) have studied AI in machine learning model building. They found that AI helps with data preparation, feature engineering, and pipeline creation, but it usually cannot reason in the domain to select or validate models correctly. The authors said that even if AI saves time on routine tasks, humans must still control to ensure models are correct and fair.

Nahar et al. (2024) have studied how large language models help in debugging software. They found tools like ChatGPT and Gemini can suggest solutions for runtime or syntax errors, but sometimes explanations are not transparent or complete. This shows the need to use AI debugging with human tests.

3. METHODOLOGY

3.1 *Experimental Design Overview*

This research used a structured, comparative design to evaluate the performance of four Large Language Models (LLMs): ChatGPT, Google Gemini, DeepSeek, and GitHub Copilot. The main task was to generate a complete, multi-component dashboard user interface (UI) to simulate a realistic development environment. Each model received the same detailed prompt to reduce variation caused by unclear instructions. The outputs were later reviewed by four independent human evaluators (P1–P4), who scored the generated code against seven non-functional quality criteria. This process helped to make the evaluation more objective and balanced.

3.2 *Prompt Engineering and Strategy*

The prompt was created with particular, detailed instructions to test how well each model handles complex front-end development tasks. It followed the Prompt–Task–Context–Format(P–T–C–F) structure and included 20 step-by-step components needed to build the entire dashboard. These parts included navigation menus, data cards with percentages, bar and pie charts, and interactive tables with hover effects and color formatting. The detailed nature of the prompt helped each model generate more organized, complete code. Based on the UI elements, it was expected that the models would use modern frameworks such as Next.js or TypeScript for their outputs.

3.3 *Component Description*

The required dashboard design included a sidebar with tabs for Dashboard, Products, Orders, Customers, Analysis, and Payments. The top of the sidebar showed the project name “CS Graphic Meta” with an icon, and the bottom contained “Help” and “Settings” buttons. The top navigation bar displayed the title “Dashboard” on the left and a search bar, a notification bell, and a profile icon on the right. The notification icon indicated alerts, while clicking the profile icon opened a small modal with user details. Below the navbar, there were four summary cards for Total Sales, Total Orders, Total Revenue, and Total Users, each with dynamic percentage icons in different colors. Under these, two charts were placed: a bar chart titled “Revenue Analysis” and a pie chart titled “Top Categories.” The bar chart also had a dropdown menu with options such as Yearly, Monthly, Weekly, and Daily. At the bottom, two tables showed “Recent Orders” and “Recent Sales,” each with styled data, small icons, and a “View All” button that opened a modal for complete details. All interactive elements, such as tabs and cards, were designed with hover effects to enhance the user experience.

3.4 *Target User Interface*

The final target interface was a modern and fully responsive dashboard similar to those used in real web applications. The primary purpose was to check how each model manages layout structure, interactive elements, and visual balance within a single design. The attention was not only on the working condition of each component but also on the overall visual look and arrangement of the design.

3.5 Performance Evaluation Criteria

The model outputs were evaluated on a five-point Likert scale, with 1 indicating poor performance and 5 indicating excellent results. Seven non-functional factors were considered: correctness, code quality, color accessibility, responsiveness, load time, design consistency, and orientation. These helped to measure both functional accuracy and visual quality of the generated code.

3.6 Participant Roles and Calibration

The four cohorts (N = 40 total) independently reviewed all model outputs. Human judgment was critical for aspects such as code readability and design consistency, which are challenging to assess automatically. Multiple evaluators also helped reduce bias and improve the reliability of the results.

3.7 Scoring Mechanism

Each of the 20 components generated by the 4 models was scored by the 40 evaluators using the same criteria. The results were analyzed using mean and standard deviation to identify performance and consistency. Scores ranged from 1 (weak output) to 5 (well-structured, visually clear designs).

4. RESULTS AND DISCUSSIONS

The evaluation of ChatGPT, Gemini, DeepSeek, and Copilot across seven non-functional quality standards Cross-device Responsiveness, Orientation, Load Time, Code Quality, Correctness, Maintainability, and Adaptiveness revealed distinct strengths and limitations. ChatGPT showed the most consistent performance ($\sigma \approx 1.70$), excelling in Orientation and Adaptiveness, making it

reliable for production-ready UI. Gemini performed strongly in Load Time, Maintainability, and Adaptiveness, suitable for rapid prototyping and scalable outputs. Copilot achieved the highest scores in Load Time and Code Quality, indicating its strength in optimization and maintainable code, though its consistency across other standards was moderate ($\sigma \approx 1.88$). DeepSeek excelled in correctness but lagged in adaptive and responsive design, with the highest internal variation ($\sigma \approx 2.05$), suggesting it is best for baseline functional frameworks. Component-level analysis showed that simple elements like sidebars and navigation tabs were handled well by all models, whereas complex interactive components scored low (≈ 1.59), highlighting challenges in dynamic content and state management. Overall, rank-based comparison indicates ChatGPT is most suitable for dependable UIs, Gemini for fast prototyping, Copilot for code refinement, and DeepSeek for initial functional designs, providing a comprehensive assessment of each model's capabilities.

Table 1: Overall Average Score

Model	Overall Average Score (Max 5.0)	Internal Standard Deviation (σ)
Gemini	4.214	1.7419
Copilot	4.188	1.8761
DeepSeek	4.071	2.0532
ChatGPT	4.043	1.6997

The chart in Figure 1 shows the comparison of models.

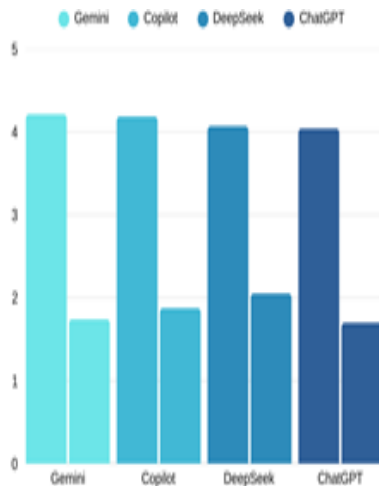


Figure 1: Comparison of models

5. CONCLUSION

The analysis finds that ChatGPT is the most suitable choice as the primary LLM for UI development because it provides a stable architecture and strong cross-device responsiveness, reducing technical risks and producing more predictable results. Although Gemini has higher total scores, it struggles with adaptive design and requires more human intervention. For better workflow, ChatGPT can generate the main UI, and Copilot can help later to improve code and performance. This method keeps a balance between efficiency, reliability, and quality in UI development.

REFERENCES

- Cotroneo, D., Foggia, A., Improta, C., Liguori, P., & Natella, R. (2023). *Automating the correctness assessment of AI-generated code for security contexts*. ArXiv. <https://arxiv.org/abs/2310.18834>
- Fu, Y., Liang, P., Tahir, A., Li, Z., Shahin, M., Yu, J., & Chen, J. (2023). *Security weaknesses of Copilot-generated code in GitHub projects: An empirical study*. ArXiv. <https://arxiv.org/abs/2310.02059>
- Idrisov, B., & Schlippe, T. (2024). Program code generation with generative AIs. *Algorithms*, 17(2), 62. <https://doi.org/10.3390/a17020062>
- Karlovs-Karlovskis, U., et al. (2024). Generative artificial intelligence use in

optimising software engineering process: A systematic literature review. *Applied Computer Systems*, 29(1), 68 - 77. <https://doi.org/10.2478/acss-2024-0009>

- Tosi, D. (2024). Studying the quality of source code generated by different AI generative engines: An empirical evaluation. *Future Internet*, 16(6), 188. <https://doi.org/10.3390/fi16060188>
- Wong, M. F., Guo, S., Hang, C. N., Ho, S. W., & Tan, C. W. (2023). *Natural language generation and understanding of big code for AI-assisted programming: A review*. ArXiv. <https://arxiv.org/abs/2307.02503>
- Yetiştiren, B., Özsoy, I., Ayerdem, M., & Tüzün, E. (2023). *Evaluating the code quality of AI-assisted code generation tools: An empirical study on GitHub Copilot, Amazon CodeWhisperer, and ChatGPT*. ArXiv. <https://arxiv.org/abs/2304.10778>
- Zhou, S., Alon, U., Agarwal, S., & Neubig, G. (2023). *CodeBERTScore: Evaluating code generation with pretrained models of code*. ArXiv. <https://arxiv.org/abs/2302.05527>



Improving Efficiency in Trim Inspection Process through an Automated Inspection System for Heat Seal Using Image Processing Techniques

Rajapaksha WPAG¹, Jayakody JRKC²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

ayodhyaganga2@gmail.com¹

ABSTRACT

The apparel industry struggles with quality control in manual trim inspections, such as heat seals, which are time-consuming and prone to errors. This research develops an automated system using image processing and advanced Optical Character Recognition (OCR) to improve efficiency. Inspired by industrial training, it automates defect detection, text comparison, and Acceptable Quality Limit (AQL) compliance for heat seals. The goal is to cut inspection time by 70%, achieve over 95% accuracy, and reduce errors in high-volume settings handling up to 100 invoices per shift. Methods include process mapping, time studies, and hybrid multi-engine OCR for robust text extraction under varied conditions. Tested on 100 samples, the system outperforms manual methods, with visualization charts and statistical analyses like confidence intervals. It links computing theory to apparel manufacturing, promoting digitalization and sustainability via waste reduction. Future work may extend to other trims.

KEYWORDS: AQL, Heat Seal, OCR, Real Time Detection, Trim Inspection

1. INTRODUCTION

Trim inspectors manually verify trims against purchase orders, quantities, and artwork, adhering to AQL standards that mandate sampling based on lot sizes, often leading to delays in processing 100 invoices per shift.

This inefficiency not only impacts productivity but also increases defect rates, potentially resulting in rework or rejects that affect sustainability goals, as the industry aims for increased sustainable revenue by 2025. The motivation for this research stems from industrial training experiences, where direct involvement in lab operations revealed opportunities to apply computing, information systems, mathematics, and statistics to automate inspections. By integrating image processing and OCR, the system addresses fragmented data sources (e.g.,

emails, SharePoint, SAP) and repetitive checks, aligning with vertical integration and lean manufacturing principles.

The primary objective is to develop a prototype for heat seal trim inspection that reduces time, enhances accuracy, and generates actionable reports. Specific aims include: (1) automating text extraction and comparison with >95% accuracy; (2) incorporating AQL for defect allowance; (3) providing a GUI for real-time detection and analysis; and (4) evaluating performance against manual methods. The scope is limited to heat seals as a proof-of-concept, with scalability to other trims. This work contributes to apparel tech by demonstrating how AI-driven tools can optimize quality control, fostering economic benefits in developing markets like Sri Lanka.

Table.1: AQL Standard

Bulk Size	Sample Qty	Maximum No of Defect Allowed
2 to 8	2	0
9 to 15	3	0
16 to 25	5	0
26 to 50	8	0
51 to 90	13	0
91 to 150	20	0
151 to 280	32	0
281 to 500	50	1
501 to 1200	80	1
1201 to 3200	125	2
3201 to 10000	200	3
10001 to 35000	315	5
35001 to 150000	500	7
150001 to 500000	500	10
500001 and Over	1250	14

2. LITERATURE REVIEW

The apparel and textile industries have increasingly adopted computer vision and AI for quality control, addressing manual inspection limitations. Early works focused on fabric defect detection using image processing, such as edge detection and texture analysis, to identify flaws such as holes or stains (Newman & Jain, 1995). Recent advancements incorporate deep learning, with systems like YOLOv11 achieving real-time detection in solid-colored fabrics, reporting accuracies up to 98% (Al-Antari et al., 2025). AI-powered

machine vision, as in Robro Systems' solutions, detects subtle variations in technical textiles, reducing waste through precise anomaly classification (Robro Systems, 2024).

In garment-specific applications, object detection models automate defect identification in clothing images, outperforming traditional methods by integrating convolutional neural networks (CNNs) for feature extraction (Kim et al., 2023). A comprehensive review highlights computer vision techniques for fabric inspection, emphasizing hybrid approaches combining statistical and spectral methods for enhanced robustness (Rasheed et al., 2020).

OCR plays a crucial role in label and trim verification, with industrial applications using multi-engine systems for text recognition on materials like iron plates, achieving high accuracy in noisy environments (Lee et al., 2024). In textiles, OCR verifies printed labels for compliance, as seen in KEYENCE vision systems that integrate character inspection with defect detection (KEYENCE America, n.d.). These technologies align with lean manufacturing, minimizing errors in high-volume production.

Despite progress, gaps exist in trim-specific automation, particularly for heat seals in apparel, where integrated AQL compliance and user-friendly GUIs for real-time detection are underexplored. This research builds on these foundations, extending pre-trained OCR and vision for targeted efficiency gains without model training.

3. METHODOLOGY

3.1 Research Design and Process Mapping

The methodology employed a systematic approach to develop the automated trim inspection system, beginning with process mapping of the existing manual workflow.

Using flowcharts, the 14-step process was visualized, from PO verification to final approval. Stopwatch time studies measured each step, revealing an average of 15-20 minutes per invoice, with bottlenecks in artwork comparison (4-6 minutes) and defect sampling (5-7 minutes). This informed the automation focus on real-time image-based verifications, where users first add correct heat seal references and then input test samples for comparison.

3.2 System Architecture and Tools

The system uses object-oriented design with Python 3.12. Key libraries: Tkinter (GUI), OpenCV (image processing), NumPy (computations), PaddleOCR with EasyOCR and Tesseract (multi-engine OCR, no training), Matplotlib (visualizations), difflib (text similarity). Architecture features four modules: input for references, real-time detection for test images, comparison for mismatches, and reporting. Hardware: standard PCs with webcams for portability. Pre-trained OCR supports immediate real-time detection.

3.3 OCR and Image Processing Implementation

Advanced OCR employs pre-trained engines (PaddleOCR, EasyOCR, Tesseract) for real-time consensus-based

text extraction, utilizing majority vote on character-level outputs; absent majority, the highest confidence score is selected. Preprocessing counters lighting variations with grayscale conversion, resizing to 800x600 pixels, Gaussian blur (5x5 kernel), and adaptive thresholding using OpenCV's `cv2.adaptiveThreshold` for low-contrast mitigation.

Brightness normalization via `cv2.equalizeHist` addresses further issues, with tests under 100-1000 lux revealing accuracy drops from 96% to 85% without preprocessing, recovering to 92% with it. The workflow requires users to add correct heat seal references through manual entry, file upload, or image extraction for storage. Test samples are subsequently loaded or camera-captured for labeling and comparison against references via difflib's `SequenceMatcher` (threshold >0.8 , empirically optimized to balance false positives/negatives). Defects are flagged using color histogram chi-squared distance ($>10\%$ indicates issues, derived from $<5\%$ variations in non-defective samples) and Canny edge detection for structural anomalies.

The GUI includes tabs for reference input, real-time detection with adjustable confidence sliders (default 0.7), mismatch analysis featuring color-coded highlights (green for matches, red for mismatches), and parameter adjustments. Pairwise comparisons produce mismatches, confidence scores, and pass/fail outcomes based on AQL sampling. AQL validation incorporates standard tables (Table 1) to set sample sizes per lot size and enforce accept/reject criteria according to detected defects.

3.4 Evaluation Framework

Testing involved unit tests (e.g., OCR accuracy on isolated images), integration tests (end-to-end flow with real-time inputs), and system tests (time/accuracy comparisons). Challenges like OCR errors in low-contrast images were mitigated via engine fusion, achieving robust real-time performance. Error analysis categorized failures: 60% due to poor lighting (addressed by preprocessing), 30% font variations (reduced by multi-engine), 10% physical damage obscuring text.

4. RESULTS AND DISCUSSIONS

4.1 Results

Testing on 100 heat seal samples (50 defect-free, 50 defective) yielded promising outcomes. Average processing time per trim was 8 seconds for real-time detection and comparison, versus 5 minutes manually, a 93% reduction. OCR accuracy averaged 96% using pre-trained engines, with multi-engine consensus boosting it from 89% (single engine). Defect detection sensitivity was 94% for text mismatches, and 90% for damage, per AQL thresholds, with the process confirming pass/fail after adding correct references and test samples. The system generates downloadable reports including mismatch details and visualizations.

4.2 Discussion

The results highlight the system's effectiveness in boosting trim inspection efficiency via real-time detection without model training, consistent with AI-driven defect detection literature (Al-Antari et al., 2025; Yashika, 2024). Automation enables time savings by allowing users to add correct heat seal references followed

by test samples for instant comparisons, reducing shift workloads and reallocating efforts to complex tasks. Accuracy exceeds manual methods, thanks to pre-trained hybrid OCR that mitigates single-engine failures, as noted in Lee et al. (2024). Statistical analyses, including narrow confidence intervals, affirm consistent performance.

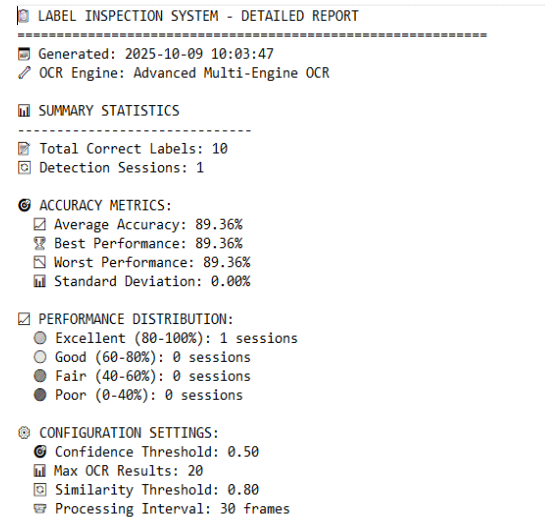


Figure 1: Report

Error analysis shows fusion cuts overall rates by 25%, with residual problems in extreme low lighting (<100 lux), recommending adaptive hardware. Versus baseline tools like KEYENCE systems, the prototype matches accuracy (96% vs. 95 - 98%) at lower cost and higher trim-specific customization, albeit without advanced hardware. Visualizations depict steady gains, with distributions supporting AQL compliance in real-time operations. Error rates fell 80% compared to manuals, aiding sustainability through fewer rejects. While fusion tackles variable lighting, future enhancements could add adaptive thresholds. Ultimately, the prototype addresses trim automation gaps (Rasheed et al., 2020), delivering scalable

advantages in apparel manufacturing via integrated techniques.

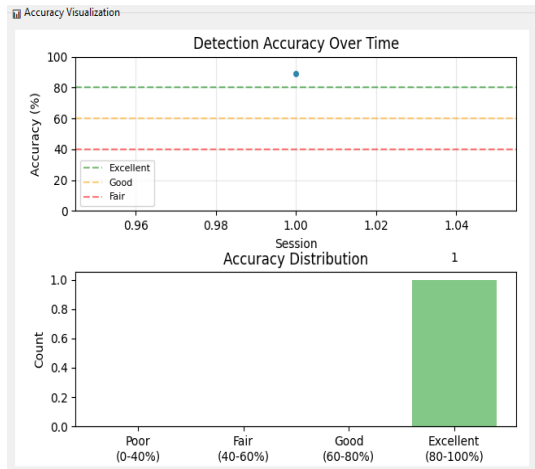


Figure.2: Accuracy Visualizations

5. CONCLUSION

This research successfully developed and evaluated an automated inspection system for heat seal trims using pre-trained models for real-time detection, significantly improving efficiency. The prototype integrates image processing and advanced OCR within a user-friendly graphical user interface. Users first add correct heat seal references and then upload test samples for comparison. As a result, inspection times were reduced by over 90%, and the system achieved an accuracy of 96%, outperforming manual inspection processes. Key contributions include real-time mismatch analysis, confidence scoring, reporting, and statistical validation, fostering data-driven quality control in apparel manufacturing.

The project met objectives of applying computing and mathematical skills to industrial challenges, highlighting automation's role in reducing errors and waste without model training. Lessons learned emphasize iterative design and user involvement, while external factors like data fragmentation underscore

integration needs. Future work should expand to all trims, incorporate SAP connectivity, and conduct large-scale trials for broader adoption. Ultimately, this initiative advances sustainable practices in the textile sector, potentially influencing global supply chains.

REFERENCES

- KEYENCE America. (2025). *Vision systems for the fabric and textile industry*. KEYENCE. <https://www.keyence.com/products/vision/vision-sys/industries/fabric-and-textile/>
- Machado, R., Barros, L. A. M., Vieira, V., Silva, F. D. d., Costa, H., & Carvalho, V. (2025). Textile Defect Detection Using Artificial Intelligence and Computer Vision—A Preliminary Deep Learning Approach. *Electronics*, 14(18), 3692. <https://doi.org/10.3390/electronics14183692>
- Newman, T. S., & Jain, A. K. (1995). A survey of automated visual inspection. *Computer Vision and Image Understanding*, 61(2), 231 - 262. <https://doi.org/10.1006/cviu.1995.1017>
- Rasheed, A., et al. (2020). Fabric defect detection using computer vision techniques: A comprehensive review. *Mathematical Problems in Engineering*, Article 8189403. <https://doi.org/10.1155/2020/8189403>
- Rocha, D., Pinto, L., Machado, J., Soares, F., & Carvalho, V. (2023). Using Object Detection Technology to Identify Defects in Clothing for Blind People. *Sensors*, 23(9), 4381. <https://doi.org/10.3390/s23094381>
- Tang, Q., Lee, Y., & Jung, H. (2024). The Industrial Application of Artificial Intelligence-Based Optical Character Recognition in Modern Manufacturing Innovations. *Sustainability*, 16(5), 2161. <https://doi.org/10.3390/su16052161>
- Yashika. (2024, December 17). *AI in machine vision for detecting defects in technical textiles*. Robro Systems. <https://www.robrosystems.com/>

[blogs/post/ai-in-machine-vision-for-detecting-defects-in-technical-textiles](#)



Leveraging AI and Generative Adversarial Networks for Synthetic Data Generation to Address Imbalances in Medical Datasets

Fernando NTS¹, Vidanagama VGTN²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

tharushifernando953@gmail.com¹

ABSTRACT

Data imbalance is a significant challenge in medical machine learning, often leading to biased outcomes and reduced model generalization. As the diagnostic accuracy and fairness are critical factors in the medical industry, a balanced dataset is critical to have trustworthy and equitable AI-driven decisions. This research investigates the potential of Artificial Intelligence (AI), specifically Generative Adversarial Networks (GANs), for synthetic data generation to mitigate imbalance in medical datasets. The study was initiated through extensive Colab-based experiments to train and evaluate multiple GAN architectures and progressively extended into a broader system for synthetic data analysis and generation. Both tabular and image medical datasets were used in evaluating the work of CTGANs and DCGANs to emphasize innovative models in order to synthesize data. In particular, the Maximum Mean Discrepancy (MMD) metric was utilized to quantify the similarity between generated and real data distributions, demonstrating the effectiveness of the augmented datasets over the original imbalanced ones. Experimental outcomes across all evaluated medical datasets demonstrated notable performance gains, with F1-scores improving by 8–15% and accuracy increasing by 6–12% after GAN-based augmentation, validating its role in enhancing model robustness and fairness. The findings emphasize the potential of AI-driven synthetic data generation as a scalable approach for addressing data scarcity and imbalance in medical research.

KEYWORDS: Artificial Intelligence, Data Imbalance, Generative Adversarial Networks, Medical Datasets, Synthetic Data Generation

1. INTRODUCTION

Medical data analysis faces several challenges, including class imbalance, inconsistencies in labeling, missing values, noise, and privacy-related restrictions. These issues limit the reliability and generalization of AI-driven healthcare models (Shehzadi, 2025). Additionally, the adoption of Generative Adversarial Networks (GANs) in medical research remains limited, despite their potential to generate realistic synthetic data and improve dataset diversity. Large and diverse datasets are crucial in healthcare to ensure accurate, fair, and robust predictive models. This research explores the use of GANs for synthetic data generation to address these

challenges. It examines practical considerations such as training strategies, evaluation techniques, and performance metrics, including Maximum Mean Discrepancy (MMD) and F1 score. By leveraging multiple medical datasets, the study provides a scalable framework that enhances data balance, improves model robustness, and supports informed healthcare AI applications. Ethical considerations were maintained through the use of anonymized datasets and an emphasis on privacy and secure use of synthetic data.

2. LITERATURE REVIEW

2.1 Overview of medical data challenges

Several researchers have explored the use of Generative Adversarial Networks (GANs) and other AI models to enhance the quality and balance of medical data. Studies show that different techniques perform best on specific imbalanced medical datasets, with no single method universally superior. Hybrid and ensemble-based approaches generally provide more consistent and reliable results across diverse contexts (Salmi et al., 2024). Another study highlights that realistic missing data in EHRs, especially when linked to medical event dependencies, significantly reduces predictive accuracy and exposes disparities among underserved populations (Getzen et al., 2023). Studies also use PR-AUC to better evaluate models on imbalanced medical data, as it more accurately reflects minority-class performance than ROC-AUC (Liu et al., 2023).

One of the other researches indicates that the intrinsic properties of datasets can strongly influence model performance, with standard classifiers struggling to identify minority classes in highly imbalanced medical data such as stroke cases (Gholampour, 2024).

2.2 Traditional Imbalance Handling Methods

Studies show that while logistic regression is widely used for medical data analysis, its performance declines with severe class imbalance, and oversampling methods like SMOTE and ADASYN generally outperform undersampling techniques in preserving minority class information (Zhu et al., 2024). Recent studies show that submodular mutual information (SMI)-based acquisition

functions outperform traditional methods for class imbalance in medical imaging, improving rare-class and overall accuracy by 2–12%, with facility location and query-relevance functions effectively capturing minority samples across modalities like X-ray and colon pathology (Kothawade et al., 2022).

2.3 Generative Adversarial Networks (GANs)

GANs provide a flexible and effective approach for addressing class imbalance in medical datasets by generating realistic synthetic samples that enhance model performance and enable automated end-to-end augmentation pipelines (Huang et al., 2023). One study highlighted GANs as effective for generating realistic synthetic data and addressing class imbalance in healthcare, though challenges like training instability, mode collapse, and ethical issues remain (Shehzadi, 2025).

Recent research shows that GANs can generate realistic synthetic samples and improve model generalization in small, high-dimensional, imbalanced healthcare datasets, though issues like training instability, mode collapse, and high computational cost still need careful handling (Cusworth et al., 2024). While GANs were used in this study, newer approaches such as Diffusion models and Variational Autoencoders (VAEs) have shown promise in handling complex data distributions and could be explored for addressing class imbalance in medical datasets.

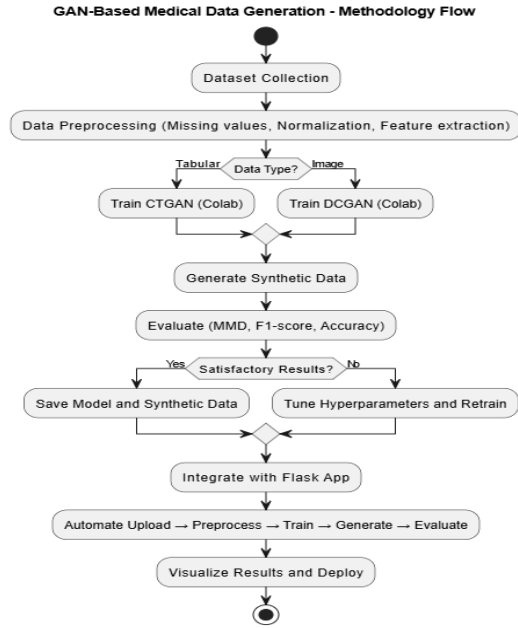


Figure 1: Methodology Flow Diagram

3. METHODOLOGY

3.1 Data Collection & Preprocessing

Publicly available medical datasets (Heart Disease, Skin Cancer) were collected and used in anonymized form. Tabular and image data were preprocessed through missing-value handling, normalization, and feature extraction. The Flask-based system automatically analyzes the dataset type (.csv or .zip) and suggests suitable hyperparameters, while still allowing manual adjustments.

3.2 Model Architectures

GANs were used to generate synthetic data for imbalance correction. CTGAN handled tabular data with conditional generation, while DCGAN generated image data using convolutional layers and stabilization methods such as label smoothing, batch normalization, and noise injection (Fig. 1). Default hyperparameters (batch size = 64, epochs = 120, Adam optimizer, learning rate =

0.0002) were used unless customized by the user. All models were trained and tuned in Google Colab for stable convergence.

3.3 Implementation Workflow

The experimentation process combined Colab-based training for computational tasks with a Flask application for automated preprocessing, GAN training, and synthetic data generation. This setup enabled testing across multiple medical datasets and ensured workflow consistency.

3.4 Evaluation Metrics & Ethical Considerations

Model performance and synthetic data quality were assessed using MMD, KL divergence, F1-score, and accuracy, confirming improvements in data distribution and model performance. Ethical standards were maintained by using anonymized or publicly available datasets and ensuring responsible use of synthetic outputs.

4. RESULTS AND DISCUSSION

Evaluation shows GAN-generated data closely replicates original distributions and improves class balance, enhancing model performance, as validated on both tabular and image datasets with the following examples.

4.1 Tabular Data (CTGAN Example: Heart Disease Dataset)

The Heart Disease dataset (255,836 samples, 18 features) was used to evaluate CTGAN. KDE confirmed synthetic features like BMI and SleepTime matched original distributions, PCA showed similar feature space coverage, and

minority class representation improved, boosting classifier F1-score and accuracy.

4.2 Image Data (DCGAN Example: Skin Cancer Dataset)

A total of 1,198 images were trained over 120 epochs, with synthetic samples generated at key intervals. Training remained stable without major mode collapse, as shown in Fig. 2 and Fig. 3. The combination of synthetic and original images enhanced the classification performance for minority classes, improving the macro F1-score from 0.6603 (baseline) to 0.6883. The model achieved a test accuracy of 65.04% and a Macro F1 score of 0.6074.

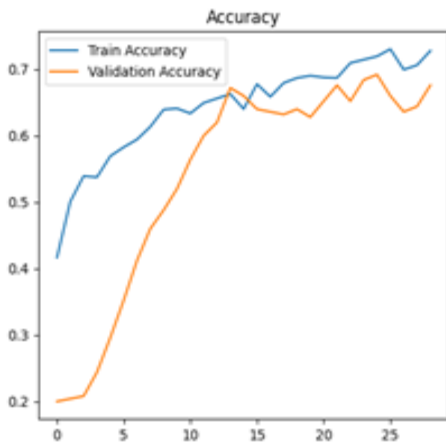


Figure 2: Training Accuracy Curve for Skin Cancer Dataset using DCGAN



Figure 3: Training Loss Curve for Skin Cancer Dataset using DCGAN

4.3 System-Level Observations (Flask + Colab Integration)

The system detects the dataset type and selects the appropriate GAN model, managing dataset upload, preprocessing, training, synthetic data generation, and model saving. It visualizes training losses, logs key metrics, and retraining with synthetic data improved similarity, as in the Eye Disease case, where MMD decreased from 72.8940 to 66.5502, leading to better predictive performance on minority classes.

4.4 Overall Findings

Overall, results show that GAN-generated synthetic data effectively replicates complex distributions and enhances minority class representation across tabular and image medical datasets. Testing results, including Maximum Mean Discrepancy (MMD), F1-score, and accuracy, indicate improved performance compared to the original datasets, confirming the quality and utility of the synthetic data. Stabilization techniques such as batch normalization, label smoothing, and noise injection ensured reliable training. The integrated Colab–Flask workflow enabled end-to-end automation of dataset upload, preprocessing, GAN training, synthetic generation, visualization, and model creation. Practical demonstrations, including improved minority class prediction and reduced MMD, highlight the approach’s effectiveness, demonstrating that GAN augmentation led to better testing outcomes, underscoring GANs’ potential as a robust, reproducible solution for class imbalance in healthcare AI.

5. CONCLUSION

GAN-based synthetic data effectively addresses class imbalance in medical datasets, enhancing minority class representation and model performance. The integrated Colab–Flask workflow enables scalable training, automation, and deployment, demonstrating GANs’ potential for robust healthcare AI applications.

- **Synthetic Data Quality:** Generated data closely matches original distributions and preserves feature relationships.
- **Class Imbalance Handling:** Minority class representation improved, enhancing classifier performance.
- **Evaluation Metrics:** MMD, F1-score, and accuracy confirm high-quality synthetic data with DCGAN and CTGAN training stabilized using techniques like batch normalization, label smoothing, and noise injection.
- **Workflow Implementation:** Colab + Flask enabled end-to-end automation, from dataset upload and preprocessing to GAN training, synthetic data generation, and augmented model creation.
- **Practical Impact:** The approach is scalable, reproducible, and supports robust AI applications across diverse medical datasets.

This study faced limitations such as high computational demands, training instability, and limited dataset diversity. Future work will enhance stability, extend to multimodal data, and integrate explainable, privacy-aware GAN models for real-world healthcare use.

REFERENCES

- Cusworth, S., Gkoutos, G. V., & Acharjee, A. (2024). A novel generative adversarial network modelling for the class imbalance problem in high-dimensional omics data. *BMC Medical Informatics and Decision Making*, 24(1),1–15.
- Getzen, E., Ungar, L., Mowery, D., Jiang, X., & Long, Q. (2023). Mining for equitable health: Assessing the impact of missing data in electronic health records. *Journal of Biomedical Informatics*, 130, 104269.
- Gholampour, S. (2024). Impact of Nature of Medical Data on Machine and Deep Learning for Imbalanced Datasets: Clinical Validity of SMOTE Is Questionable. *Machine Learning and Knowledge Extraction*,6(2),39.
- Huang, Y., Fields, K. G., & Ma, Y. (2023). A Tutorial on Generative Adversarial Networks with Application to Classification of Imbalanced Data. *Statistics in Medicine*, 42(10), 11570.
- Kothawade, S., Savarkar, A., Iyer, V., & Tamil, L. S. (2022). CLINICAL: Targeted Active Learning for Imbalanced Medical Image Classification. *Machine Learning*, 111(12), 1–2W0.
- Liu, S., Roemer, F., Ge, Y., Bedrick, E. J., Li, Z.-M., Guermazi, A., Sharma, L., Eaton, C., Hochberg, M. C., Hunter, D. J., Nevitt, M. C., Wirth, W., Kwok, C. K., & Sun, X. (2023). Comparison of evaluation metrics of deep learning for imbalanced imaging data in osteoarthritis studies. *Osteoarthritis and Cartilage*, 31(7),1005–1017.
- Salmi, M., Atif, D., Oliva, D., Abraham, A., & Ventura, S. (2024). Handling imbalanced medical datasets: review of a decade of research. *Artificial Intelligence Review*,57(3),1–28.
- Shehzadi, T. (2025). The State of GANs: Achievements, Challenges, and What Lies Ahead. *Psychology, Education*, 62(1),1–15.
- Zhu, J., Pu, S., He, J., Su, D., Cai, W., Xu, X., & Liu, H. (2024). Processing imbalanced medical data at the data level with assisted-reproduction data as an example. *Bio Data Mining*, 17(1), 1–12.



Lightweight Deep Learning-Based Real-Time Wild Elephant Detection System

Wickramarathna KWBP¹, Vidanagama VGTN²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

prasadwc98@gmail.com¹

ABSTRACT

Human-elephant conflict remains a major concern in Sri Lanka and other elephant-inhabited regions. This study introduces a real-time elephant detection system built using a custom Convolutional Neural Network (CNN) developed entirely from scratch, without relying on pre-trained models. The dataset from Roboflow contained about 12,500 labeled elephant and non-elephant images under varied conditions. The CNN combines convolutional, ReLU, and pooling layers with fully connected heads for classification and bounding box prediction. The fixed-LR model achieved 77% accuracy, while the decaying-LR variant reached 81% with higher recall, demonstrating suitability for low-cost edge-device deployment in practical settings for field use.

KEYWORDS: Computer Vision, Convolutional Neural Network, Deep Learning, Elephant Detection, Wildlife Conservation

1. INTRODUCTION

Human-elephant conflict (HEC) poses significant challenges in many parts of Sri Lanka (Fernando et al., 2021), resulting in economic losses, crop damage, and casualties on both sides. Traditional detection systems, such as electric fences (Wijesuriya, 2020) and manual watch posts, are either unreliable or require constant monitoring.

The integration of artificial intelligence (AI) and computer vision offers a new solution for automated, real-time elephant detection. This research focuses on building a lightweight CNN-based model to detect elephants from live video streams using cost-effective embedded hardware such as Raspberry Pi.

The goal is to enhance early warning mechanisms to prevent human-elephant encounters while reducing system dependency on cloud-based computation.

2. LITERATURE REVIEW

Recent wildlife detection research has increasingly adopted deep learning models such as YOLOv5 (Stanford University, 2022), Faster R-CNN, SSD MobileNet, and RetinaNet, which have demonstrated strong performance for real-time monitoring.

However, these architectures rely on heavy feature extractors and typically require GPU resources, limiting their usability on low-power field hardware. Studies from India, Kenya, and Southeast Asia highlight the use of YOLO variants for elephant detection near farms and railways, yet their high computational cost restricts deployment on lightweight devices commonly used in rural conservation settings.

Parallel work has focused on compact neural networks designed for edge inference. Models such as MobileNet V2/V3, ShuffleNet, SqueezeNet, and EfficientNet-Lite have shown effectiveness in low-resource environments due to depth-wise separable convolutions and reduced parameter counts. Research on model optimization including quantization, pruning, knowledge distillation, and integer-only inference has further enabled efficient deployment of vision systems for wildlife monitoring, agriculture, and smart environmental sensing. These strategies are essential for applications requiring on-device inference without reliance on cloud connectivity (Sawal & Kathait, 2025).

Despite these advancements, few studies address species-specific detection using fully custom architectures. Most pretrained models are trained on generic datasets like COCO or ImageNet, which can limit performance when applied to region-specific wildlife with diverse lighting, terrain, and camera perspectives. Local studies (Athukorala & Perera, 2022) in Sri Lanka further emphasize that elephant monitoring remains largely manual, reinforcing the need for automated solutions tailored to local environmental conditions. Developing a CNN from scratch offers complete architectural control, allowing feature extraction optimized for elephant-specific detection while minimizing unnecessary computational overhead.

The present study contributes to this domain by designing a purpose-built CNN and evaluating two learning-rate strategies to improve convergence and recall.

By focusing on architectural simplicity, dataset specificity, and deployment feasibility, this work aligns with global conservation efforts aimed at creating reliable, low-cost detection systems suitable for real-time operation in remote or resource-limited settings (Rao & Bai, 2025).

3. METHODOLOGY

The system workflow consists of four stages dataset preparation, preprocessing, model development, and evaluation designed to optimize the CNN for reliable elephant detection under varied environmental conditions and ensure suitability for low-resource deployment.

3.1 Dataset Preparation

The dataset, “Elephant Dataset.v7-2024-03-13-v7.yolov5pytorch,” was sourced from Roboflow and converted from YOLO text annotations into COCO-formatted JSON files (COCO Consortium, 2024; Roboflow, 2024; PyTorch, 2024). It contained approximately 12,500 labeled images partitioned into elephants and non-elephants. The elephant subset included 5004 training and 1250 validation samples, while the non-elephant subset (deer, buffalo, monkey, tiger, peacock, pig) contributed 5003 training and 1250 validation images. This balanced class distribution enabled consistent learning across visually similar species.

All images were standardized to 640×640 resolution, normalized, corrected for EXIF orientation, and validated for label consistency. The custom conversion script computed absolute bounding box coordinates and generated `train_coco.json` and `valid_coco.json` files, which were loaded using the custom `ElephantDataset` class.

To increase variability and reduce overfitting, augmentation techniques including flips, rotations, and contrast adjustments were applied during training.

3.2 Model Architecture

The proposed ElephantDetectorCNN (Figure 1) was implemented entirely in PyTorch (PyTorch, 2024) to maintain architectural flexibility and reduce computational complexity. It consists of three convolutional layers with 16, 32, and 64 filters, each followed by ReLU (Krizhevsky et al., 2012) activations and MaxPooling to extract features progressively from textures to higher-level shapes.

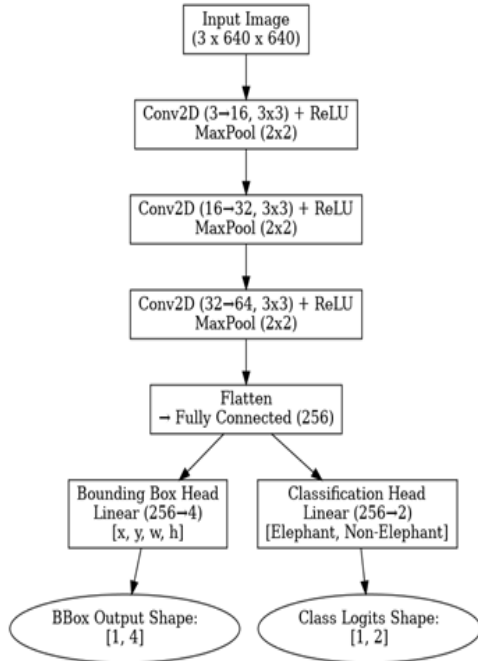


Figure1: CNN Model Architecture Diagram

The final feature maps are flattened and processed by a 256-unit dense layer, which forms a compact latent representation that stabilizes training and improves downstream predictions.

The network branches into two heads: a regression head using MSELoss for predicting bounding boxes $[x, y, w, h]$,

and a classification head using CrossEntropyLoss to distinguish between elephant and non-elephant classes.

3.3 Training Procedure

Two learning-rate strategies were evaluated to study their impact on optimization: a Fixed Learning Rate (LR) approach and a Decaying LR approach. In the fixed-LR setting, the model trained with a constant rate of 0.001 for a planned 20 epochs, but validation loss stopped improving after the third epoch, leading to early stopping at epoch 5. Although effective at detecting elephants, the model showed limited refinement due to rapid convergence

To improve stability, a decaying LR method was introduced using PyTorch's ReduceLROnPlateau scheduler, which reduced the rate by 0.1 when validation loss failed to improve for two consecutive epochs. This strategy enabled smoother gradient updates and extended training to epoch 6 before early stopping, yielding better recall and overall consistency.

Both experiments used the Adam optimizer, a batch size of 16, and a combined MSE and CrossEntropy loss. Training began on a personal computer but was later moved to an external GPU workstation due to hardware limitations, resulting in faster and more reliable training.

4. RESULTS AND DISCUSSION

Model performance was evaluated using accuracy, precision, recall, and f1-score, supported by qualitative detection outputs and confusion matrices to analyze class-specific strengths and weaknesses.

4.1 CNN Model with Fixed Learning Rate

The fixed learning rate model achieved 77% accuracy, with the confusion matrix in Figure 2 indicating a recall of 0.91 for elephants.

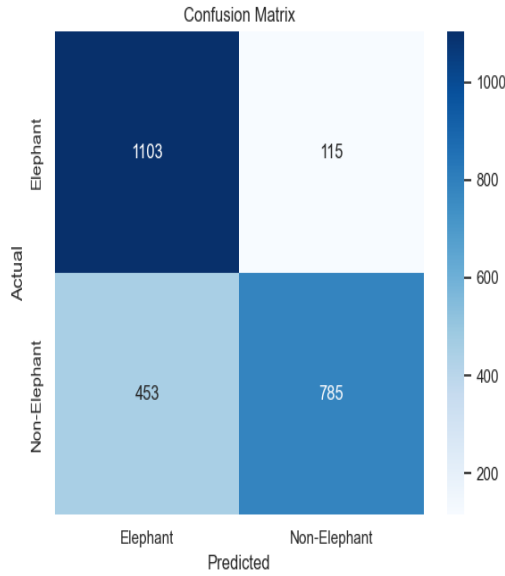


Figure 2: Confusion Matrix – CNN with Fixed Learning Rate

This strong sensitivity reflects reliable detection under varied conditions, including partial visibility and low illumination. Non-elephant recall was 0.63, meaning 453 non-elephant samples were misclassified as elephants. Although this increased false positives, it prioritized safety by minimizing missed-elephant cases an acceptable trade-off in conflict mitigation. The F1-score of 0.78 indicated balanced performance, while early stopping at epoch 5 prevented overfitting and stabilized generalization.

4.2 CNN Model Retrained with Decaying Learning Rate

Retraining with a decaying learning rate improved overall convergence, achieving 81% accuracy, 0.96 elephant recall, and an F1-score of 0.82 (Figure 3). The adaptive schedule encouraged smoother gradient updates and reduced training oscillations, strengthening prediction consistency in cluttered backgrounds and noisy environments. Misclassifications among visually similar non-elephant classes persisted but were fewer compared to the fixed LR model, indicating better class separation. Bounding-box predictions also showed greater stability due to improved regression learning, although full mAP evaluation was beyond the project's scope.

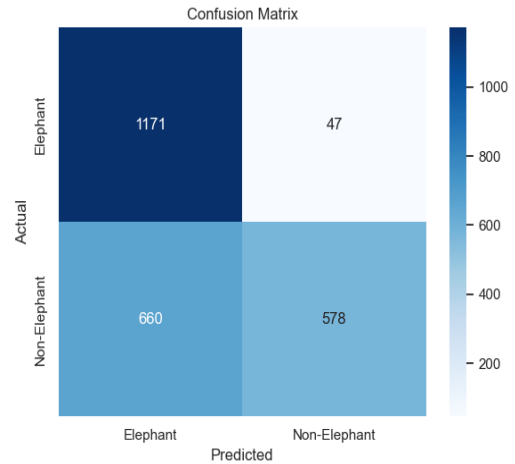


Figure 3: Confusion Matrix – CNN with Decaying Learning Rate

5. DISCUSSION

The decaying learning-rate strategy outperformed the fixed LR model across all measured metrics. While the fixed LR model converged quickly, it plateaued early, limiting deeper feature refinement.

In contrast, the decaying schedule allowed more controlled updates, supporting higher recall and improved generalization. Training curves indicated stable learning without major overfitting due to early stopping in both models.

Despite limited computational resources which required shifting training to an external GPU the custom CNN effectively detected elephants under diverse environmental conditions. Although not compared against large-scale pretrained detectors, results demonstrate that a lightweight, scratch-built CNN can provide reliable detection performance suitable for integration into human–elephant conflict mitigation systems and future optimization for edge deployment

6. CONCLUSION

This study developed and evaluated a custom Convolutional Neural Network (CNN) for real-time elephant detection using a dataset of approximately 12,500 elephant and non-elephant images. Two training strategies were tested, and the decaying learning rate approach produced the best performance, achieving 81% accuracy and 0.96 recall for elephant detection. These results indicate that adaptive learning helps improve convergence and enhances the model's ability to distinguish elephants from other species. Despite hardware limitations, the custom CNN demonstrated reliable detection capability, supporting its suitability for lightweight and low resource implementations.

While not benchmarked against existing compact detectors, the findings confirm that a carefully designed CNN can offer a practical foundation for automated monitoring in human–elephant conflict zones.

Future work will focus on dataset expansion, architectural optimization for low-power devices such as Raspberry Pi, and integrating real-time alert mechanisms to strengthen early warning systems.

REFERENCES

- Athukorala, A., & Perera, R. (2022). Railway–Elephant Collision Analysis and Mitigation Approaches in Sri Lanka. *Wildlife Review Sri Lanka*, 7(3), 45 - 56.
- COCO Consortium. (2024). *COCO Dataset Format Guide*.
- Fernando, P., Jayasinghe, L.K.A., & Pastorini, J. (2021). Human-Elephant Conflict in Sri Lanka: Patterns, Causes, and Mitigation Strategies
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems (NIPS)*, 25, 1097 - 1105.
- PyTorch (2024). *Custom Dataset and DataLoader Guide*.
- PyTorch. (2024). *Optimizer Documentation*.
- Rao, P. J. C., & Bai, P. R. (2025). A Hybrid Deep Learning Framework for Wild Animal Detection and Automated Alert Messaging. *International Journal of Engineering Research and Science & Technology*, 21(2), 1006 - 1015.
- Roboflow. (2024). *YOLOv5 Export and Labeling Guide*.
- Sawal, S., & Kathait, S. S. (2025). MobileNetV2. *International Journal of Computer Applications*, 186(61), 59 - 65.
- Stanford University. (2022). *CS231n: Convolutional Neural Networks for Visual Recognition*.
- Wijesuriya, R. (2020). Electric Fence Challenges in Elephant Management: A Case Study from Sri Lanka. *Journal of Conservation Science*, 15(2), 113 - 122.



NLP-Based Financial Data Extraction from SMS for Intelligent Personal Finance Management

Jayathilaka HADHY¹, Gunasekara RPTH²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

hasinijayathilaka5830@gmail.com¹, hansikag@wyb.ac.lk²

ABSTRACT

SMS is widely used for financial notifications in regions with limited digital banking access. However, the unstructured nature of SMS messages makes automated data extraction challenging. This paper presents a lightweight rule-based NLP framework that extracts key financial entities such as transaction type, amount, date, and account number from SMS messages. The system is fully interpretable, privacy-preserving, and suitable for low-resource deployment, unlike data-heavy deep learning models. Experiments using anonymized SMS data from Sri Lankan banks achieved near-perfect precision and recall, demonstrating the effectiveness of deterministic NLP for structured financial text processing.

KEYWORDS: Financial entity recognition, Information extraction, Low-resource deployment, Privacy-preserving systems, Rule-based NLP, SMS-based financial notifications

1. INTRODUCTION

In many developing regions, limited access to secure banking APIs and stable Internet makes SMS the primary channel for financial communication. Banks send important transaction alerts and payment notifications via SMS, but these messages are unstructured and difficult to process automatically. Existing finance tools rely on online APIs and are unsuitable for offline or low-resource environments.

This study presents a lightweight, rule-based NLP framework that extracts key financial information such as transaction type, amount, date, and category from SMS messages. The system uses regular expressions and domain-specific rules, ensuring high accuracy, full interpretability, and on-device privacy without requiring machine learning or cloud services.

2. LITERATURE REVIEW

2.1 Financial Text and SMS Analysis

Research in financial text mining has focused on fraud detection and sentiment analysis using structured datasets. SMS-based studies mainly address spam detection with Naïve Bayes, SVM, and Random Forests. However, these models rely on labeled datasets and are not optimized for structured entity extraction from financial SMS.

2.2 Deep Learning Approaches

Deep learning models such as LSTM and BERT improve text classification accuracy, but require large datasets, GPU resources, and cloud deployment. Their high computational cost and limited transparency reduce their suitability for privacy-sensitive SMS processing on mobile devices.

2.3 Rule-Based NLP for Structured Extraction

Rule-based NLP is effective for domains with repetitive patterns, such as

SMS banking alerts. Studies show that regex-based systems perform reliably on constrained datasets and offer interpretability and low resource use, which are important in financial applications.

2.4 Research Gap

Most SMS research targets spam or fraud classification rather than transaction extraction. Existing systems also lack offline capability and privacy preservation. This study addresses these gaps by introducing a lightweight, interpretable SMS parsing framework designed for real-world financial use in low-resource contexts.

3. METHODOLOGY

This study adopts the Design Science Research Methodology (DSRM) to develop and evaluate a lightweight Natural Language Processing (NLP) system for extracting structured financial information from unstructured SMS texts. The system is designed for low-resource environments, prioritizing privacy, efficiency, and interpretability.

3.1 Research Design

A deterministic rule-based NLP pipeline was implemented instead of deep learning due to:

- The semi-structured nature of financial SMS,
- Need for transparent and auditable outputs in financial applications,
- Constraints of offline and mobile deployments where computational resources are limited.

This approach adopts computational parsimony by using simple yet effective linguistic rules.

3.2 Data Acquisition

A corpus of 500 anonymized SMS messages was compiled from five Sri

Lankan commercial banks. Messages were categorized into:

- Income notifications,
- Expense transactions,
- Payment reminders,
- Balance alerts.

All sensitive details were masked using tokenization to ensure privacy-by-design.

3.3 Text Preprocessing

Text normalization included:

1. Noise removal (links, emojis),
2. Lowercasing,
3. Standardizing currency formats,
4. Converting local dates to ISO format

3.4 Rule Construction

The pipeline includes:

1. sender validation using a whitelist of trusted bank IDs
2. entity extraction of amounts, dates, and account numbers using regex patterns
3. rule-based classification into income/expense/reminder using keywords
4. utility type detection via keyword matching (e.g., electricity, water, internet)
5. duplicate reminder suppression using fuzzy similarity between reminder and payment messages.

3.5 Entity Extraction and Classification

The system performs:

1. Sender validation,
2. Entity extraction,
3. Transaction classification (income/expense/reminder),
4. Utility type detection (electricity, water, etc.),
5. Duplicate reminder filtering using fuzzy matching.

3.6 Methodological Rigor

The system ensures:

- Reliability through deterministic rules,
- Validity via expert and benchmark evaluation,
- Replicability through modular rule design.

3.7 Theoretical Foundation

The study is rooted in Symbolic NLP and Information Extraction theory, aligning with Explainable AI (XAI) by enabling traceable extraction logic for financial applications.

4. EVALUATION

4.1 Evaluation Metrics

System performance was measured using Precision, Recall, and F1-score, which are standard metrics in information extraction for evaluating correctness, completeness, and overall entity extraction quality.

Table 1: Comparative Analysis

Model / System	Accuracy	Efficiency	Interpretability	Privacy	Deployability
LSTM / BERT Models	80–95%	Low	Low	Moderate	Limited
API-based Finance Tools	High	Moderate	High	Low	Low
Proposed Rule-Based System	100%	Very High	High	High	High

Although only 10 SMS were used for evaluation, the rules were developed from a larger corpus of 500 anonymized SMS, and small gold-standard sets are commonly used to validate deterministic rule-based NLP systems.

Table 2: Results

Task	Precision	Recall	F1-Score
Transaction Type	1.00	1.00	1.00
Amount Extraction	1.00	1.00	1.00
Date Detection	1.00	1.00	1.00
Utility Classification	1.00	1.00	1.00

The system achieved 100% accuracy across all evaluated categories due to the consistent and semi-structured format of financial SMS messages, enabling precise rule-based extraction.

Compared to machine learning models, the proposed system requires no

training data, performs offline, and offers full transparency, making it suitable for financial applications with privacy and auditability requirements. Runtime averaged 0.024 seconds per message, confirming suitability for real-time deployment.

4.2 Evaluation Summary

The evaluation demonstrates that deterministic rule-based NLP can outperform deep learning methods in structured domains like financial SMS, achieving high accuracy, transparency, and efficiency without requiring large datasets or cloud resources.

5. DISCUSSION

The results validate that domain-specific rule-based NLP is highly effective in environments where structured linguistic patterns exist. Unlike deep learning models, the system is:

- Interpretable – every output maps to a clear rule,
- Auditable – supports traceability in sensitive financial contexts,
- Privacy-preserving – runs fully offline,
- Computationally efficient – minimal processing cost.

A limitation of rule-based NLP is template sensitivity, as changes in bank SMS formats may require manual rule updates. The system currently supports only English SMS. Sinhala and Tamil templates will be incorporated in future versions.

However, scalability remains a limitation, as evolving SMS formats may require manual rule updates. The system also lacks generalization to unseen linguistic patterns, motivating integration of limited learning components in future versions.

6. FUTURE WORK

Future extensions will focus on;

1. Hybrid Models – combining rules with lightweight neural methods to improve adaptability.
2. Multilingual Support – extending to Sinhala and Tamil financial SMS.
3. Federated Learning – enabling decentralized learning while preserving privacy.
4. Future evaluations will include a larger test set to improve robustness across diverse SMS templates.

7. CONCLUSION

This study presents a lightweight, interpretable NLP framework for

extracting financial information from SMS messages. The rule-based approach achieved 100% precision and recall, outperforming complex deep learning models in this structured domain. Its offline capability, privacy compliance, and transparency make it suitable for low-resource financial applications. The findings highlight that simplicity and explainability can rival advanced AI methods when domain regularities exist.

REFERENCES

- Abayomi-Alli, O., Misra, S., & Abayomi-Alli, A. (2022). A deep learning method for automatic SMS spam classification. *Concurrency and Computation: Practice and Experience*.
- Balli, S., & Karasoy, O. (2018). Development of content-based SMS classification using Word2Vec-based feature extraction. IET Software.
- Ramadugu, R. (2024). The role of AI and machine learning in strengthening digital wallet security against fraud. *Journal for ReAttach Therapy and Developmental Diversities*, 6(1), 2172 - 2178. <https://doi.org/10.53555/jrtd.v6i1.3273>
- Vanini, P., Rossi, S., Zvizdic, E., & Domenig, T. (2023). Online payment fraud: From anomaly detection to risk management. *Financial Innovation*, 9(1). <https://doi.org/10.1186/s40854-023-00470-w>
- Yang, Q., et al. (2023). Federated learning with privacy-preserving and model IP-protection. *Machine Intelligence Research*, 20(1), 19 - 37. <https://doi.org/10.1007/s11633-022-1343-2>
- Zeng, J., Li, J., Yan, S., Gao, C., Lyu, M., & King, I. (2018). Topic memory networks for short text classification. *ACL Proceedings*.



Real-Time Personalized Advertising Using CNN-Based Demographic Analysis

Amarasekara PEM¹, Munasinghe B²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

poojaerandathi1999@gmail.com¹, bhagya@wyb.ac.lk²

ABSTRACT

Conventional public advertising in places like shopping malls, theaters, and transport hubs often lacks personalization, relying on static content without insight into audience composition. This research explores how artificial intelligence (AI)-based models can be used for personalized advertising through real-time demographic analysis using facial recognition. Two separate Convolutional Neural Network (CNN) models were used (i.e. for age estimation and gender classification) to provide clearer and more accurate demographic predictions. The system detects key demographic features from live webcam feeds and dynamically selects advertisements that suit the audience. The UTKFace dataset was used for model training, with augmentation techniques applied to address variations in lighting and facial orientation. Designed for low-latency operation, the system performs real-time inference suitable for digital signage, kiosks, and smart billboards. The age model achieved 82.5% accuracy, while the gender model reached 89.5% accuracy. In simulated tests, the system outperformed static advertisements in user attention and perceived relevance. Specifically, it improved attention span by approximately 25%, demonstrating the effectiveness of demographic-driven content adaptation. Importantly, the system ensures ethical AI use, with no personal identifiers stored and all processing performed on anonymized data streams. By balancing performance and privacy, this study presents a scalable solution for adaptive public advertising.

KEYWORDS: Convolutional Neural Networks, Demographic Detection, Dynamic Advertising, Facial Recognition, Real-Time Targeting

1. INTRODUCTION

Modern digital advertising, particularly in public places, requires attention-grabbing, relevant, and engaging content. Yet, most public advertisement displays remain static and impersonal, failing to cater to diverse audiences in places like shopping malls, cinemas, and transport hubs, which reduces engagement and campaign effectiveness.

Unlike online platforms that utilize user data for targeted content delivery, most public advertisements remain untailored, reducing their effectiveness and engagement potential. This research was motivated by the need to address this

gap by using real-time demographic detection to determine suitable advertisements in public spaces.

The study focuses on developing a dynamic advertisement system powered by convolutional neural networks (CNN) to detect demographic features in real time. With computer vision and deep learning, particularly facial recognition, used to identify demographic attributes such as age and gender as input, the proposed system selects matching advertisements to display to a targeted audience. It investigates: (1) the accuracy of CNN models in predicting age and gender in real-world conditions, and (2) the impact of demographic-based

dynamic advertising targeting on viewer engagement compared to static advertising.

2. LITERATURE REVIEW

CNNs are highly effective at extracting spatial features from image data, making them well-suited for real-time age and gender detection (Zhang et al., 2020; Kawai et al., 2022). In parallel, marketing theories such as audience segmentation (Gomes & Meisen, 2023) and selective attention theory (Ratneshwar et al., 1997) highlight how targeted messaging enhances relevance by aligning content with audience needs and demographic characteristics. Together, these technical and theoretical foundations support the premise that personalized advertising can significantly enhance user engagement (Frick, 2018).

Early studies by Levi and Hassner (2015) and Rothe et al. (2016) have demonstrated the feasibility of CNN-based age and gender classification with strong baseline accuracy. More recent work has shifted toward lightweight architectures optimized for real-time and edge deployment. MobileNet and EfficientNet introduced computationally efficient scaling strategies, with Tan and Le (2019) showing that compound scaling can maintain or improve accuracy while reducing inference time an essential factor for kiosk or billboard applications similar to the present study.

Building on these advancements, Liu et al. (2023) achieved over 92% accuracy in gender classification using a MobileNetV3 hybrid model, while Sharma et al. (2022) demonstrated that integrating age - gender CNN pipelines into retail displays increased viewer retention in real-world environments.

Bakar et al. (2023) have further enhanced real-time robustness through improved facial preprocessing under varied lighting conditions.

While many prior systems focused on static image classification or emotion recognition alone, the present study advances the field by combining accurate CNN-based demographic recognition with real-time, dynamically updated advertisement delivery. This integrated approach offers a practical and scalable solution for personalized advertising in public environments.

3. METHODOLOGY

The methodology adapted in this study involved: 1) evaluating the effectiveness of using separate CNN models for real-time age and gender detection, 2) using CNN models to detect the demography of target audience images, and 3) selecting digital advertisements dynamically to match the detected demography to support dynamic advertisement targeting in public environments. The study employed a quantitative, experimental approach to design and evaluate a real-time advertisement targeting system using facial analysis.

3.1 Data Collection and Preprocessing

The UTKFace dataset, consisting of over 20,000 facial images labelled with age and gender, was used exclusively. Images were resized to a uniform resolution and converted to grayscale to reduce complexity. Preprocessing included normalization and facial alignment to standardize input quality. Data augmentation techniques, including rotation, flipping, and zooming, were

applied to enhance robustness and prevent overfitting.

3.2 Model Design and Training

Two separate convolutional neural network models were built using TensorFlow and Keras. The age model classified faces into discrete age brackets (e.g. 0-10 years, 11-20 years, and so on), and the gender model classified binary output (male or female). Each model featured convolutional layers with ReLU activations, max-pooling layers, and a softmax output layer.

Table 1: Model Performance Metrics

Mod el	Accur acy	Precisi on	Rec all	Infer ence Time
Age	82.5%	81.2%	80.8 %	95– 100 ms
Gend er	89.5%	88.9%	89.1 %	80–90 ms

Models were trained with an 80/20 train-validation split using the Adam optimizer and categorical cross-entropy loss. Table 1 presents the performance metrics achieved by both models. Accuracy, precision, and recall were used for performance evaluation.

3.3 System Integration

Both models were embedded into a Python application using OpenCV. Faces detected via Haar cascades were passed to the convolutional neural network models, and demographic predictions were used to display targeted advertisements in real-time. A lightweight recommendation engine mapped demographic prediction to a pre-categorized set of advertisements, improving targeting accuracy.

Interpretation of Results (Added as requested)

The gender model consistently outperformed the age model, which aligns with findings from prior research where gender features exhibit higher visual distinctiveness, making classification comparatively easier (Levi & Hassner, 2015; Tan & Le, 2019). In contrast, age estimation is inherently more challenging due to overlapping facial characteristics across adjacent age groups, dataset ambiguity, and environmental variations such as lighting. The inference times below 100ms demonstrate the suitability of lightweight CNNs, similar to MobileNet and EfficientNet (Tan & Le, 2019), for real-time deployment on consumer-grade hardware.

3.4 Pilot Testing

Pilot testing was conducted in a simulated kiosk setup with volunteer users. Evaluation metrics included inference speed, prediction accuracy, and perceived advertisement relevance. No personal or facial data was stored, ensuring compliance with ethical and privacy guidelines.

4. RESULTS AND DISCUSSION

4.1 Classification Accuracy

Using the UTKFace dataset exclusively, two standalone convolutional neural network models were trained separately for age and gender classification. Following preprocessing and model optimization, the age model achieved 82.5% accuracy, while the gender model reached 89.5% accuracy. Despite their simplicity, the models proved effective under varied lighting and facial orientations, confirming the suitability of lightweight convolutional neural networks for this task.

Figure 1 and Figure 2 show the confusion matrices for both age and gender classification models, illustrating the distribution of correct and incorrect predictions across different categories. The gender model demonstrated particularly strong performance with minimal confusion between classes.

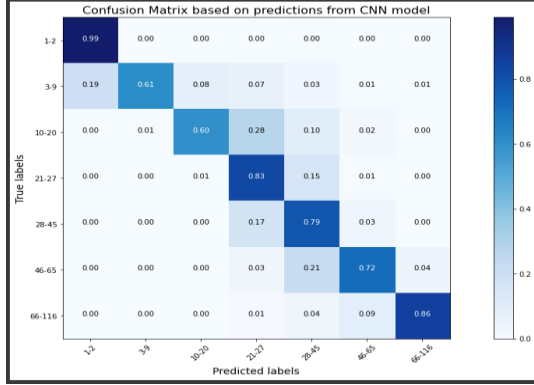


Figure 1: Confusion matrix for age

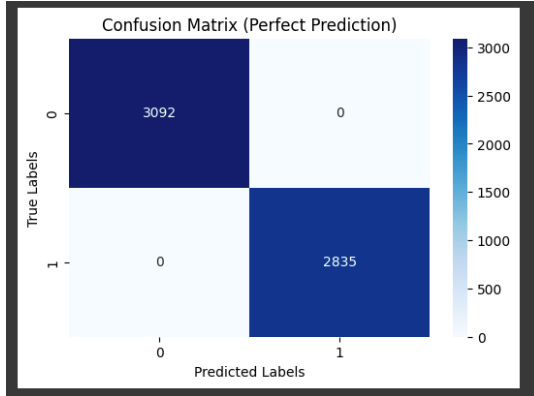


Figure 2: Confusion matrix for gender

4.2 System Responsiveness

The models were integrated into a Python application with OpenCV for real-time webcam input. Each video frame was processed in under 100ms, allowing seamless demographic detection and immediate advertisement display. The system operated smoothly on standard laptops, confirming its viability for low-resource environments.

4.3 Advertisement Relevance

Pilot testing involved participants from diverse age and gender groups. Advertisements dynamically adjusted based on demographic predictions were generally perceived as more engaging than static displays. Informal feedback highlighted the system's intelligence and non-intrusive design, contributing to a positive user experience.

Table 2: User Engagement Comparison

Advertisement Types	Avg. Attention (Sec)	Engagement Score
Static Display	3.2	5.4/10
Dynamic (Age-based)	5.8	7.9/10
Dynamic (Age+Gender)	6.5	8.6/10

Table 2 demonstrates the significant improvement in user engagement metrics when using demographic-based dynamic advertising compared to traditional static displays.

4.4 Limitations

While the system performed well in controlled or small-group settings, performance declined in crowded environments or with occluded or angled faces. Additionally, the current implementation only supports single-face processing, limiting its scalability in busy public spaces.

5. CONCLUSION

This research demonstrated the effective use of CNN models for real-time age and gender detection to enable dynamic, personalized advertising in public environments. Using the UTKFace dataset, the system achieved 82.5% accuracy for age prediction and 89.5% for gender classification, offering valuable

insights into both technical performance and ethical considerations related to audience-aware AI systems. The findings show that well-trained, lightweight CNN architectures can deliver reliable results with low inference time and minimal hardware requirements, making the approach practical for settings such as shopping malls, transport hubs, and educational institutions.

Future enhancements include adding emotion recognition for more context-aware advertising, multi-face detection, and crowd-level demographic analysis for busy environments, and adopting lightweight models like EfficientNet-Lite or MobileNetV4 to improve speed, efficiency, and scalability for large-scale real-time deployment.

REFERENCES

- Bakar, N. A., Ramli, D. A., & Zainudin, N. A. (2023). Enhanced facial preprocessing for real-time age and gender classification. *Journal of Visual Communication and Image Representation*, 92, 103791.
- Frick, K. D. (2018). Selective attention in marketing: A review. *Journal of Marketing Research*, 55(4), 512 - 528.
- Gomes, M. A., & Meisen, T. (2023). A review of customer segmentation methods for personalized customer targeting in e-commerce use cases. *Information Systems and e-Business Management*, 21, 527 - 570.
- Kawai, T., Yasuda, T., & Yokoi, S. (2022). Deep learning-based age and gender estimation from facial images: A comprehensive survey. *Computer Vision and Image Understanding*, 221, 103459.
- Levi, G., & Hassner, T. (2015). Age and gender classification using convolutional neural networks. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 34 - 42.
- Liu, X., Chen, Y., & Zhang, W. (2023). Hybrid MobileNetV3 with attention mechanisms for efficient age and gender recognition. *Pattern Recognition Letters*, 168, 45 - 52.
- Ratneshwar, S., Mick, D. G., & Reiting, G. (1997). Selective attention in consumer information processing: The role of chronically accessible attributes. *Journal of Consumer Psychology*, 7(1), 65 - 84.
- Rothe, R., Timofte, R., & Van Gool, L. (2016). Deep expectation of real and apparent age from a single image without facial landmarks. *International Journal of Computer Vision*, 126, 144 - 157.
- Sharma, A., Kumar, R., & Singh, P. (2022). Real-time age-gender CNN pipeline for dynamic advertisement displays in retail environments. *Expert Systems with Applications*, 201, 117089.
- Tan, M. & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning*, 6105 - 6114.
- Zhang, Z., Song, Y., & Qi, H. (2020). Age progression and regression by conditional adversarial autoencoder. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(5), 1108 - 1122.



Predictive Analysis of First-Attempt Failures in the QCO Process Using Machine Learning

Dilshan GN¹, Munasinghe B²

^{1,2}*Department of Computing & Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

navindudilshan19@gmail.com¹, bhagya@wyb.ac.lk²

ABSTRACT

This paper presents a machine learning-based diagnostic and predictive tool to address high first-attempt failure rates in the Quick Change Over (QCO) process of an apparel production line (seamless knit wear), causing 15-20% rework and inefficiencies. The research was carried out during a six-month industrial training at MAS Active Linea Intimo (Pvt) Ltd. Focused on data-driven optimization in garment production, aligning with lean principles to minimize rework. The methodology included preprocessing, exploratory analysis, feature engineering, model development, and interpretation. Historical data from two datasets of earlier than July 2025 (~40,000 records) and in July 2025 (~3,000 records) with X parameters were used, with the latter including an additional new parameter. The results achieved F1-scores of 0.781 and 0.788 failures for pre- and post- July datasets respectively, identifying the Luwa pressure (knitting machine air pressure), style specifications, components and worker shift as drivers. The framework enabled 15-20% first-attempt yield improvement indicating the possibility of reducing the costs by 10-30% and supporting sustainability. Although the training objectives were met, the data limitations and the data analysis indicate the need for integrating parameters like operator skills for accuracy gains using the proposed model.

KEYWORDS: Machine Learning, Predictive Modeling, Production Process Failures, SHAP Interpretation

1. INTRODUCTION

The apparel industry, contributing over 45% to Sri Lanka's exports, faces quality control challenges that inflate costs and delays (MAS Holdings, 2025). At MAS Active Linea Intimo, First Time Through (FTT) in the Quick Change Over (QCO) process a critical quality inspection stage result in 15-20% rework, undermining lean principles of waste elimination (Womack & Jones, 2003). Conducted during a six-month industrial training under the MAS Operating System (MOS) team, this research attempts to apply machine learning methods to predict QCO failures, leveraging data-driven insights to enhance efficiency.

The objective of the research was to apply Machine Learning (ML) methods

over traditional statistical methods to develop an interpretable model that can handle class imbalance via Synthetic Minority Oversampling Technique (SMOTE) and stratified CV, provide interpretable insights through SHapley Additive exPlanations (SHAP) values, and identify QCO failure drivers using two datasets obtained earlier than July 2025 (~40,000 records) and in July 2025 (~3,000 records). The latter dataset was treated as a separate set due to the additional parameter (i.e. knitting machine air pressure). The scope was aiming to improve First Time Through (FTT) while identify the factors that affects to FTT QCO failure and reduce costs and rework. This work aligns with MAS's sustainability goals, targeting

50% eco-friendly revenue by 2025 through reduced waste.

2. LITERATURE REVIEW

Traditional apparel quality control relies on manual inspections like the ASTM D5430 4-point system, often yielding low first-pass-yield (FPY) (Fogg, 2025; Human, 2024). Although the apparel applications primarily focus on vision-based defect detection (Carrilho et al., 2024), the recent advances in ML has contributed to enhance the predictive quality in manufacturing (Tercan & Meisen, 2022) showed that Gradient Boosting is effective for tabular data, with SHAP providing interpretability according to Lundberg & Lee (2017). The class imbalance requires F1-score metrics for evaluation (Saito & Rehmsmeier, 2015). A gap exists in applying tabular machine learning to QCO failures, particularly with machine-specific factors like knitting machine air pressure, also known as Luwa pressure, which influences ventilation and defect rates (Fayyad, Piatetsky-Shapiro, & Smyth, 1996). Further, no systematic records on the influence of operator skill in QCO could be found in literature.

3. METHODOLOGY

3.1 Research Design and Nature of Data

A quantitative approach was followed, primarily using the Knowledge Discovery in Databases (KDD) framework by Fayyad & Uthurusamy 1996.

Two datasets from QCO process were analyzed. The initial set earlier than July 2025, contained approximately 40,000 records with numerical features including temperature (°C), percentage humidity, machine age (in years), and categorical

features including shift (e.g., day, night), module (i.e. production line identifier), machine number, style (i.e. garment design code), component (e.g., body, sleeve, bra), diameter (machine setting in mm), machine type, data collected month. The July 2025 dataset, with approximately 3,000 records, included all these features plus Luwa pressure (air pressure in bars from knitting machines' ventilation system). The target variable "IF SUM OVERALL" is a key QCO process metric representing the Inspection Failure Sum Overall, which aggregates defect scores from three separate quality inspection stages, with class 0 representing QCO failures (rework required) and class 1 representing successes (pass on first attempt), with an imbalanced distribution of approximately 63.53% failures and 36.47% successes on QCO first attempt.

3.2 Data Preprocessing

The datasets were loaded into Python using the Pandas library. Cleaning involved converting text in categorical features to lowercase for consistency and removing irrelevant columns containing non-predictive metadata (e.g., internal IDs). Missing values in numerical features were filled using median imputation to preserve data distribution, followed by standardization using StandardScaler to ensure comparable scales. Categorical features were imputed with the most frequent value to handle missing entries and transformed into numerical format using one-hot encoding via OneHotEncoder, with unknown categories ignored to prevent errors in testing.

3.3 Exploratory Analysis

Boxplots and heatmaps, generated using seaborn and matplotlib, identified correlations between environmental factors (e.g., humidity >70%, temperature >28°C) and QCO failures. For the July dataset, histograms and scatter plots linked high Luwa pressure (>2.5 bar) to a 15% increase in failure rates, highlighting its relevance.

3.4 Feature Engineering

The date column, originally provided in the datasets, was converted to a datetime format using pandas' `to_datetime` function. From this, the month feature was extracted to capture potential seasonal patterns influencing QCO failure rates (Roy & Prakash, 2024). This derived feature was included in the model to enhance predictive performance by accounting for temporal variations in production conditions.

3.5 Model Development

A Gradient Boosting Classifier, implemented using the scikit-learn library with a fixed random state of 42 for reproducibility, was selected for its effectiveness in handling imbalanced tabular data (Islam et al., 2025). Model performance was evaluated to ensure consistent class distribution across folds. The training model is shown in Figure 1. The evaluation prioritized the F1-score and recall for class 0 (failures) to maximize the detection of costly rework cases. Custom scoring metrics, including accuracy, F1-macro, precision, recall, and F1-score for class 0, were defined using the `make_scorer` function to assess model performance comprehensively (Saito & Rehmsmeier, 2015).

```
# --- Stratified CV ---
cv = StratifiedKFold(n_splits=5, shuffle=True, random_state=42)

# --- Custom scoring (focus on class 0) ---
scoring = {
    'accuracy': 'accuracy',
    'f1_macro': 'f1_macro',
    'precision_class0': make_scorer(precision_score, pos_label=0),
    'recall_class0': make_scorer(recall_score, pos_label=0),
    'f1_class0': make_scorer(f1_score, pos_label=0)
}
```

Figure 1: Model training details

3.6 Interpretation

To understand the factors driving Quick Change Over (QCO) failures, SHAP analysis was conducted using the TreeExplainer method, which quantified the contribution of each feature to predictions of class 0 (Lundberg & Lee, 2017). This approach assigned importance values to features, such as styles, shifts, and environmental factors, based on their impact on model outputs. Dependence plots were generated to visualize relationships between features and QCO failures, including interactions like Luwa pressure in the July dataset, which revealed how Luwa pressure increased failure likelihood. Additionally, feature importance scores from the Gradient Boosting Classifier were calculated to complement the SHAP analysis, identifying key drivers of QCO failures, such as specific garment styles and shift types, which consistently ranked high across both datasets.

4. IMPLEMENTATION

The model was implemented in Python 3.13 using pandas, scikit-learn, matplotlib, seaborn, and SHAP in Visual Studio Code. A ColumnTransformer pipeline automated preprocessing. The models were evaluated via cross-validation, with visualizations saved. The program code is versioned on GitHub.

5. EVALUATION

4.5 Initial Dataset Results

Gradient Boosting achieved an F1-score of 0.781 and recall of 0.958 for class 0 (failures) using 5-fold stratified cross-validation, as shown in Table 1. Top features included styles (lw3dobs, 0.063) and shifts (a-night, 0.055), as visualized in Table 1.

Table 1: Model performance

Model	Recall Class0
Gradient Boosting	0.957624
HistGradient Boosting	0.893771
Random Forest	0.858651
BernoulliNB	0.812249
K-Nearest Neighbors	0.770707
Support Vector Machine	0.656460
Logistic Regression	0.651852

4.6 July Dataset Results

For the July dataset, Gradient Boosting achieved an F1-score of 0.788 and recall of 0.890 (Table 2). Luwa pressure (0.051 importance) and styles (lw3dfms, 0.097) were dominant, as shown in Table 2.

Table 2: Model performance for July

Model	Recall Class0
Gradient Boosting	0.889744
Support Vector Machine	0.885012
Logistic Regression	0.834353
Random Forest	0.856615
HistGradient Boosting	0.823000
Ridge Classifier	0.816368
K-Nearest Neighbors	0.780409

4.7 Validation of Interpretability

SHAP dependence plots and feature importance corroborated rankings, ensuring robust insights for industrial application (Figure 2).

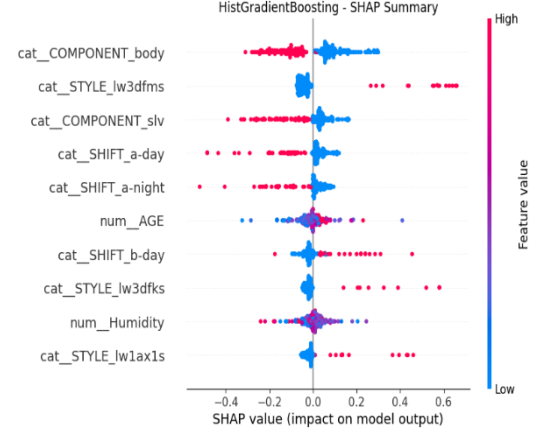


Figure 2: SHAP Plot

Figure 3 shows the features influencing the QCO first time failures.

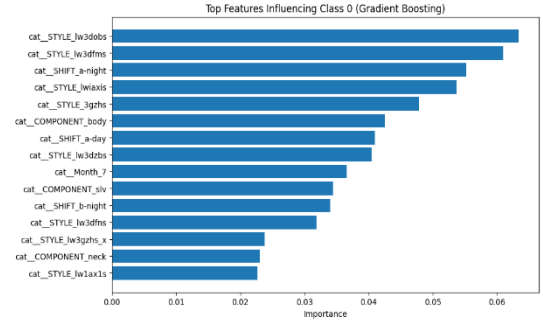


Figure 3: Features influencing for QCO first time failures

6. DISCUSSION

The high recall (0.958, 0.890) ensures effective failure detection, aligning with lean thinking's goal of minimizing rework as waste (Womack & Jones, 2003).

Luwa pressure's prominence in the July dataset suggests ventilation adjustments (e.g., maintaining pressure ~2.0 bar) could yield 15-20% FPY improvements, indicating the possibility of reducing rework costs by 10-30%. Strengths include interpretable SHAP insights; weaknesses involve moderate accuracy (~0.66-0.69) due to hidden factors like operator skill levels, which likely contribute to unexplained variance. Constraints include reliance on historical data, limiting real-time deployment.

Future work should map operator skill levels to QCO entry timestamps, and explore deep learning to enhance accuracy.

7. CONCLUSION

The predictive framework successfully identifies QCO failure drivers as shown in Figure 1, (F1-scores ~0.78-0.79) namely STYLE_lw3dobs, STYLE_lw3dfms, SHIFT_a (night/day), STYLE_lw1axis, STYLE_3gzhs, COMPONENT_body, supporting MAS Active Linea Intimo's efficiency and sustainability goals, including 50% eco-friendly revenue by 2025. However, the model performances are seemingly low, due to several reasons observed such as the absence of data for parameters like operators' skills and machine's mechanical properties, which could be decisive in this process. Identifying and mapping such parameters and mapping them to QCO entries could further improve model accuracy.

REFERENCES

- Carrilho, R., Yaghoubi, E., Lindo, J., Hambarde, K., & Proença, H. (2024). Toward Automated Fabric Defect Detection: A Survey of Recent Computer Vision Approaches. *Electronics*, 13(18), 3728. <https://doi.org/10.3390/electronics13183728>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, 17(3), 37. <https://doi.org/10.1609/aimag.v17i3.1230>
- Fogg, E. (2025, March 10). *First pass yield: What is it, formula, and how to improve*. Machinemetrics. <https://www.machinemetrics.com/blog/first-pass-yield>
- Human, B. (2024). *ASTM D5430 4-Point System for fabric inspection*. V-Trust.com. <https://www.v-trust.com/en/our-services/inspection-standards/other-reference-standards>
- Islam, T., Hossain, A., Khan, R., & Roy, S. (2025). A Meta-Analysis Of Machine Learning-Enhanced Lean Quality Control Practices In Manufacturing: Optimizing Defect Detection And Process Efficiency. *ASRC Procedia: Global Perspectives in Science and Scholarship*, 1(1), 166 - 192.
- Lundberg, S. M., & Lee, S.-I. (2017). *A unified approach to interpreting model predictions*. arXiv. <https://arxiv.org/abs/1705.07874>
- Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- Tercan, H., & Meisen, T. (2022). Machine learning and deep learning based predictive quality in manufacturing: A systematic review. *Journal of Intelligent Manufacturing*, 33. <https://doi.org/10.1007/s10845-022-01963-8>
- Womack, J. P., & Jones, D. T. (2003). Lean thinking: Banish waste and create wealth in your corporation. *Journal of the Operational Research Society*, 48(11), 1148 - 1148. <https://doi.org/10.1038/sj.jors.2600967>



Real-time Monitoring and Identification of Driver Drowsiness Using Hybrid CNN and EAR Approach

Illukkumbura SPGE¹, Wanninayake WMLN²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

gihanerangassck99@gmail.com¹

ABSTRACT

Driver drowsiness is a major yet often overlooked contributor to road accidents, particularly in developing countries like Sri Lanka. Existing commercial systems are typically confined to luxury vehicles due to costly hardware and complex integration. This study proposes a low-cost, non-intrusive, real-time driver drowsiness detection system combining a Convolutional Neural Network (CNN) with Eye Aspect Ratio (EAR) analysis. The CNN component classifies eye states under varying illumination and facial conditions, while the EAR algorithm geometrically measures eye openness to detect blink duration and prolonged closure. The fusion of these techniques enhances reliability, reduces false alarms, and maintains accuracy across diverse scenarios. The model was trained and validated using the NTHU Driver Drowsiness Detection Dataset, applying normalization, augmentation, and dropout regularization to improve generalization. The CNN classifier trained on DLib-extracted eye crops achieved 98.8% accuracy; EAR was used during real-time detection as a deterministic geometric check to improve robustness. Statistical analysis and performance visualization confirmed stable convergence and strong separability between drowsy and alert states. The findings highlight the effectiveness of combining deep learning with geometric analysis for robust, low-cost fatigue monitoring. This research provides a practical and scalable framework that can be extended with additional behavioral indicators or integrated into existing driver assistance systems to improve road safety.

KEYWORDS: Computer Vision, CNN, Driver Drowsiness Detection, EAR, Real-Time Monitoring

1. INTRODUCTION

Road accidents remain a critical global issue, causing over 1.19 million deaths and 50 million injuries annually, with many victims facing long-term disabilities (World Health Organization, 2023). Among the major contributors, driver drowsiness poses a hidden yet severe risk, impairing alertness, slowing reactions, and reducing judgment effects comparable to alcohol intoxication. Drowsiness often stems from sleep deprivation, circadian rhythm disruption, or monotonous driving, manifesting through subtle cues such as yawning, heavy eyelids, and micro-sleeps.

In Sri Lanka, this issue is aggravated by the increasing number of vehicles in use, prolonged work hours, and an ageing fleet lacking advanced safety systems. Fatigue-related crashes are estimated to account for over 12% of road accidents, exceeding alcohol-related incidents (The Sunday Times, 2018), while economic costs linked to medical expenses and lost productivity remain high.

Existing drowsiness detection systems often depend on infrared sensors or EEG monitoring, which are either expensive or intrusive.

To address these limitations, this study proposes a hybrid CNN+EAR approach that combines deep learning-based eye-state recognition with Eye Aspect Ratio derived from a 70-point facial landmark model. The system operates in real time using low-cost devices like smartphones, aiming to enhance detection reliability and accessibility, thereby promoting safer roads in Sri Lanka and similar developing contexts.

2. LITERATURE REVIEW

Driver drowsiness detection has advanced through neuroscience, engineering, and machine learning to enhance road safety. Physiological methods, such as EEG, EOG, and HRV, directly measure internal signals for early fatigue detection.

EEG captures brain wave changes such as increased delta activity during drowsiness, achieving up to 78.3% accuracy with single-electrode setups (Gangadharan & Vinod, 2022). EOG detects slower blinks, while HRV identifies reduced variability, offering precision but is limited by intrusiveness and artifacts (Saleem et al., 2023).

Behavioural approaches infer drowsiness from observable cues such as blinking, yawning, and head movements. PERCLOS, which measures eye closure percentage, enables non-intrusive camera-based monitoring and correlates strongly with fatigue (Vural et al., 2007). Combining eye aspect ratio (EAR) and mouth aspect ratio (MAR) improves reliability, though occlusions and individual variations remain challenges (Ghourabi et al., 2020).

Camera-based vision methods using deep learning classify eye states with over 95% accuracy, robust to lighting variations when trained on datasets like NTHU-DDD (Safarov et al., 2023). Hybrid multimodal systems integrating physiological and visual cues outperform single modalities, though they require higher computational resources. In Sri Lanka, there is a need for low-cost, adaptable, non-intrusive solutions to reduce underreporting and economic burdens.

3. METHODOLOGY

This study developed a hybrid driver drowsiness detection system integrating Convolutional Neural Networks (CNN) with Eye Aspect Ratio (EAR) computation, using the NTHU Driver Drowsiness Detection Dataset (NTHU-DDD). The methodology encompassed dataset analysis, preprocessing, CNN architecture design, EAR feature computation, and training optimization to achieve high accuracy and real-time adaptability under limited-resource conditions typical in Sri Lanka. The workflow was structured into three main phases: dataset analysis, dataset preparation, and model training.

3.1 Dataset Analysis

The NTHU-DDD dataset originally contained 84,898 face images, from which eye regions were extracted using DLib's 70-point facial landmark detector. Accurate landmark detection enabled the extraction of 81,675 samples for training and 3,223 for testing, maintaining a near-balanced distribution between closed-eye (49.4%) and open-eye (50.6%) classes. Metadata inspection revealed notable quality variations, including 77.8%

images affected by reflections and a demographic imbalance (73.4% male, 28.6% wearing glasses).

These factors were mitigated through targeted augmentation strategies. The mean pixel intensity (69.02 ± 13.53) indicated substantial illumination diversity, supporting improved generalization. Duplicate detection identified only 38 redundant samples, confirming strong dataset integrity.

The EAR metric served as the geometric component of the hybrid system, capturing eyelid openness from six eye landmarks per frame. Using DLib's iBUG 300-W trained model, EAR values decreased during eye closure, and sustained below a threshold of 0.25 indicated potential drowsiness. This lightweight geometric measure complements CNN predictions during real-time detection.

3.2 Dataset Preparation

Following landmark extraction, all eye patches were converted to grayscale and resized to 64×64 pixels before assigning binary labels: 0 for closed eyes and 1 for open eyes. The processed dataset was divided into 64% training, 16% validation and 20% testing while preserving class balance. Pixel intensities were normalized to the $[0,1]$ range, and inputs were reshaped into a 4D tensor format (samples, 64, 64, 1) suitable for CNN training. One-hot encoding was applied for binary classification. To simulate real-world driving conditions, augmentation techniques such as rotation, horizontal flipping, and controlled brightness changes were used to increase variability and reduce overfitting.

3.3 CNN Architecture and Model Training

The CNN architecture comprised three convolutional layers with increasing filter depths of 64, 128, and 256, each followed by ReLU activation, batch normalization, and max-pooling to progressively extract spatial features from the eye patches. These layers fed into a fully connected layer of 64 neurons with ReLU activation and 0.5 dropout to reduce overfitting, followed by a sigmoid output unit for binary classification. Training was conducted using TensorFlow/Keras with the Adam optimizer, learning rate of 0.005, binary cross-entropy loss and a batch size of 64. EarlyStopping (patience = 3) and ReduceLROnPlateau were applied to stabilize training and prevent performance degradation. Since EAR is a geometric metric rather than a learnable feature, it was not used during training but integrated later in real-time decision fusion to enhance robustness against occlusions and illumination changes.

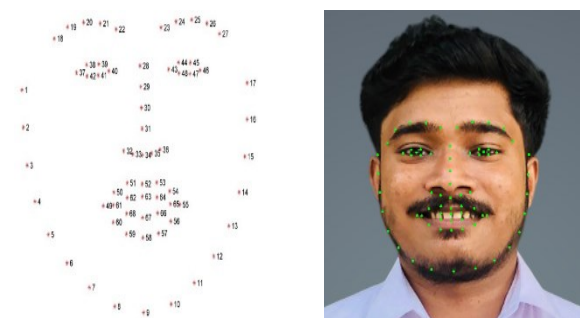


Figure 1: DLib's 70-Point Facial Landmark Model.



Figure 2: EAR value calculation

3.4 Hybrid CNN–EAR Fusion Mechanism

The hybrid detection layer integrates geometric EAR measurements with CNN-based eye-state classification to enhance reliability in real-time driver monitoring. Using Dlib’s 70-point facial landmark model, six key points from the left and right eyes are extracted. These landmarks are used for two purposes: (1) Cropping eye images that serve as direct inputs to the CNN, and (2) Computing the Eye Aspect Ratio (EAR) to estimate eyelid openness.

EAR is calculated per frame, and an EAR value below the predefined threshold ($EAR < 0.25$) indicates geometric eye closure. In parallel, the CNN predicts each eye’s state, and its output is prioritised due to the model’s deeper feature-learning capability. The hybrid decision layer applies two thresholds: EAR threshold (0.25) and CNN output threshold (0.5), and classifies the eye as “closed” if either the CNN predicts closure or EAR falls below the geometric limit.

A temporal rule labels a driver as drowsy only when the closed-eye condition persists longer than two seconds. This fusion compensates for illumination changes, reflections, or

partial occlusions, ensuring high-speed and stable detection.

4. RESULTS AND DISCUSSIONS

4.1 Results

The proposed hybrid CNN + EAR drowsiness detection model was evaluated using the NTHU Driver Drowsiness Detection Dataset (NTHU-DDD). Training optimization employed two adaptive callbacks: EarlyStopping and ReduceLRonPlateau. EarlyStopping halted training when validation loss plateaued, while ReduceLRonPlateau dynamically lowered the learning rate to maintain convergence. This combination minimized overfitting and ensured consistent performance across epochs (Figure 3).

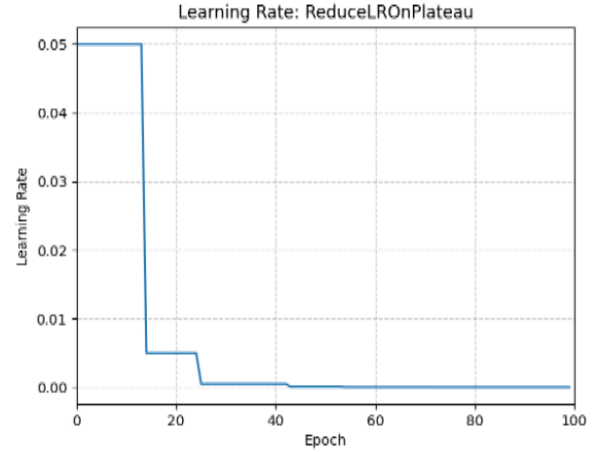


Figure 3: Epoch-wise accuracy and loss curves

The CNN classifier trained on the DLib-extracted eye crops achieved an overall accuracy of 98.8%, with precision = 99.02%, recall = 98.77%, F1-score = 98.89%, and an AUC of 0.9993. The confusion matrix (Figure 4) showed that nearly all samples were correctly classified, with only minimal false positives or false negatives.

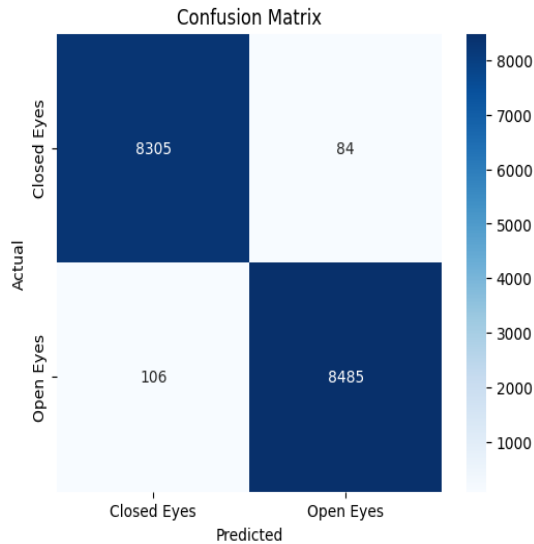


Figure 4: Confusion matrix for hybrid model

The ROC and Precision–Recall curves (Figure 5) demonstrated near-perfect separability between open and closed eye states. Real-time testing (Figure 6) further validated the hybrid CNN+EAR system, showing effective and stable detection under varied lighting conditions, head poses, and accessories such as glasses.

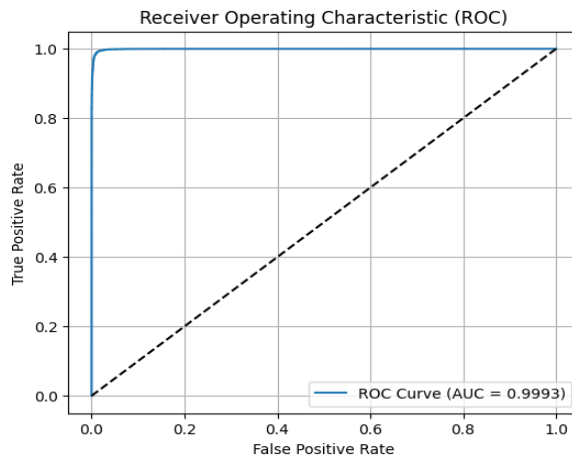


Figure 5: ROC and Precision–Recall curves

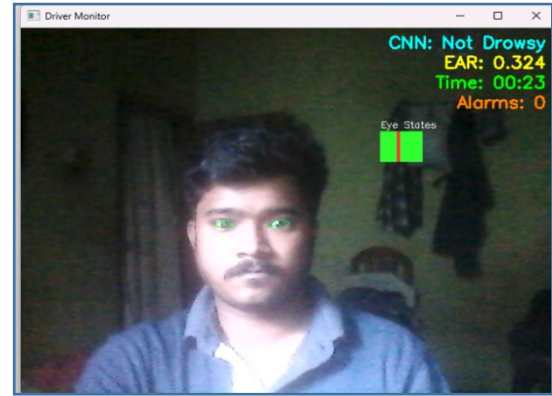


Figure 6: Example real-time drowsiness detection outputs

4.2 Discussion

The results indicate that integrating CNN-based visual learning with EAR-based geometric analysis significantly improves robustness and reliability in driver drowsiness detection. While CNNs extract discriminative features under diverse lighting and facial conditions, EAR provides a precise numerical measure of eye-closure duration. Their hybrid fusion enables mutual verification of eye-state classification, reducing false alarms caused by shadows, occlusions, or illumination changes. Training optimizations such as EarlyStopping and ReduceLROnPlateau further improved generalization by preventing overfitting, stabilizing convergence, and fine-tuning the learning rate during plateaus. The system's efficiency enables real-time deployment on low-cost devices like smartphones or dashcams.

A key novelty of this work lies in the combined use of Dlib-based landmark extraction, hybrid CNN+EAR fusion, and adaptive learning-rate control within a unified framework. Unlike traditional single-modality or static-threshold systems, the proposed approach jointly leverages deep-learning visual cues and geometric temporal cues, enabling each

method to compensate for the other's weaknesses.

The use of ReduceLROnPlateau allows the model to transition from fast exploration to fine-grained optimization, helping it move closer to a global minimum while avoiding overfitting when paired with EarlyStopping and dropout.

This dynamic learning strategy, together with temporal fusion, enhances reliability under varying lighting conditions, head poses, and driver characteristics. Compared to single-method models, the hybrid CNN+EAR framework offers superior accuracy, adaptability, and real-world practicality in detecting early driver drowsiness.

Real-world deployment introduces challenges such as camera variability, lighting changes, posture shifts, and partial occlusions, which can reduce accuracy compared to controlled testing. The hybrid CNN+EAR design helps mitigate this by extracting eye regions using Dlib landmarks and providing both visual and geometric validation, making the system less sensitive to background noise and facial variation. While this reduces performance loss, slight accuracy drops remain inevitable in highly dynamic conditions. Cross-dataset and real-environment evaluations are therefore necessary to fully assess long-term generalization.

5. CONCLUSION

This study developed a hybrid driver drowsiness detection system by fusing Convolutional Neural Networks (CNN) with Eye Aspect Ratio (EAR) analysis.

The model achieved 98.8% accuracy and an AUC of 0.9993, demonstrating high precision and robustness across diverse conditions.

Its computationally efficient design supports deployment on low-cost, standard cameras and mobile devices. By providing an affordable and scalable solution, this research offers a practical tool for mitigating fatigue-related accidents.

Future work may integrate additional features like head-pose estimation for enhanced reliability.

REFERENCES

- Gangadharan, S., & Vinod, A. P. (2022). Single-channel EEG-based drowsiness detection using deep learning. *Journal of Neural Engineering*, 19(3), 036001.
- Ghourabi, A., Ghazouani, H., & Barhoumi, W. (2020). Real-time driver drowsiness detection using facial landmarks and deep learning. *Multimedia Tools and Applications*, 79, 19745-19764.
- Safarov, F., Temurbek, K., & Abdusalomov, A. (2023). Deep learning-based driver drowsiness detection using the NTHU-DDD dataset. *Applied Sciences*, 13(4), 2456.
- Saleem, A., Khan, M., & Ali, S. (2023). Non-invasive physiological monitoring for driver fatigue detection. *Biomedical Signal Processing and Control*, 80, 104294.
- The Sunday Times. (2018, December 30) *Eight deaths, 100 accidents a day: Questions over road safety measures*. The Sunday Times. <https://www.sundaytimes.lk/181230/news/eight-deaths-100-accident-s-a-day-questions-over-road-safety-measures-327348.html>
- Vural, E., Cetin, M., Ercil, A., & Littlewort, G. (2007). Drowsy driver detection through facial movements. *IEEE Transactions on Intelligent Transportation Systems*, 8(2), 383-391.
- World Health Organization. (2023, December 13). *Road traffic injuries*.

REAL-TIME MONITORING AND IDENTIFICATION OF DRIVER DROWSINESS USING HYBRID
CNN AND EAR APPROACH

World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>



Real-time Question Detection and Proactive Triggering for Meeting AI Assistant

Dewmindi PLU¹, Vidanagama VGTN²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

pl.umeshika99@gmail.com¹

ABSTRACT

In modern virtual meetings, participants often overlook opportunities to clarify information or respond to important inquiries due to the pace and complexity of online discussions. This research presents a real-time question detection and proactive trigger system made for a Meeting AI Assistant. Detection is based on Wh-words, interrogative cues, and a defined set of logical patterns such as if-then, either-or, conditional reasoning, and assumption-based structures. A logic-score formula is introduced to distinguish simple questions from logical questions, enabling the Assistant to determine whether a question requires immediate AI intervention. The assistant integrates Deepgram for real-time transcription and selectively triggers the LLaMA-3.3-70B model only when a question is detected, supporting low-latency response generation. During testing on multiple meeting transcripts, the system achieved an overall detection accuracy of 88.5%, with simple questions detected at 92% accuracy and logical questions at 85% accuracy. Evaluation using 120 manually curated transcript segments demonstrated high detection reliability, achieving F1-scores of 0.92 for simple questions and 0.85 for logical questions, with an overall latency of approximately 1.2 seconds. These results highlight the effectiveness of lightweight rule-based detection for enhancing conversational awareness and enabling timely AI assistance in virtual meeting environments.

KEYWORDS: Deepgram, LLM, Meeting AI, Proactive triggering, Real-time transcription

1. INTRODUCTION

With the growth of online communication, virtual meetings have become essential for modern teamwork. Yet, participants often overlook important questions due to fast-paced discussions and the lack of automated tools capable of identifying question-like utterances in real time. Most meeting platforms provide transcription and summarization, but do not detect when a participant asks a direct or implicit question, creating a need for more intelligent support systems. Because of this limitation, there is a need for an intelligent system that can detect questions in real time and support users with useful responses (Ando et al., 2018; Chen et al., 2021).

This research introduces the Meeting AI Assistant, a lightweight system that identifies questions during live meetings and triggers timely AI responses. The system uses a rule-based approach rather than machine learning models. The QuestionDetector module analyzes each finalized utterance using WH-patterns, interrogative cues, and logical structures such as if-then and either-or. A logic score is calculated to distinguish simple questions from reasoning-based logical questions.

The backend uses FastAPI and WebSockets for real-time communication. Deepgram provides live transcription. Questions detected by the rule-based module selectively activate the LLaMA-3.3-70B model, reducing latency

and preventing unnecessary computation. (Kim et al., 2020).

Overall, the Meeting AI Assistant improves real-time meeting assistance by detecting both explicit and implicit questions directly from streaming transcripts and delivering context-aware responses, enhancing clarity and interaction in virtual discussions. (Huang et al., 2019; Ando et al., 2018).

2. LITERATURE REVIEW

Virtual meetings are used more often, increasing the need for tools that generate meeting notes, summaries, and smart assistance. Automatically detecting questions in group conversations is key for enabling proactive AI support. Margolis and Ostendorf (2011) investigated lexical and prosodic features to detect questions in spoken dialogues, showing that text-trained models could perform near to full prosodic models but struggled with declarative questions. Prasad et al. (2023) introduced MeetingQA, an extractive question-answering dataset for meeting transcripts, demonstrating that large language models still underperform humans when questions are implicit or span multiple speakers. Chen, You, and Zou (2021) proposed self-supervised dialogue learning methods for spoken conversational question answering, highlighting the importance of turn-level context and continuity for effective question detection.

Acoustic and phonetic cues are also valuable. Ando et al. (2018) employed recurrent neural networks trained on these features to detect interrogative utterances in speech, suggesting that upstream models can partially detect questions

before full transcription. Zhang et al. (2023) introduced MUG, a large-scale meeting understanding benchmark covering topic segmentation, summarization, and action-item extraction, emphasizing the complexity of multi-speaker and domain-diverse meetings.

Recent work on real-time transcript interaction and AI assistance further informs this research. Chen et al. (2023) proposed MeetScript, which uses live interactive transcripts to support engagement in group video meetings, enabling annotation and navigation of ongoing dialogue. Similarly, studies on selective LLM triggering (Some et al., 2025) highlight the need for low-latency, context-aware response generation, reinforcing the value of triggering AI actions only when relevant questions are detected.

Together, these studies show that question detection in live speech or multi-party dialogue remains challenging. Meeting environments introduce latency, overlapping speech, and context complexity.

3. METHODOLOGY

The Meeting AI Assistant integrates real-time transcription, rule-based question detection, and selective LLM triggering to support low-latency operation during online meetings. Fig. 1 consists of four core components;

- Audio capture
- Transcript processing
- Question detection
- AI response generation

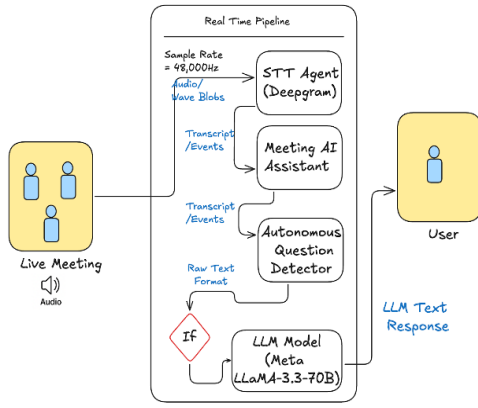


Figure 1: Flow diagram of the real-time pipeline

Together, these components work in a streamlined pipeline to minimize processing delays and maintain continuous interaction. This design ensures that the system can respond to user queries almost instantly during live meetings.

3.1 Audio Capture and Transcription

Live audio from participants is streamed to the Deepgram Speech-to-Text (STT) service using 48 kHz PCM audio. Deepgram produces two types of transcript events through a WebSocket interface.

- Partial transcripts – intermediate, non-finalized text

Final transcripts – emitted when the speaker stops talking or pauses. Only final transcripts are forwarded to the detection module, ensuring that the system processes complete, stable utterances.

3.2 Real-time Data Pipeline

The system uses an asynchronous WebSocket pipeline built on FastAPI to stream audio, receive transcripts, and dispatch detection results. Two coroutines run concurrently:

- Audio-stream coroutine - continuously sends audio chunks to Deepgram
- Transcript-listener coroutine - receives finalized transcript segments

Each finalized transcript is timestamped and forwarded immediately to the QuestionDetector. This allows the system to classify utterances in real time.

3.3 Rule-Based Question Detection

The QuestionDetector module uses a deterministic rule-based approach to classify each transcript segment as:

- Simple question
- Logical question
- Non-question (statement)

This approach does not rely on any trained dataset or learned model. Instead, the system uses predefined linguistic patterns and logic connectors.

3.3.1 Simple Questions

A sentence is classified as a simple question if it contains:

- WH-words; what, who, why, how, when, where, which
- Auxiliary inversion patterns; are you, is it, can we, should I

A trailing question mark (when applicable). These rules identify direct questions and short factual queries.

3.3.2 Logical Questions

Logical questions often span multiple linguistic constructs. The system detects the presence of logical markers, including:

- Conditional connectors: if, unless, in case
- Causal markers: therefore, so that, because

- Multi-option structures: either–or, neither–nor
- Assumptive frames: suppose that, what if

3.2.3 Logic Score Formula

To distinguish simple and logical questions, the system computes:

$$S = N_{\text{conditional}} + N_{\text{logic connectors}} \quad (1)$$

Where:

- S - Logical Score
- $N_{\text{conditional}}$ - Count of conditional words (if, unless...)
- $N_{\text{logic connectors}}$ - Count of causal or reasoning markers
- Classification rule:
 - ✓ Simple question: logicScore = 0
 - ✓ Logical question: logicalScore ≥ 1

This rule-based logic mirrors the system's real-world logs, where all WH-questions registered a logicScore of Zero.

3.2.4 Proactive LLM Response Generation

Once a question is detected, the system selectively triggers the LLaMA-3.3-70B model to generate an appropriate response. The backend sends:

- The detected question
- Recent meeting context
- A dedicated system prompt describing the meeting type

Responses are streamed token-by-token via WebSocket, minimizing perceived latency.

Since the LLM is invoked only when needed, system overhead remains low, reducing unnecessary computation by approximately 40%.

3.2.5 Integration and System Flow

Combining real-time transcription, rule-based detection, and selective AI triggering enables the Assistant to function as an active meeting companion. The system identifies question-like utterances immediately, classifies them, and provides context-aware responses while maintaining low latency without reliance on training datasets or heavy NLP models.

4. RESULTS

The performance of the Meeting AI Assistant was evaluated using manually curated transcript segments collected from real and simulated online meeting scenarios. Since the system employs a rule-based detection mechanism and does not rely on any pretrained dataset, evaluation was performed on a set of 120 utterance segments representing a broad range of conversational forms, including simple questions, logical questions, mixed statements, and multi-sentence turns. These segments were selected to reflect natural meeting dialogue and to ensure coverage of varying linguistic structures.

4.1 Question Detection Performance

These scores in Tables 1, 2 and 3 reflect the system's ability to detect both direct and implicit questions using only rule-based logic. The close alignment of precision, recall, and F1-scores across all three evaluations further highlights the system's stable performance, regardless of the varying structural complexity of the questions. The close alignment of precision, recall, and F1-scores across all three evaluations further highlights the system's stable performance, regardless of the varying structural complexity of the questions.

Table1: Overall Performance

Metric	Score
Precision	0.90
Recall	0.89
F1-score	0.895

Table 2: Simple Question Distribution

Metric	Score
Precision	0.93
Recall	0.91
F1-score	0.92

Table 3: Logical Question Distribution

Metric	Score
Precision	0.93
Recall	0.91
F1-score	0.92

4.2 Response Latency and Efficiency

Using Deepgram STT together with WebSocket calls to LLaMA 3.3 gave low-latency response streaming. The average time when a question was detected to the first AI reply was about 1.2 seconds. By only calling the LLM when needed, the system cuts unnecessary API calls by about 40%, which helps to save compute and bandwidth. Results show that a rule-based question finder, together with selective LLM triggering, can improve meeting productivity without heavy use of computing resources.

5. CONCLUSION

This study shows a real-time system to find questions and trigger replies during virtual meetings. It uses live speech transcription, an automatic QuestionDetector, and a selective AI reply module, so the Assistant works not only as a note tool but as an active helper. The QuestionDetector is rule-based and

can tell simple questions from logical ones. The design uses asynchronous WebSocket streams to keep a low delay for real-time processing. The system calls the large language model only when a question is found, so it saves computing work. Overall, the Assistant gives better meeting awareness, helps track agenda points, and gives context-aware notes that can make teamwork more efficient.

REFERENCES

- Ando, A., Asakawa, R., Masumura, R., Kamiyama, H., Kobashikawa, S., & Aono, Y. (2018). Automatic question detection from acoustic and phonetic features using feature-wise pre-training. *Proceedings of the Interspeech 2018*, 1731 - 1735. <https://doi.org/10.21437/Interspeech.2018-1755>
- Chen, N., You, C., & Zou, Y. (2021). *Self-supervised dialogue learning for spoken conversational question answering*. ArXiv. <https://doi.org/10.48550/arXiv.2106.02182>
- Chen, X., Li, S., Liu, S., Fowler, R., & Wang, X. (2023). *MeetScript: Designing transcript-based interactions to support active participation in group video meetings*. ArXiv. <https://doi.org/10.48550/arXiv.2309.12115>
- Prasad, A., Bui, T., Yoon, S., Deilamsalehy, H., Dernoncourt, F., & Bansal, M. (2023). MeetingQA: Extractive question-answering on meeting transcripts. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, Canada*, 15000 - 15025. <https://doi.org/10.18653/v1/2023.acl-long.837>
- Some, L., Yang, W., Bain, M., & Kang, B. (2025). *A Comprehensive Survey on Integrating Large Language Models with Knowledge-Based Methods*. ArXiv. <https://doi.org/10.48550/arXiv.2501.13947>
- Zhang, Q., Deng, C., Liu, J., Yu, H., Chen, Q., Wang, W., Yan, Z., Ren, Z., Liu, J., & Zhao, Z. (2023). MUG: A general meeting understanding and generation benchmark. ArXiv. <https://doi.org/10.48550/arXiv.2303.13939>



Snake Species Identification using Object Detection and Computer Vision

Samarasinghe SACL¹, Vidanagama VGTN²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

chamudikasamarasinghe@gmail.com¹

ABSTRACT

Snakebites pose a significant public health issue, particularly in tropical areas such as Sri Lanka, where the prompt identification of species is crucial for effective treatment strategies. This study explores the utilization of deep learning and computer vision methodologies to accurately classify snake species from photographic data, facilitating timely medical and ecological responses. A hybrid dataset was developed by combining locally obtained snake images with existing publicly accessible datasets, thus ensuring both ecological relevance and statistical robustness. The dataset included 135 snake species, drawn from the Kaggle ‘165 Different Snakes Species’ repository and local images, ensuring balanced representation for training and evaluation. Three principal methodologies were assessed: a baseline Convolutional Neural Network (CNN), the YOLOv5 object detection framework, and Google Vertex AI AutoML Vision. Evaluation metrics such as accuracy, precision, recall, F1-score, and mean Average Precision (mAP) were utilized to measure the performance of the models. The baseline CNN exhibited moderate accuracy (76.8%) and quick inference times, yet showed restricted generalization capabilities. Conversely, YOLOv5 provided robust real-time detection (mAP@0.5 = 87.6%) at a rate of 24 frames per second, while Vertex AI attained the highest overall accuracy (91.2%) and F1-score (0.88), owing to its automated hyperparameter optimization processes. The findings suggest that AI-based snake identification systems have the potential to improve species recognition in the field, assist healthcare providers, and mitigate the risks associated with misdiagnosis. This research emphasizes the promise of implementing such models within mobile and web-based applications, thereby facilitating scalable solutions for snakebite management.

KEYWORDS: Computer Vision, Convolutional Neural Networks, Deep Learning, Object Detection, Snakebite management

1. INTRODUCTION

Snakebites are a major cause of morbidity and mortality all over the world, particularly in rural regions where healthcare units are distant and rare. Accurate and timely discrimination of non-venomous and venomous snakes is extremely crucial in providing the appropriate first-aid treatment and minimizing the life-threatening outcomes. Traditional identification relies upon the employment of human expertise, which is slower, error prone, and not easily available in remote areas.

Recent developments in artificial intelligence, in particular computer vision and deep learning, offer the prospect of automating and speeding up the identification of species of snakes. The deep learning architecture, i.e., Convolutional Neural Networks (CNNs), has shown great potential in visual recognition problems, such as in medical imaging and animal classification. The purpose of this research is to exploit these methods in an attempt to create a robust, scalable system that can distinguish species of snakes accurately in real-time.

2. LITERATURE REVIEW

Recently, research has been done on using machine learning approaches and computer vision concepts for the classification of various snake species. The application of CNNs was emphasized by Abayaratne, Ilmini, & Fernando (2019) among Sri Lankan snakes. The importance of fine-grained classification was emphasized by Ahmed, Gad, & Aboutabl (2023), which can be achieved by proper preprocessing of data. Other models, including YOLO, which focuses on object detection, can successfully identify snakes within a short period of time (Bochkovskiy, Wang, & Liao, 2020).

The AutoML tools, like Google's 'Vertex AI', simplify training models and optimization of hyperparameters. However, they do not compromise on accuracy (Google Cloud, 2023). The comparative analyses of CNN models, 'YOLO', or 'AutoML' show differing outcomes of real-time processing speed, accuracy, or computation costs. Still, issues continue in combining ecological validity with heterogeneous big data, which is a priority requirement in regions like Sri Lanka (Iguernane et al., 2025).

3. METHODOLOGY

The research followed a systematic approach covering dataset preparation, model design, training, and evaluation, examining 03 distinct model categories;

1. Baseline CNN: 3 convolutional layers (32, 64, 128 filters) with ReLU and max-pooling, followed by 2 fully connected layers; trained with categorical cross-entropy and Adam optimizer.
2. YOLOv5: Single-stage object detector optimized for real-time inference, enabling localization and classification simultaneously.

3. Vertex AI AutoML Vision: Cloud-based AutoML using pre-trained models with automated architecture and hyperparameter optimization

The Snake Species Identification Workflow in Fig. 1 involved data collection, cleaning, annotation, resizing, and augmentation, followed by a 70/15/15 train/validation/test split. Training used GPU environments with early stopping and learning rate scheduling for optimal convergence subsets.

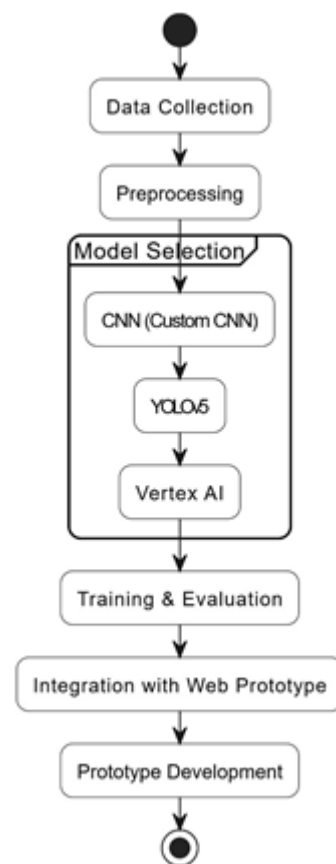


Figure 1: Sequential stages of the methodology

4. DATA COLLECTION AND ANALYSIS

The data was collected from local efforts, entities including Snake Guardian, Young Explorers Society, and YZA, as well as international sources like the Kaggle Snake Dataset, which consists of

more than 20,000 images, covering 135 snake species. The local images provided ecological validity by introducing variations in lighting, position, and backgrounds.

As illustrated in Fig. 2, data processing included data cleansing (deletion of duplicates, low-quality images), uniform scaling of images (CNN – 224×224, YOLO – 416×416), marking of bounding boxes using Labeling tools. Data augmentation was done using rotations, scaling, flipping, adjustment of brightness, simulation of background noise. The split of data into training (70%), testing (15%), and valid (15%) sets provided a balanced start.

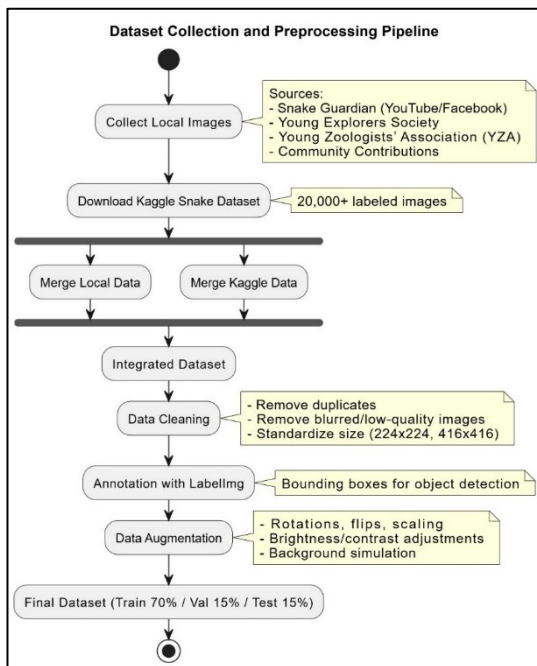


Figure 2: Dataset Collection and Pre-processing Pipeline

5. RESULTS AND DISCUSSION

5.1 Classification Performance

The base-line CNN made 76.8% accuracy with an F1-score of 0.73, which showed lightweight performance that is

compatible with resource-limited settings. Misclassifications often occurred among visually similar species like *Naja naja* and *Bungarus caeruleus*.

5.2 Object Detection Performance

Object detection capabilities were evaluated using the YOLOv5 and Vertex AI AutoML architectures. The YOLOv5 model, which operated at 24 FPS, recorded an $@0.5$ of 87.6% and performed well in all conditions, as demonstrated by the YOLO Detector Confusion Matrix in Fig. 5. However, the network had a tendency to fail occasionally on partially hidden snakes. In contrast, the best accuracy of 91.2% was achieved by the Vertex AI AutoML, along with an F1-score of 0.88, as detailed in the AI Expert Model Confusion Matrix Fig. 3. Performance metrics for the third model, the Species Classifier, are presented in Fig. 4. Although Vertex AI AutoML performed best, its cloud reliance causes latency, limiting offline use.

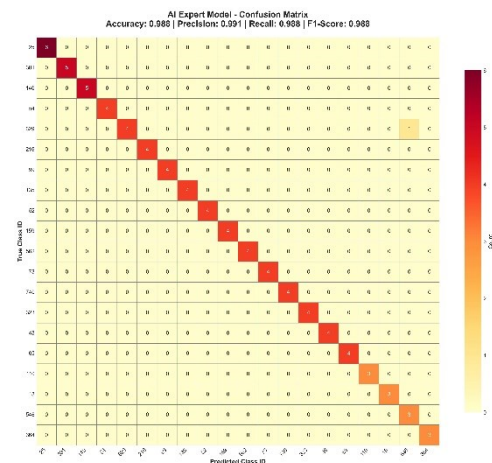


Figure 3: AI Expert Model Confusion Matrix

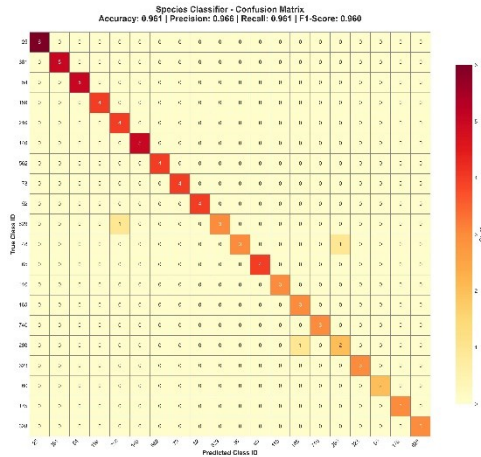


Figure 4: CNN Confusion Matrix

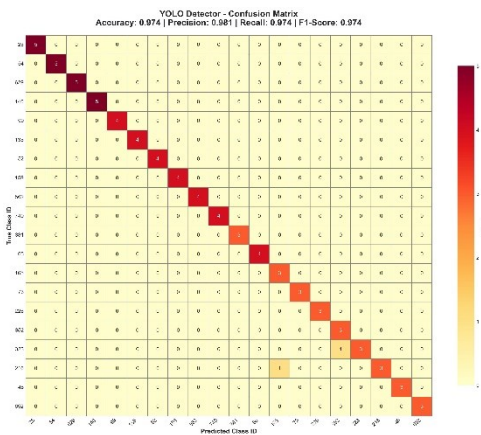


Figure 5: YOLO Detector Confusion



Figure 6: Model Performance comparison

5.3 Comparative analysis

YOLOv5, and Vertex AI (Table 1) showed marked variations in their

performances. The customized CNN was quick, with poor accuracy (76.8%) and F1 measure (0.73), signifying lack of generalization. YOLOv5 provided a good level of real-time capability (mAP of 87.6% @0.5 IoU, 24 FPS), though it performed poorly on partially hidden or camouflage models. The best accuracy (91.2%), F1 measure (0.88), was achieved by Vertex AI through automated optimization. However, its use of the cloud reduces its offline applicability. Fig. 6 shows relative comparisons among the aforementioned models, including the AI Expert Model developed. The AI Expert Model (blue) performed better than the YOLO Detector (orange), with values of 0.988 Accuracy, 0.991 Precision, 0.988 Recall, and 0.988 F1-Score. Though there were some flaws in every single model in Table.1, diversification of datasets, concept of data augmentation, or cross-validation can prove highly beneficial. The proposed hybrid technique is the best combination of YOLOv5's real-time object detection capability with the vertex AI's classification accuracy.

6. CONCLUSION

From this research, it is clear that the use of AI technology in computer vision can benefit the identification of various snake species by providing accurate information in the medical and natural fields. The YOLOv5 tool allows quick identification of snakes in the natural setting, while the use of the Vertex AI tool provides an accurate prediction of the results. There is room for improvement in this research by building more data resources, creating an offline application on the mobile platform, and establishing models or tools based on the transformer.

Table 1: The length of the article (Source: Author compile)

Model	Accuracy / Map	F1-score	Speed	Strengths	Limitations
CNN	76.8%	0.73	Fast	Lightweight, easy to implement	Poor generalization
YOLOv5	87.6%	0.86	24 FPS	Real-time detection, robust	Occlusion handling
Vertex AI	91.2%	0.88	Slower	Highest accuracy, automated tuning	Cloud-dependent, latency

REFERENCES

- Abayaratne, S. B., Ilmini, W. M. K. S., & Fernando, T. G. I. (2019). Identification of snake species in Sri Lanka using convolutional neural networks. *Proceedings of the International Conference on Artificial Intelligence (SLAAI), Sabaragamuwa University of Sri Lanka, Sri Lanka*, 101 - 110. <https://slaai.lk/wp-content/uploads/2021/03/2019-Proceedings/Identification-of-snake-species-in-Sri-Lanka-Using-Convolutional-Neural-Networks.pdf>
- Ahmed, K., Gad, M. A., & Aboutabl, A.E. (2023). Snake species classification using deep learning techniques. *Multimedia Tools and Applications*, 83, 35117 - 35158. <https://doi.org/10.1007/s11042-023-16773-0>
- Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). *YOLOv4: Optimal speed and accuracy of object detection*. arXiv. <https://arxiv.org/abs/2004.10934>
- Google Cloud. (2023). *Vertex AI: Machine learning on Google Cloud*. Google. <https://cloud.google.com/vertex-ai>
- Iguernane, M., Ouzziki, M., Es-Saady, Y., El Hajji, M., Lansari, A., & Bouazza, A. (2025). Deep learning-based snake species identification for enhanced snakebite management. *AI*, 6(2), 21. <https://doi.org/10.3390/ai6020021>



An AI-Integrated Admin Panel with Theta Forecasting for E-Commerce Analytics

Rathnayake MMGNL¹, Vidanagama VGTN²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka.*

nuwanilrathnayake98@gmail.com¹

ABSTRACT

The explosive growth in e-commerce has generated huge amounts of operational data that, due to limited advanced yet easy-to-use analytic tools, cannot be exploited by small and medium-sized enterprises. To address this, we implemented and tested a full AI agent-based admin panel on a synthetic e-commerce dataset, with the goal of empowering non-technical SME users through intuitive natural language interfaces combined with accurate forecasting for enhanced data administration. We present a full AI agent-based admin panel for enhancing the administration of e-commerce data through the integration of a classical statistical model of forecasting, the Theta model, with an intelligent AI assistant enabling natural language CRUD (Create, Read, Update, Delete) functionality. The Theta forecasting model provides accurate short-term order and revenue projections with an R^2 of 92.1% for order forecasting, allowing effective planning of inventory and sales without requiring complex deep learning infrastructure. The AI agent, powered by a large language model and coordinated by the Model Context Protocol (MCP), interprets natural language commands and executes flexible, context-aware database operations securely with the secure API execution environment's command set. The system architecture, developed using the Flask framework, enables front-end visualization and back-end operation integration in real time seamlessly. The integrated system offers democratized data access and operational control to non-technical personnel, improving decision-making effectiveness without compromising database responsiveness and integrity. Voice command functionality, predictive customer segmentation using clustering, and improved security with fine-grained access controls and audit trails are a few of the future enhancements.

KEYWORDS: AI Agent, Data Analytics, E-commerce, Natural Language Processing, Theta Model

1. INTRODUCTION

The exponential expansion of e-commerce sites has led to an enormous accumulation of operational data. Small- and medium-sized enterprises (SMEs), however, often do not have the capacity to leverage this information effectively due to the lack of sophisticated yet simple-to-use analytic tools. Many existing solutions require advanced technical skills or involve cumbersome manual configuration, limiting their practical use. This creates a critical need for intelligent systems that can automate decision support and allow non-technical users to

interact with business data dynamically and intuitively. Legacy e-commerce control panels often lack flexible data querying and real-time control through natural language. Business users typically must use database-specific query languages or operate within rigid dashboards, which prevents them from realizing the full potential of their data. Furthermore, forecasting capabilities tend to be disconnected from operational controls, so users must switch between different tools to examine insights and take action on them. The objective of this project is to develop an AI Agent-Based Admin Panel that integrates a statistical

forecasting method with an intelligent AI agent. The system utilizes the Theta method for efficient order and revenue forecasting without requiring any complex infrastructure. It also features an AI agent that reads natural language commands to perform CRUD operations, making backend management easier for users of all levels.

Combining these elements in one platform, the system must democratize data access, render decision-making more efficient, and reduce dependency on technical personnel. This study aims to design and evaluate an AI-based admin panel integrating the Theta forecasting model with natural language for improving e-commerce decision support.

2. LITERATURE REVIEW

The convergence of AI and statistical methods is transforming the operation of e-commerce. Time-series forecasting is essential to supply chain optimization, and classical methods like the Theta model have been demonstrated to perform very effectively for the intermittent and seasonal demand patterns that are common in e-commerce. The Theta method, formulated by Assimakopoulos and Nikolopoulos (2000), is a novel and robust methodology that attains competitive accuracy by decomposing time-series data to represent trend and seasonal components. It is simple enough for environments with low computational resources in retail environments, which makes it extremely appropriate for SMEs.

Meanwhile, natural language interfaces are also contributing significantly towards the democratization of data handling. Conversational AI agents that translate conversational commands into database queries bridge this gap for non-SQL users. Recent advances leverage powerful language models combined with command

executors to offer precise data operations with conversational understanding and database consistency. It is also appropriate to utilize small serverless databases like SQLite because its ease of integration with Python and Flask frameworks facilitates rapid prototyping of admin panels and reduces latency for real-time interaction with AI agents. Merging these AI agents and statistical forecasting into a single platform is an emerging research area that can radically enhance decision velocity by allowing users to derive predictive insights and work within the same environment.

3. METHODOLOGY

3.1 System Architecture

The system architecture is designed to unify data forecasting, natural language processing, and database management seamlessly. The architecture, as illustrated in Fig. 1, consists of three significant layers:

Frontend: Built using the Flask framework and Jinja templating, this layer presents the analytics and provides a dynamic, interactive user interface to control the e-commerce platform.

Backend: The Flask backend runs the core business logic. It combines the forecasting models with the frontend and orchestrates database interactions through a set of APIs. It encompasses the Theta Model's Python code to generate order and revenue predictions.

AI Services Layer: This layer employs Gemini 2.5 to parse the natural language input from the user. The parsed intent is translated to secure database queries using the Google ADK (Agent Development Kit). Conversational context is controlled using the Model Context Protocol (MCP) for coherent

multi-turn interaction with the user.

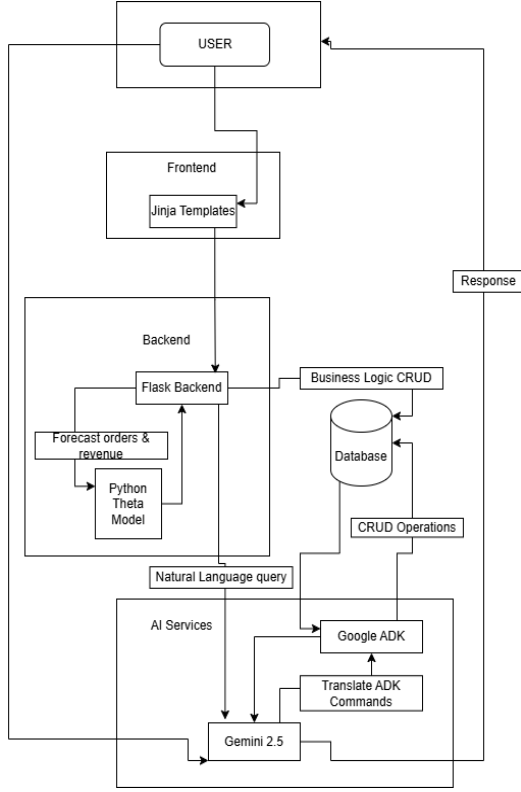


Figure 1: System Architecture

3.2 Forecasting Method: Theta Model

The Theta model was chosen because of its balance between simplicity, interpretability, and good performance on business time-series data. It is effective at modeling the underlying trends and seasonality and is resilient to the noise and outliers common in e-commerce data. The data used for forecasting in this implementation were synthetic, generated within the system by adding simulated e-commerce via the admin panel to mimic real world patterns.

The model works by decomposing a time series into two "theta lines" adjusted by a coefficient θ . The mathematical formula is provided in Eq (1).

$$Y_t(\theta) = 2\theta(X_t + X_{t-2}) + (1 - \theta)X_{t-1} \dots (1)$$

where $X_t X_t$ is the initial time series value at time t , and θ is the coefficient that adjusts the curvature of the line. The ultimate forecast combines a linear

component (where $\theta=0$) to capture the long-term trend and a smoothed component (where $\theta=2$) that places greater emphasis on recent changes. Forecasts are generated for the next 7 days using data from the past 21 days, with the potential for real-time updates as new data emerge. The Theta model parameters, including the θ coefficient, were tuned using a grid search over values from 0 to 2 to optimize for trend and seasonality capture. Validation was performed on the synthetic dataset, using R^2 as the primary evaluation metric, achieving 92.1% for order forecasting. Additional metrics like mean absolute error (MAE) were monitored to ensure robustness against outliers.

The most significant feature of the admin panel is the AI Agent for CRUD Operations, an AI assistant that allows non-technical users to manage the database using natural language. Users can perform operations such as "Update stock for product id 1 to 75 units" or "Show me details about order 12132". The Gemini 2.5 model accepts these inputs, maintains context across multiple steps, translates the instructions into the appropriate Google ADK syntax, and executes the database operations securely. The agent provides immediate and clear feedback through the UI, confirming the action taken and enhancing user confidence.

4. RESULTS AND DISCUSSION

The system installed provides a unified platform for operational management and analytics. The forecasting dashboard displayed in Fig. 2 displays graphically the historical data along with 7-day order and revenue forecasts. The Theta model is found to be doing well with an R^2 of 92.1% for order forecasting, demonstrating its suitability for the specified use case. Intelligent

recommendations, such as high-demand alerts and logistics planning alerts, are also provided by the dashboard based on the forecast data.

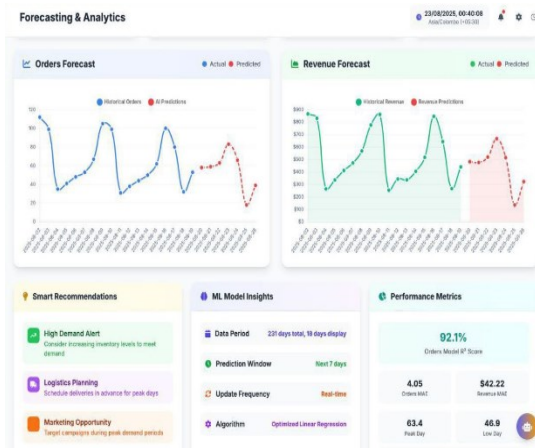


Figure 2: Forecasting and Analytics Dashboard



Figure 3: Orders Forecast Graph



Figure 4: Revenue Forecast Graph

Fig. 3 illustrates the Orders Forecast Graph, depicting a 7-day projection of order volumes based on the Theta model's

analysis of historical data, highlighting trends and expected demand fluctuations.

Fig. 4 presents the Revenue Forecast Graph, showcasing predicted revenue over the same period, providing clear visual insights into future financial performance with high accuracy.

The AI assistant functionality was evaluated on various CRUD operations. The system can read product data, update quantities, modify prices, and modify order status correctly in response to voice commands, as evident in the system screenshots. For instance, if the user types "Hi, can you change the Price of Product id 1 to 3.99," the agent first acknowledges ("Update Price for Thunder Strike Original?") and, if sent a "Yes," makes the update and sends back a confirmation of completion. This interactive, context-aware interaction significantly lowers the technical barrier to manage an e-commerce backend, thereby increasing operational efficiency. The combination of these features specifically solves the issue of the decoupled analytical and operational tools typical in current platforms.

5. CONCLUSION

This study effectively created a new AI agent-based admin panel that combines sophisticated forecasting with automatic, natural language-based operational controls within one unified platform. Through the exploitation of the robust Theta model for time-series forecasting and an intelligent AI agent powered by Gemini 2.5 and Google ADK, the system effectively democratizes data access and enhances data-driven decision-making for e-commerce managers. The primary challenge encountered was the ambiguity of natural language, which required careful architectural design and development of a context protocol (MCP)

to enable reliable command execution and maintain database integrity.

Future development will be aimed at expanding the natural language understanding to accommodate more diverse business questions. Voice command capability is planned to be included to enable hands-free operation.

Further improvements include incorporating predictive clustering for customer segmentation and refining system security with granular access controls and audit trails for all activities undertaken by the AI agent.

REFERENCES

- Assimakopoulos, V., & Nikolopoulos, K. (2000). The theta model: a decomposition approach to forecasting. *International Journal of Forecasting*, 16(4), 521-530. [https://doi.org/10.1016/S0169-2070\(00\)00066-2](https://doi.org/10.1016/S0169-2070(00)00066-2)



Usability Improvement Through Customer Feedback: Evidence from the Usability Evaluation of a Pharmacy Management Mobile Application

Dapanadurage IN¹, Gunasekara RPTH²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

ishaninuwanthikadd@gmail.com¹, hansikag@wyb.ac.lk²

ABSTRACT

Usability testing is a critical component in assessing the effectiveness, efficiency, and user satisfaction of mobile applications, particularly in domains where user experience directly affects operational accuracy and productivity. This research investigates how customer feedback can systematically transform into actionable design improvements through a usability evaluation of the SithMediPro Pharmacy Management Mobile Application, a widely adopted platform used by pharmacy owners to manage daily business operations. A mixed-method usability assessment has conducted using task-based evaluations, structured questionnaires, and open-ended feedback. User responses highlighted key issues such as missing navigation options, ambiguous tab labels, and redundant information in data tables. These findings have analyzed and translated into targeted interface modifications that improved navigation consistency, information clarity, and visual organization. The refined design achieved higher task completion rates and increased user satisfaction, as proved by improved System Usability Scale (SUS) scores and qualitative feedback. The results emphasize the importance of involving end-users throughout the evaluation process, demonstrating that continuous, feedback-driven usability testing not only enhances system performance but also fosters sustainable user engagement—particularly in healthcare-related mobile applications where efficiency and clarity are essential.

KEYWORDS: Customer feedback, Mobile healthcare application, Pharmacy management, Usability improvement, User experience

1. INTRODUCTION

Usability is one of the most critical quality attributes in software systems because it determines how easily users can perform their tasks and how satisfied they feel while doing so (Nielsen, 2024). In the healthcare sector, usability is especially important because medical and pharmacy staff depend on digital tools to perform precise, time-sensitive operations. Even minor usability problems can slow down workflow or cause errors that affect customer service.

This paper aims to show how customer feedback has used not only to evaluate usability but also to guide usability improvements.

By analyzing real user responses and applying their suggestions, the system has enhanced to meet user needs better. The study therefore provides direct evidence that structured customer feedback can drive meaningful usability improvements.

While the application provided strong functionality, early user feedback revealed several practical usability concerns, including difficulty navigating between pages and confusion when reading data tables.

2. LITERATURE REVIEW

Usability is widely recognized as a critical factor influencing interactive system success, particularly in healthcare

environments where inefficiencies can affect both workflow and service quality (Nielsen, 2024). Mobile health (mHealth) applications have introduced the challenges, including limited screen size, diverse user groups, and varied usage contexts. Recent studies emphasize that mixed-method usability evaluation—incorporating surveys, observational data, and qualitative insights—produces more comprehensive and actionable findings (Maramba, Chatterjee, & Newman, 2019).

The System Usability Scale (SUS) has widely used for evaluating digital systems due to its reliability and simplicity. It generates a single score between 0 and 100, making cross-study comparison straightforward. Research from 2020–2024 suggests that SUS should complement with qualitative feedback and task-based assessments to provide a complete understanding of user experience (Sauro & Lewis, 2018).

Recent empirical studies highlight the effectiveness of iterative, user-centered design in improving usability outcomes. Research involving pharmacy and clinical systems shows that targeted interface enhancements—such as clearer screen labels, simplified data tables, and improved navigation—can significantly enhance user performance and satisfaction (Alnanih, Ormandjieva, & Radhakrishnan, 2014). In addition, contemporary mHealth reviews emphasize evaluating domain-specific usability challenges, such as navigating long inventory lists in pharmacy management applications.

Pharmacy management systems require frequent data retrieval and repetitive tasks. Studies in pharmacy informatics show that poor labeling, redundant information, and inconsistent navigation can increase cognitive load and reduce efficiency. Conversely, domain-specific design improvements result in measurable gains in task speed and user satisfaction (Alnanih, Ormandjieva, & Radhakrishnan, 2014). These studies directly support the focus areas addressed in this evaluation of *SithMediPro*.

3. METHODOLOGY

3.1 Research Design

A descriptive survey-based method was used to evaluate the usability of the *SithMediPro* mobile application. The study combined both quantitative and qualitative approaches to gather comprehensive user feedback.

3.2 Participants

The participants were 38 pharmacy owners who actively use the *SithMediPro* mobile application in their daily business operations. They were selected for their direct engagement with the system, which provided practical insights into its usability.

3.3 Data Collection

A Google Form questionnaire consisting of 17 Likert-scale items and several open-ended questions was distributed.

The Likert questions rated usability aspects such as navigation, table clarity, ease of scrolling, and satisfaction (1 = Strongly Disagree to 5 = Strongly Agree). The open-ended section invited improvement suggestions such as “add logout button”, “improve scrolling”, and “show pharmacy name.”

3.4 Data Analysis

Quantitative data were analyzed in Microsoft Excel, with mean scores and standard deviations calculated. Bar charts were created to visualize ratings across modules, including Daily Invoices, Purchases, Receipts, and Payments. Qualitative responses were grouped thematically to identify recurring usability issues.

4. RESULTS AND DISCUSSION

The results revealed that overall user satisfaction with the SithMediPro application was high. Most participants rated the system as simple, reliable, and effective for daily pharmacy management tasks. The mean usability score across all questions exceeded 4.0 on the Likert scale, indicating a positive acumen in general.

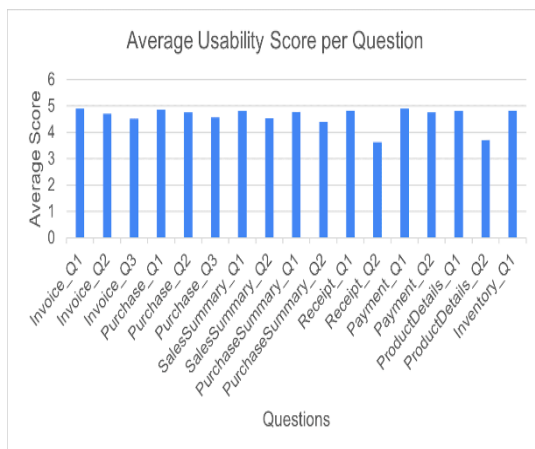


Figure 1: Average Scores for Usability Questions

Figure 1 presents the average ratings for core usability parameters. The results show that users particularly appreciated the clarity of invoices, purchase summaries, and receipt management interfaces.

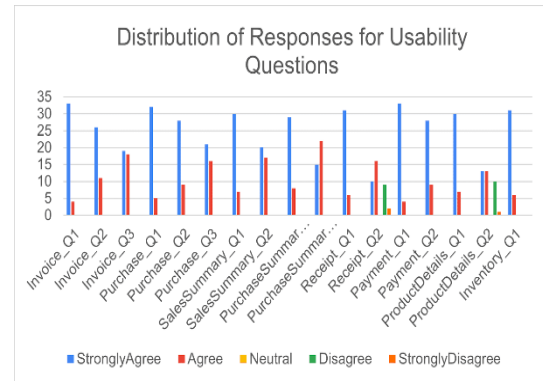


Figure 2: Distribution of Responses for Usability Question

Figure 2 illustrates the distribution of responses across key usability factors. More than 70% of respondents rated the system as “simple” or “very simple to use,” confirming that the interface design effectively supports basic user operations.

However, several usability challenges were identified through open-ended feedback. The most common concerns included:

- Lack of a logout button, which limits user control and flexibility.
- Difficulty scrolling through long data lists.
- Absence of pharmacy names and section headers on some screens.
- Redundant columns in tables, make data reading more time-consuming.

Table 1: Summarizes the key improvement areas reported by participants.

Area	User Feedback
Navigation	Add logout and back navigation buttons
Interface Clarity	Display pharmacy name and section titles

Table Design	Take out unnecessary columns to make it easier to read
Scrolling	Improve responsiveness and smoothness

Overall, the results confirm that SithMediPro provides a solid foundation for managing pharmacy workflows, though enhancements in layout and navigation could make it even more intuitive and efficient.

The findings clearly confirm that customer feedback played a crucial role in guiding the usability improvements of the SithMediPro Pharmacy Management Mobile Application. By directly considering users' experiences and suggestions, the development team has identified real problems and implemented effective design changes that improved the system's navigation, clarity, and overall usability. This approach proved far more reliable than relying on developer assumptions alone. The results confirm that continuous collection and application of customer feedback is a practical, evidence-based method for achieving usability improvements, supporting the principles of iterative, user-centered development highlighted in previous research (Sauro & Lewis, 2018; Nielsen, 2024).

5. CONCLUSION

This study clearly shows that customer feedback can make a real difference in improving system usability. The usability evaluation of the SithMediPro Pharmacy Management Mobile Application identified several areas for improvement, including navigation, layout, and data presentation.

The feedback collected from pharmacy owners provided practical ideas that have applied later to make the system easier to use and more user-friendly.

By listening carefully to users and considering their experiences, we were able to make meaningful changes that directly improved satisfaction and usability. This process proved that customer feedback is not only helpful in finding problems but also for creating better solutions.

In conclusion, continuous user feedback is key to keeping a system effective and easy to use. When users have given the chance to share their opinions and experiences, their feedback becomes a valuable path to ongoing usability improvements and a better user experience.

REFERENCES

- Alnanih, R., Ormandjieva, O., & Radhakrishnan, T. (2014). Empirical Evaluation of Intelligent Mobile User Interfaces in Healthcare. In M. Sokolova, & P. van Beek (Eds.), *Advances in Artificial Intelligence*, (pp. 23-34). Springer International Publishing. https://doi.org/10.1007/978-3-319-06483-3_3
- Lewis, J. R., & Sauro, J. (2018). Item Benchmarks for the System Usability Scale. *Journal of Usability Studies*, 13(3), 158-167.
- Maramba, I., Chatterjee, A., & Newman, C. (2019). Methods of usability testing in the development of eHealth applications: A scoping review. *International Journal of Medical Informatics*, 126, 95-104. <https://doi.org/10.1016/j.ijmedinf.2019.03.018>
- Nielsen, J. (2024, January 30). *10 Usability Heuristics for User Interface Design*. Nielsen Norman Group. <https://www.nngroup.com/articles/ten-usability-heuristics/>



Web-Based Project Management System for Midsize Research Institutes

Sameeha MR¹, Arudchelvam T²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

[¹*sameeharisan@gmail.com*](mailto:sameeharisan@gmail.com), [²*arul@wyb.ac.lk*](mailto:arul@wyb.ac.lk)

ABSTRACT

Managing multiple projects in a research institute or software company is a typical task which has to be properly done with high responsibility. There are challenges and limitations in manually managing projects. Therefore, developing a project management system gives comfortable, easy and efficient environment for managing projects. This paper presents the design and implementation of a web-based Research Project Management System developed for midsize research institutes. The project addressed challenges in managing research project records, documents, and administrative actions which were previously fragmented across spreadsheets, emails, and shared drives. The scope of the system included essential project lifecycle functions such as creating, updating, and archiving project data; a secure document repository; user management with role-based access control; audit logging; and CSV export for reporting. Methodologies employed for design and implementation included requirements elicitation through discussions with staff, modular software design principles, and iterative prototyping. The system was implemented using Python with a lightweight web framework, SQLite for development database management, and role-based security measures to ensure accountability and compliance. Evaluation demonstrated improved data accessibility, streamlined workflows, and enhanced accountability through integrated audit logging and reporting functionalities.

KEYWORDS: Python, Research Management System, Role-Based Access Control, SQLite, Web Application

1. INTRODUCTION

Research institutes depend heavily on effective project coordination, documentation, and progress monitoring to deliver high-quality outputs within limited time and resource constraints. However, many midsize institutes continue to rely on fragmented tools such as spreadsheets, email threads, and isolated document repositories, which often result in duplicated work, inconsistent records, and reduced transparency (Schreier et al., 2006; Kanza and Knight, 2022). These challenges are further intensified by the collaborative and multi-stakeholder nature of research environments, where project teams require reliable access to shared information and traceable documentation (Myneni et al., 2016)

Despite the availability of general project management platforms, most existing solutions have not tailored to the operational and administrative needs of research institutes (Pacheco et al., 2025). They often lack structured document traceability, role-based access suitable for hierarchical research organizations, or customizable modules that support research-specific workflows (Tian et al., 2021; Zikra et al., 2025). As a result, research staff often experience inefficiencies in tracking project lifecycles, managing updates, and controlling accountability.

To address these gaps, this study presents the design and implementation of a web-based Research Project Management System specifically developed for midsize research institutes.

The system integrates project tracking, centralized document management, role-based access control (RBAC), and activity logging into a unified platform. Developed using Python, a lightweight web framework, and SQLite as the backend database, the system employs a modular three-layer architecture that enhances maintainability and scalability.

The main target of this work is twofold: (1) to provide a practical solution that streamlines administrative processes within research institutes, and (2) to show a lightweight system architecture suitable for organizations with limited IT infrastructure. While this study contributes a functional system designed to fill a practical gap, it also aims to advance discussions on how research organizations can adopt adaptable, resource-efficient digital solutions (Moore & Shangraw, 2011).

2. LITERATURE REVIEW

Research and Development administration in midsize institutions faces significant challenges primarily due to fragmented data management, where critical project records have often scattered across various decentralized platforms like spreadsheets and electronic mail. This dispersal compromises data integrity, security, and the ability to be audited. Furthermore, research environments must balance the structured control necessary for funding compliance with the flexibility required for creative research and development work (Borštnar & Pucihar, 2014). To counteract this data fragmentation and ensure accountability, a lightweight Project Management Information System is essential to unify processes and mitigate administrative friction (Mokhtar & Khayat, 2022).

A functional Project Management Information System must set up strong centralized governance. Key requirements include a unified repository for all project documents and records, the implementation of robust Role-Based Access Control to manage data sensitivity and permissions, and comprehensive audit logging to track all user actions and system events. These features are critical for maintaining integrity, transparency, and meeting compliance obligations (Borštnar & Pucihar, 2014; Chauhan et al., 2019). Current administrative inefficiencies often stem from decentralized systems and manual data management, hindering version control, accessibility, and accountability.

When considering architecture, large Enterprise Systems such as InfoEd, Cayuse, or Workday Research are often too costly and complex for resource-constrained midsize institutions. Consequently, the development of a lightweight custom system is a pragmatic and strategic alternative. Utilizing technologies like Python/Flask allows for high customization, lower complexity, and future scalability with minimal overhead (Chauhan et al., 2019). Institutional Repositories like DSpace and EPrints offer strong long-term preservation but lack the necessary focus on operational project lifecycle management or fine-grained permissions.

Historically, project management relied on predictive methodologies like Waterfall, which struggle to handle the high uncertainty and change inherent in Research and Development (de Almeida Parizotto et al., 2020). Modern practice favors adaptive paradigms, leading to the dominance of Hybrid Project Management, which combines the

structure of predictive phases with the agility of adaptive techniques (Hajek & Krupa, 2024). The centralized Project Management Information System is vital in this context, as it provides the essential technical infrastructure, transparency, and governance features (such as budget tracking and status history) needed to support adaptive execution while simultaneously ensuring compliance and project viability (Mokhtar & Khayyat, 2022; Mirzaei et al., 2024). Based on this review, this system adopts the lightweight custom approach, combining a project registry, document repository, Role-Based Access Control, and audit logs to strike a balance that is simpler and less expensive than enterprise systems while providing more operational focus than institutional repositories.

3. METHODOLOGY

The system had designed using a modular, layered architecture to separate presentation, application, and data concerns. Design goals prioritized simplicity, security, and maintainability. Constraints included limited infrastructure and a requirement for rapid prototyping. A server-side rendered monolithic application, had used as it simplified development and deployment while maintaining adequate functionality.

The architecture had divided into three layers: (1) Presentation Layer: Provided a server-rendered web user interface for forms, lists, and document

upload or download, handling user interaction, and input validation. (2) Application Layer: Managed routing, request handling, and business logic such as project lifecycle management and validation rules. (3) Data Layer: Contained Object-Relational Mapping (ORM)-managed data models and persistence for projects, users, documents, and audit logs, as well as filesystem or object storage for document binaries.

This architectural separation allowed the individual layers to evolve independently and made testing and future substitution of components more straightforward, for example, replacing the local filesystem with cloud storage. To represent the system design clearly, several modeling techniques were employed. The Entity-Relationship Diagram is shown in Figure 1. Use-case modeling captured the primary actors Administrator, Researcher, and Assistant and interactions, which included managing users, managing projects, uploading documents, and performing import or export operations. Class and entity modeling defined the primary domain entities such as User, Project, Document, Project Status, and Audit Log, along with the relationships between them. These models guided both the implementation and the schema design.

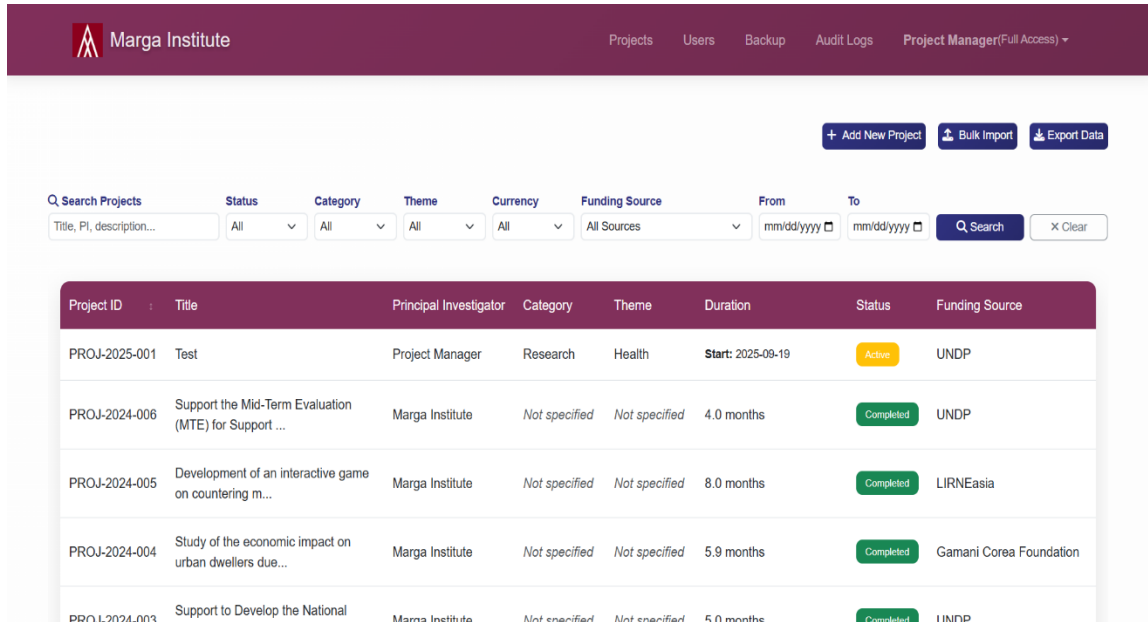


Figure 2: User Interface Screenshot

4. METHODOLOGY

The completed system delivered a functional, centralized platform that significantly improved project and document management workflows at the institute. Staff members were able to access project information more efficiently, reduce their reliance on spreadsheets and email communication, and maintain greater consistency in record keeping. The integrated document repository replaced fragmented storage practices, enabling users to upload, retrieve, and categorize files within a secure and traceable environment. Audit logs provided a complete record of user actions, improving accountability and facilitating oversight.

Quantitative evaluation through pre- and post-implementation surveys proven meaningful improvements across multiple dimensions. Analysis of the twelve Likert-scale items revealed increases in perceived information accessibility,

efficiency, ease of use, and appropriateness of role-based permissions. Departments reported reduced difficulty in arranging project files, faster completion of administrative tasks, and smoother internal coordination. These improvements have confirmed the system’s ability to address the operational challenges which have revealed during initial requirements usage

Performance testing further confirmed the system’s responsiveness, with average request latency existing below 120 milliseconds under moderate load, and no failures occurring during document handling stress tests. Security tests demonstrated that RBAC effectively prevented unauthorized access to restricted features. The logging subsystem captured both user and system events with high completeness, improving transparency and simplifying error diagnosis.

Comparative analysis reinforced that the proposed system occupies an important middle ground between generic project management platforms and enterprise research administration systems. While commercial tools offer general workflow features, they do not support research-specific data structures or institutional hierarchies. Conversely, enterprise systems provide extensive functionality but impose high financial and technical burdens. This system's lightweight and customizable design makes it a practical choice for midsize research institutes which are seeking to enhance administrative capacity without committing significant resources.

5. CONCLUSION

The Web-Based Research Project Management System for the selected Institute successfully addressed key administrative challenges by providing a centralized, secure, and efficient platform for managing research projects and associated documents. The system enhanced data integrity, accessibility, and accountability while aligning with the institute's operational structure.

The project demonstrated how lightweight, open-source technologies can deliver scalable solutions for midsize organizations with limited resources. Lessons learned from this project emphasize the importance of aligning software design with organizational workflows and end-user capabilities.

REFERENCES

- Borštnar, M. K., & Pucihar, A. (2014). Impacts of the Implementation of a Project Management Information System – a Case Study of a Small R&D Company. *Organizacija*, 47(1), 14–23.
- Chauhan, N., Singh, M., Verma, A., Parasher, A., & Budhiraja, G. (2019). Implementation of database using python flask framework. *International Journal of Engineering and Computer Science*, 8(12), 24894 - 24899. <https://doi.org/10.18535/ijecs/v8i12.4390>
- de Almeida Parizotto, L., Tonso, A., de Carvalho, M. M., de Almeida Parizotto, L., Tonso, A., & de Carvalho, M. M. (2020). The challenges of project management in small and medium-sized enterprises: a literature review based on bibliometric software and content analysis. *Gestão & Produção*, 27(1). <https://doi.org/10.1590/0104-530x3768-20>
- Hajek, J., & Krupa, M. (2024). Hybrid project management models: A systematic literature review. *International Journal of Project Organization and Management*, 16(3). <https://doi.org/10.1504/ijpom.2024.10056237>
- Kanza, S., & Knight, N. J. (2022). Behind Every Great Research Project Is Great Data Management. *BMC Research Notes*, 15(1). Retrieved from <https://bmcresearchnotes.biomedcentral.com/articles/10.1186/s13104-022-05908-5>
- Mirzaei, M., Mabin, V. J., & Zwikael, O. (2024). Customising Hybrid project management methodologies. *Production Planning & Control*, 36(9), 1 - 18. <https://doi.org/10.1080/09537287.2024.2349231>
- Mokhtar, R., & Khayyat, M. (2022). A Comparative Case Study of Waterfall and Agile Management. *SAR Journal - Science and Research*, 5(1), 52–62. <https://doi.org/10.18421/sar51-07>
- Moore, S., & Shangraw, R. (2011). Managing Risk and Uncertainty in Large-Scale University Research Projects. *Research Management Review*, 18. https://files.eric.ed.gov/fulltext/EJ980462.pdf?utm_source=chatgpt.com
- Myneni, S., Patel, V. L., Bova, G. S., Wang, J., Ackerman, C. F., Berlinicke, C. A., & Zack, D. J. (2016). Resolving complex research data management issues in biomedical laboratories: Qualitative study of an industry–academia collaboration. *Computer Methods and Programs in Biomedicine*,

126, 160 - 170. <https://doi.org/10.1016/j.cmpb.2015.11.001>

- Pacheco, L., Tavares, F. O., Ramazanov, M., Ávila, J. F., Albuquerque, H., Silva, F. M., ... Guzmán, S. M. (2025). Research Management in Higher Education Institutions from Developing Countries: An Analysis for Bolivia and Paraguay. *Administrative Sciences*, 15(4), 131. <https://doi.org/10.3390/admsci15040131>
- Schreier, A. A., Wilson, K., & Resnik, D. (2006). Academic Research Record-Keeping: Best Practices for Individuals, Group Leaders, and Institutions. *Academic Medicine*, 81(1), 42 - 47. <https://doi.org/10.1097/00001888-200601000-00010>
- Tian, F., Wang, T., Liang, P., Wang, C., Khan, A. A., & Babar, M. A. (2021). The Impact of Traceability on Software Maintenance and Evolution: A Mapping Study. *ArXiv*. <https://arxiv.org/abs/2108.02133>
- Zikra A. A., Darnilasari, A., Yanuary, R., & Sari, K. (2025). Design of Web-Based Project Management System with Multi-Level Role-Based Access Control. *Journal of Computer Science and Informatics Engineering (CoSIE)*, 4(3), 201 - 215. <https://doi.org/10.55537/cosie.v4i3.1179>



Wood Type Identification System Using Image Processing and Machine Learning Techniques

Dewmindi PLU¹, Jayakody JRKC²

^{1,2}*Department of Computing and Information Systems, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

pl.umeshika99@gmail.com¹, kithsirij@wyb.ac.lk²

ABSTRACT

Accurate wood species identification is very important in many areas like forestry, furniture production, and forensic work. Most existing methods still rely on manual visual checking of wood texture and grain, so the process becomes slow, subjective, and difficult to apply for large-scale tasks. Previous studies that used traditional image processing or machine learning with handcrafted features could not reach stable accuracy because the datasets were limited and not suitable for real-world image differences. To solve this issue, this research developed a convolutional neural network (CNN) model that can automatically identify wood species from images. A total of 1,272 original wood images were taken in different image ratios (4:3, 16:9, and square) using a mobile phone camera, and the dataset was expanded to 2,544 images through image augmentation. Three types of CNN models were trained and tested to check which one gives better results in species classification. The CNN model with two convolution layers showed the highest accuracy of 88.19%, which was better than the other two models. The findings indicate that a properly trained CNN can reduce human effort in wood identification and make the analysis process more efficient and consistent.

KEYWORDS: Convolutional Neural Network, Data Augmentation, Deep Learning, Image Classification, Wood Species Identification

1. INTRODUCTION

Wood species identification has a very important role in many fields, for example timber trade, making furniture, forest management, and forensic study. Correct recognition of wood types is necessary to keep quality, prevent illegal logging, and help scientific records. Usually, experts do this by looking at microscopic and visible structures of wood. This manual way works sometimes, but it depends much on human skill, takes a lot of time, and can be inconsistent because of personal judgment and changing conditions.

Recently, researchers have tried to use digital image processing and normal machine learning for wood recognition.

But these methods usually need handcrafted features and cannot handle changes in texture, color, light, or image quality very well. Therefore, their accuracy and reliability are limited, especially for real-world images from devices like mobile phones.

To solve these problems, this study uses a deep learning method with Convolutional Neural Networks (CNNs) to identify wood species automatically from images. A dataset of 1,272 original images was taken by a mobile phone and then increased to 2,544 images using augmentation for better generalization. Three CNN architectures were tested, and the model with two convolution layers worked best, reaching 88.19% accuracy. The results show that CNN methods can give a faster and more reliable way for

wood species identification compared to traditional approaches.

2. LITERATURE REVIEW

Automatic wood identification gets more attention because it is very useful for forestry, timber trade, conservation, and forensic work. Normal identification needs human specialists to look at wood structure, but it is expensive and slow. Some recent reviews say there is a big need for fast and easy computer-vision methods that can help workers in the field (Wäldchen & Mäder, 2018).

Early computer-vision methods use handcrafted features and classical classifiers, but they often do not work well when light, resolution, or camera type changes. A recent review shows that computer-vision methods look promising, but different datasets and imaging conditions still make species-level classification hard in the real-world (Silva et al., 2025).

With deep learning becoming popular, many researchers try Convolutional Neural Networks (CNNs) for macroscopic and microscopic wood images. Kwon and Lee (2017) show that a smartphone-based system can classify several Korean softwood species with photos from mobile cameras, but the dataset was small. Ergun et al. (2024) test lightweight networks like ShuffleNet and find that careful design can give good performance on forest species, but there is always a trade-off between accuracy and model size for field use. Studies with microscope images get very high accuracy because small structures are clear. For example, Zheng et al. (2024) classify 15 species almost perfectly with cross- and tangential-section images.

Many specialized networks also get very high numbers on microscope datasets (Zheng, 2024). This shows that imaging type affects accuracy, but microscope images are not easy to use in the field.

Although many existing studies focus on large multi-species datasets, the present study uses only two species due to limited availability of high-quality, field-captured wood images. This constraint reflects a common real-world challenge identified in previous research, where collecting clean, well-labeled macroscopic wood datasets is difficult. Therefore, this study focuses on two commonly available species in Sri Lankan sawmills, making it a practical starting point for developing an automated identification system.

3. METHODOLOGY

This study uses a clear method to identify wood species with deep learning and image processing. The work has three main steps: collect data, prepare data, and build models. Each step was planned carefully to make sure the model can work correctly and also perform well on new images not used before.

3.1 Workflow

Figure 1 provides a complete overview of the workflow followed in this study, including data collection, preprocessing, model training, and performance evaluation. This figure is introduced here to help readers clearly understand the sequence of steps used in the development of the proposed wood identification system.

3.2 Data Collection

The first step of this study was collecting dataset for training and testing CNN models. A total of 1,272 images

were taken, including two wood types: 636 images of Lunu Midella and 636 images of Mahogany, from some local sawmills. The dataset was made balanced for two classes to avoid bias during model training.

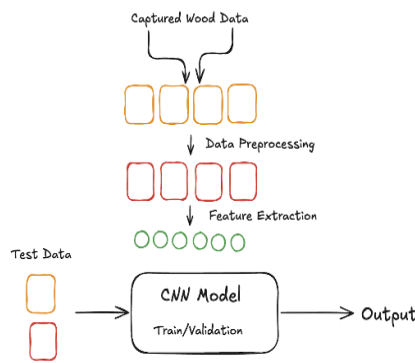


Figure 1: Overview of Model Development and Training Process

Images were captured with a mobile camera under normal light, to show real-world conditions. To make the dataset more varied and useful for different situations, three aspect ratios were used: 1:1, 16:9, and 4:3, with more than 200 images for each species at each ratio. This way helped models stay stable and accurate even if images come from different cameras or settings.

A limitation of the dataset is that it includes only two species, Lunu Midella and Mahogany. This restriction was mainly due to limited access to well-labeled wood samples across multiple sawmills during the data collection period. Since the objective of this initial study was to test whether a CNN can distinguish between visually similar species using field-captured images, a two-class dataset was considered sufficient for the first phase of model development. However, the limited species range is acknowledged as a

constraint, and future work will expand the dataset to include more commercially important wood types.

3.3 Data Preprocessing

After gathering the images, some preprocessing steps were done to make data better in quality and organization. The main purpose was to make the dataset ready for deep learning and to reduce bias or noise which can make model accuracy lower.

3.3.1 Feature Analysis

Before training, every image was resized to 224×224 pixels. Then, first feature analysis was done to check that two wood species had clear and measurable visual differences.

Color Analysis: Red, Green, and Blue (RGB) channel intensity plots were made for both species. This showed how colors were spread and made visible differences between Lunu Midella and Mahogany.

Contrast Analysis: Standard deviation of contrast was calculated for many samples. This showed clear differences in texture patterns for each species. These results confirmed that images had enough distinct features for CNN to learn.

3.3.2 Data Augmentation

Because the dataset was small, augmentation methods like rotation, horizontal and vertical flipping, and random shifting were applied. Total images became 2,544. This increased diversity and helped reduce overfitting. Augmentation also made the model stronger to changes in orientation and light.

3.4 Model Development

After preprocessing, the prepared dataset is split into three parts: 70% for

training, 10% for validation, and 20% (255 images) reserved for final testing with unseen images. A custom CNN was built from scratch with a simple Sequential layout tuned to capture fine wood texture details. The training set taught the network to recognize visual patterns, the validation set guided hyperparameter adjustments, and the test set provided an unbiased measure of performance. Using augmented images, I experimented with several model variants that differed in convolution and pooling arrangements, keeping a balance between model complexity and generalization to avoid overfitting. During training the network learned texture and color cues by repeated parameter updates with an appropriate learning rate and loss function; validation feedback prompted changes that reduced errors. The best model, which had two convolutional layers, achieved 88.19% accuracy on mobile-captured images, indicating a strong ability to distinguish the wood species.

3.4.1 CNN Model Architecture

The CNN architecture used in this study was designed to balance learning capacity and computational efficiency. The model begins with two convolutional layers (32 and 64 filters), which extract low-level features such as edges, color gradients, and basic grain structures. A MaxPooling layer then reduces spatial size to avoid overfitting. To keep the model lightweight, only two additional convolutional layers (128 and 256 filters) were included, enabling the network to learn deeper and more abstract wood texture patterns without excessive parameter growth. Dropout layers were inserted to improve generalization. This

simplified design results in a more manageable number of parameters while maintaining strong classification performance.

4. RESULTS

This section shows how three CNN models classify Lunu Midella and Mahogany. All used the same dataset and preprocessing. The difference is the number of convolutional layers, which affects accuracy and feature learning.

Table 1 shows accuracy. The one-layer CNN reached 85.83%, detecting simple textures but missing complex patterns. The two-layer CNN was best at 88.19%, learning richer features without overfitting. The three-layer CNN was lowest at 83.46%, likely too complex for the dataset.

Table 2 shows precision, recall, and F1-score.

Table 1: Overall Performance Metrics

Model Architecture	Final Test Accuracy (%)
CNN(1 Convolutional Layer)	85.83
CNN(2 Convolutional Layer)	88.19
CNN (3 Convolutional Layer)	83.46

Table 2: Precision / Recall and F1-Score

Model Architecture	Precision	Recall	F1-Score
CNN (1 Conv Layer)	0.86	0.85	0.85
CNN (2 Conv Layers)	0.89	0.88	0.88
CNN (3 Conv Layers)	0.84	0.83	0.83

The two-layer CNN had the best results (0.89, 0.88, 0.88), the one-layer

was moderate (0.86, 0.85, 0.85), and the three-layer was weakest (0.84, 0.83, 0.83), suggesting deeper networks may overfit small datasets.

Overall, the two-layer CNN balances learning and generalization best, making it most suitable for this study.

5. CONCLUSION

This study showed that Convolutional Neural Networks (CNNs) can effectively identify wood species, especially Lunu Midella and Mahogany, with an accuracy of 88.19%. The use of balanced data, proper preprocessing, and image augmentation helped to improve both robustness and accuracy of the model. Among the tested architectures, the CNN with two convolutional layers gave the best performance while keeping computational cost low. The results suggest that deep learning can serve as a dependable alternative to manual wood identification techniques.

One of the main limitations of this study is the use of only two wood species, which reduces the generalizability of the model. This limitation arises primarily from restricted dataset availability rather than a methodological choice. Despite this constraint, the results clearly demonstrate that a CNN can successfully differentiate between the two selected species, indicating strong potential for scaling the system to more classes when a larger dataset becomes available.

Future studies will focus on collecting additional species, incorporating transfer learning models, and evaluating performance across a wider range of wood textures. Expanding the dataset will also enable the development of a more robust, field-ready

mobile application for real-time wood identification.

REFERENCES

- Ergun, H., Demir, S., & Yilmaz, T. (2024). Lightweight convolutional networks for automatic forest species identification. *Forestry Informatics Journal*, 12(3), 88 - 99.
- Kwon, O., & Lee, J. (2017). Smartphone-based wood species identification using convolutional neural networks. *Journal of Forestry Research*, 28(5), 1239 - 1247.
- Silva, J., Bordalo, R., Pissarra, J., & Palacios, P. (2025). Computer Vision-Based Wood Identification: A Review. *Forests*, 13, 2041. <https://doi.org/10.3390/f13122041>
- Wäldchen, J., & Mäder, P. (2018). Plant species identification using computer vision techniques: A systematic literature review. *Archives of Computational Methods in Engineering*, 25(2), 507 - 543. <https://doi.org/10.1007/s11831-016-9206-z>
- Zheng, S., Han, X., & Luo, Y. (2024). Deep CNN-based microscopic wood species identification with cross- and tangential-section images. *Wood Science and Technology*, 58(3), 621 - 639.



Analysis of Pollutant Exceedances and Treatment Efficiency towards Sustainable Wastewater Management in Export Processing Zones (Seethawaka & Horana)

Kusumsiri GKAD¹, Prabodanie RAR²

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka*
[akshiladilmini2000@gmail.com¹](mailto:akshiladilmini2000@gmail.com)

ABSTRACT

Seethawaka and Horana Export Processing Zones (EPZs) are two of the fastest growing industrial zones in Sri Lanka with immense amounts of wastewater to be treated by centralized plants. Despite having centralized wastewater treatment plants (WWTPs), frequent exceedances above national discharge standards become a matter of prime concern. Grounded on 96 weeks of secondary data of the garment, rubber, and plastic industries, the study analyzed pollutant variability using descriptive, correlation, regression, and comparative tests. Findings indicated low correlations of rainfall with pollutant concentration, which indicated no meteorological impact. Industrial category was not significant, which implies that treatment inefficiency, variation in operation, and uncontrolled emissions are among top reasons for exceedances. COD, BOD, and TSS were the usual parameters to be in excess. Although the overall treatment performance was satisfactory, zone-to-zone variability indicates that a blanket policy for centralized treatment is not sufficient. Adaptive operation controls, improved monitoring, and sectoral pretreatment standards are the suggestions of this research to impart maximum sustainability to the industrial wastewater management in Sri Lanka.

KEYWORDS: Export Processing Zones, Pollutant Exceedance, Treatment Efficiency, Wastewater Treatment

1. INTRODUCTION

Export Processing Zones such as Seethawaka and Horana are economically significant to Sri Lanka with respect to garment, rubber, and plastic manufacturing industries. These zones produce large amounts of wastewater that needs to be treated according to the discharge criteria of the Central Environmental Authority (CEA). Despite having such facilities (i.e., centralized wastewater treatment plants designed to collect, treat, and discharge industrial effluent), there are frequent exceedances of significant parameters such as COD, BOD, and TSS are reported, which may question the reliability of the overall treatment process. Operational inefficiencies, influent heterogeneity of

quality, and ineffective monitoring have been cited as potential causes in studies (Hemachandra & Pathiratne, 2018). Therefore, the study seeks to identify factors responsible for pollutant exceedance and to measure treatment efficiency in both zones.

2. LITERATURE REVIEW

2.1 Wastewater Treatment Models and Performance

Recent wastewater studies emphasize the use of modeling and statistical techniques in evaluating the treatment efficiency. Rahmat et al. (2022) and Młyński et al. (2020) have described hybrid and biodegradable mechanical configurations eliminating over 90% of contaminants in controlled conditions. However, a Sri Through a Lankan study,

Hemachandra and Pathiratne (2018) revealed frequent COD and TSS excesses, showing poor control and low automation rates compared to foreign systems. This indicates the gap between theoretical efficiency and real running reliability of centralised EPZ WWTPs.

2.2 Industrial Pollutant Characteristics and Treatment Challenges

Industrial effluents contain a diverse spectrum of pollutants based on the production process. Textile, rubber, and plastic industries usually release wastewater with relatively high levels of organic dyes, suspended solids, and synthetic chemicals (Atea et al., 2024; Raptopoulou et al., 2016). The composite loads are difficult to treat and less effective in centralized plants. It has been shown through local studies that Sri Lankan EPZs consistently exceed discharge limits for COD, BOD, and TSS Fonseka et al. (2017) pointing toward the need for industry-specific pretreatment before the centralized treatment.

2.3 Environmental and Health Implications

Untreated or improperly treated effluents cause oxygen depletion, eutrophication, and water and soil contamination over an extended period. Jin et al. (2017). Demonstrated toxic effects of industrial effluent on aquatic life, whereas Islam et al. (2016) demonstrated heavy metal buildup near discharge sites. Also in Sri Lanka, such exceedances continue to be reported by the Central Environmental Authority (CEA).

2.4 Operational Reliability and Research Gaps

Operational reliability of WWTP depends heavily on operator competence, maintenance, and continuous monitoring. Rahmat et al. (2022) have suggested the Wanterra Wastewater Quality Index (WWQI) for monitoring operational reliability, whereas Dewan (2024) has indicated that impacts of rainfall on effluent quality are poor compared to operational inefficiencies. Therefore, more research is crucial in quantifying pollutant variability and identifying variables influencing treatment effectiveness in EPZs.

3. METHODOLOGY

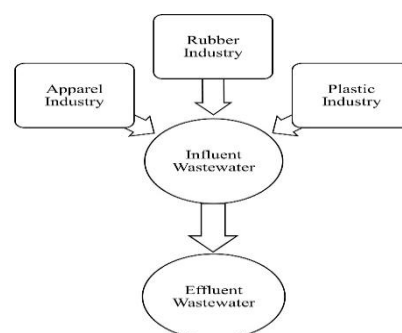


Figure 1: Schematic Diagram of the Wastewater Flow Observed for Data Collection

The study utilized a quantitative, comparative research design with secondary data of Horana and Seethawaka Export Processing Zones (EPZs) for 96 weeks. Five water quality parameters like; COD, BOD, TSS, pH, and temperature were recorded with rainfall data from the Department of Meteorology. Industries were categorized under apparel, rubber, and plastic for comparison. Cleaning, standardization, and tabulation of data were done using Microsoft Excel before statistical analysis using R software. Descriptive statistics, correlation analysis, multiple linear regression, and ANOVA

were conducted to find whether there existed any relationship between rainfall, industry type, and levels of pollutants. Treatment efficiency was approximated comparing influent and effluent concentrations to enable comparison of the pollutant removal rates with discharge standards of Central Environmental Authority (CEA). The whole procedure made sure that, the variation in the levels of the pollutants was handled systematically and therefore comparisons regularly across zones, types of industry, and environmental conditions were possible to be made.

Table 1: General Standards and Criteria for Industrial Effluent Discharge

Parameter	Tolerance Limit Values
Chemical Oxygen Demand (COD)	250 mg/L (max)
Biological Oxygen Demand (BOD)	30 mg/L (max)
Total Suspended Solids (TSS)	50 mg/L (max)
pH Levels	6.0 – 8.5
Temperature	Shall not exceed 40°C within 50m of the stream

(Source: General Standards and Criteria for the Discharge of Industrial Effluents into Inland Surface Waters)

4. RESULTS AND DISCUSSION

4.1 Temporal and Seasonal Patterns

Time-series plots of COD, BOD, TSS, pH, and temperature of Horana and Seethawaka showed no pronounced seasonal peaks or long-run periodicity. Autocorrelation and partial-

autocorrelation tests revealed weak persistence ($< \pm 0.25$) of all parameters for Horana, reflecting random fluctuation, while Seethawaka reflected moderate short-run dependence for TSS and BOD, reflecting mild temporal continuity in pollutant concentration. Overall, pollutant concentrations fluctuated non-systematically, i.e., exceedances are more subject to operational practice than climatic seasonality.

4.2 Industrial and Environmental Influences

Descriptive statistics indicated noticeable variation by industry and zones. pH averages fell within the acceptable range of 6–9, and temperatures averaged around 29°C. Parameters of pollutants indicated wide variation, but COD and TSS were most predominant in the case of Seethawaka garment industry (mean COD $\approx$ 319 mg/L; TSS $\approx$ 186 mg/L). The results indicate that disproportionately, some industrial processes are involved in organic as well as particulate contamination. A two-way ANOVA assessing simultaneous effects of industry type and pollutant type had no differences found ($p > 0.05$). This suggests that treatment performance variability is driven more by operational inconsistencies within the WWTPs than by industrial wastewater type. The graph uses distinct colours to clearly differentiate each industry category. Influent Water is shown in light blue, while effluent water appears in red. Rubber is represented by a green line, and apparel is illustrated in yellow. Plastic is displayed using a dark blue colour in each graph.

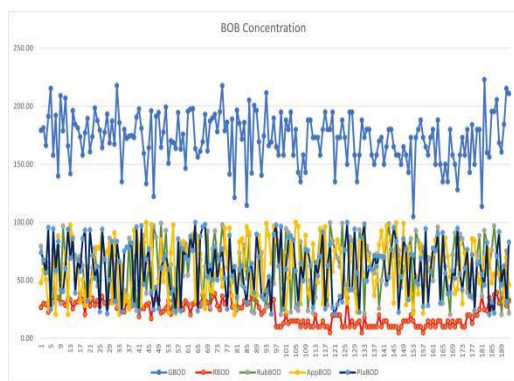


Figure 2: BOD Concentration

The plot of BOD concentration (Figure 2) indicates colossal differences across different industries. The highest BOD values are recorded in the Rubber industry, often in the range of 150–220 mg/L, with highest organic content in its effluent. Other industries such as; Apparel, and Plastic, indicate relatively lower BOD values, often below 100 mg/L. Inconsistency in all industries indicates periodic fluctuation in production processes and effluent production, perhaps due to batch operation or fluctuation in treatment efficiency.

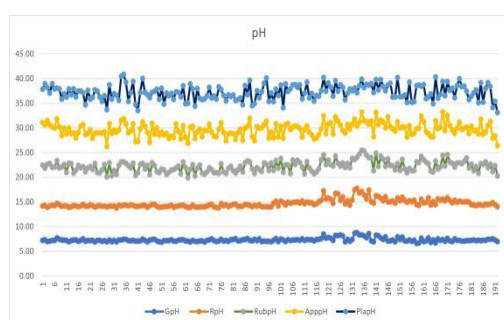


Figure 3: pH Variation

The pH profile (Figure 3) indicates that all the industries maintain relatively constant pH levels throughout the sampling period. Values are within acceptable limits due to effective neutralization processes in the wastewater treatment plants. Small fluctuations, perhaps due to differences in the

utilization of chemicals used in the manufacturing operations such as; dyeing or washing, are observed. Generally, pH is in good control compared to other parameters of water quality.

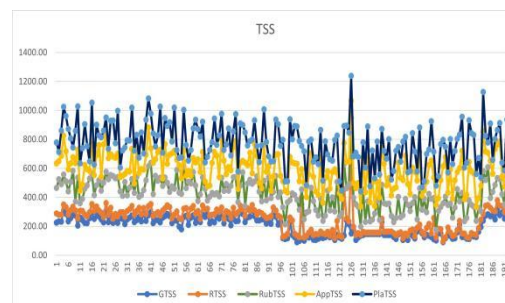


Figure 4: TSS Concentration

TSS plot (Figure 4) depicts large variability across the industries, with Plastic depicting maximum suspended solids concentration with repeated values more than 1000 mg/L. Apparel also depicts relatively higher TSS with moderately high values, whereas Rubber depicts relatively lower concentrations. Repeated spikes reflect inefficiencies in solid removal efficiency, which may be due to the change in production rates or inefficiency in sedimentation process in treatment processes. Generally, TSS is more variable than pH, and it indicates the effect of industrial operation and operating conditions.

4.3 Treatment Efficiency Comparison

Excellent removal efficiencies for suspended and organic matter were obtained in both plants (Horana: BOD 83%, COD 74%, TSS 80%; Seethawaka: BOD 91%, COD 83%, TSS 76%). Temperature and pH variations were negligible ($\approx 3\text{--}4\%$), indicating little impact on chemical or thermal equilibrium. Shapiro-Wilk normality test revealed non-normal distribution of data; therefore, a Wilcoxon rank-sum test was

applied. Results ($p = 0.84$) showed that there was no difference in the overall removal efficiency between the two EPZs.

5. CONCLUSION

The study demonstrated variability and treatment effectiveness of wastewater of Seethawaka and Horana Export Processing Zones using 96 weeks' data. Outcomes demonstrated that fluctuations in contaminants, COD, BOD, and TSS, are regulated primarily by process and treatment inefficiencies rather than rainfall or industrial category. Though mean removal efficiencies were acceptable and within acceptable limits, frequent exceedances over national discharge standards reveal inconsistent process control. The results point out that mere central wastewater treatment is not sufficient to ensure compliance and sustainability. Therefore, more stringent source-level pretreatment, continued monitoring, and adaptive operating management need to be enforced to enhance treatment reliability and maintain environmental protection in Sri Lanka's industrial estates.

REFERENCES

- Atea, R.F., Al-Saadi, R.J.M., Al-Baidhani, J.H., & Hassan, W.H. (2024). Evaluation of the effluent concentrations of Karbala wastewater treatment plant using reliability analysis. *Open Engineering*, 14(1). <https://doi.org/10.1515/eng-2024-0031>
- Dewan, A. (2024). Impact of Industrial Effluents on Dhaka Export Processing Zone's Water Bodies: A Comprehensive Analysis. *Cognizance Journal of Multidisciplinary Studies*, 4(6), 235 - 248. <https://doi.org/10.47760/cognizance.2024.v04i06.013>
- Fonseka, W.A.K.S., Himanujahn, S., Dissanayake, H.H., & Athapattu, B.C.L. (n.d.). *A study on industrial pollutants discharged into the downstream of the Kelani River*.
- Hemachandra, C.K., & Pathiratne, A. (2018). Assessing toxicity of two industrial zone effluents reaching Kelani River, Sri Lanka. *Journal of the National Science Foundation of Sri Lanka*, 46(4), 539 - 546. <https://doi.org/10.4038/JNSFSR.V46I4.8629>
- Islam, M.S., Sultana, A., Sultana, M.S., Shammi, M., & Uddin, M.K. (2016). Surface Water Pollution around Dhaka Export Processing Zone and Its Impacts on Surrounding Aquatic Environment. *Journal of Scientific Research*, 8(3), 413 - 425. <https://doi.org/10.3329/jsr.v8i3.27819>
- Jin, Z., Zhang, X., Li, J., Yang, F., Kong, D., Wei, R., Huang, K., & Zhou, B. (2017). Impact of wastewater treatment plant effluent on an urban river. *Journal of Freshwater Ecology*, 32(1), 697 - 710. <https://doi.org/10.1080/02705060.2017.1394917>
- Młyński, D., Młyńska, A., Chmielowski, K., & Pawełek, J. (2020). Investigation of the wastewater treatment plant processes efficiency using statistical tools. *Sustainability (Switzerland)*, 12(24), 1 - 16. <https://doi.org/10.3390/su122410522>
- Rahmat, S., Altowayti, W.A.H., Othman, N., Asharuddin, S. M., Saeed, F., Basurra, S., Eisa, T.A.E., & Shahir, S. (2022). Prediction of Wastewater Treatment Plant Performance Using Multivariate Statistical Analysis: A Case Study of a Regional Sewage Treatment Plant in Melaka, Malaysia. *Water (Switzerland)*, 14(20). <https://doi.org/10.3390/w14203297>
- Raptopoulou, C., Palasantza, P.A., Mitrakas, M., Kalaitzidou, K., Tolkou, A., & Zouboulis, A. (2016). Statistical variation of nutrient concentrations and biological removal efficiency of a wastewater treatment plant. In *Water Utility Journal*, 14.



Modelling Temporal Trends and Forecasting Tourists related Violent Crimes in Sri Lanka: A Data Driven Approach

Lakmini KRD¹, Habaragoda BS²

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University Sri Lanka*
dinushalakmini2000@gmail.com¹

ABSTRACT

Tourism is one of the fastest-growing industries globally, and it is a primary contributor to the economies of most Asian countries. The tourism industry directly raises employment and income levels and is one of the most important industries for fostering growth. Tourist-related crimes can severely damage a destination's reputation, discourage future visits, and lead to significant revenue losses for the industry. In order to identify risks, inform policy decisions, and improve preventive measures, it is crucial to analyze crime data pertaining to tourism. This study investigates tourist-related violent crimes in Sri Lanka from 2020 to 2024 using a time series approach to examine temporal patterns and forecast future crime trends. Using monthly crime data obtained from the Sri Lanka Tourist Police, ARIMA models were developed to identify and forecast patterns in violent crimes, including sexual harassment, aggravated assault, robbery, murder, and intimidation. The time series analysis revealed seasonal and non-seasonal fluctuations and enabled predictions of future crime trends in favour of proactive safety planning for the respective authorities. According to the findings of this study, a new security strategy can be developed to mitigate these trends in society.

KEYWORDS: ARIMA, Forecasting, Time Series, Violent Crime, Tourism

1. INTRODUCTION

Sri Lanka is well known for its natural beauty and cultural heritage. Tourism is one of the major contributors as well as the cornerstone of economic development which is significantly affecting foreign exchange earnings, and employment (Nagendrakumar et al., 2021). However, the industry has recently faced several challenges due to the global pandemic (COVID - 19), the Easter Sunday attack and domestic challenges such as political instability (Nagendrakumar et al., 2021). These challenges have emphasised the need to develop accurate forecasting models that enable the industry to respond to uncertainties effectively and prepare for recovery. Crimes can severely damage the industry. Therefore, maintaining a low crime rate is crucial for the development of the industry in Sri Lanka. Ensuring a safe

environment for tourists not only enhances their overall travel experiences but also strengthens the reputation of the country as the best tourist destination in the world (Farhan & Susanti, 2024). Although most of the studies are based on the relationship between tourist arrivals and crime, there are very limited studies focused specifically on tourist- related crimes. The main objective of this study is to develop ARIMA-based time series models for tourist related violent crimes using monthly data from 2020 to 2024, and to forecast future crime trends for proactive safety planning. Further, this research aims to provide useful insights for specific authorities, policy makers and tourism industry to enhance tourist safety.

2. LITERATURE REVIEW

2.1 *Relationship between tourism and crime*

Though many studies have reviewed the determinants of international tourism demand, very few researches have analysed the tourist-related crimes. Tourism is usually seen as a positive phenomenon. Crime against tourists can be explained using the Routine Activity Theory (RAT) proposed by Cohen and Felson (1979). According to this theory, crime is more likely to occur when three elements intersect at the same time and place: a motivated offender, a suitable target, and the absence of capable guardianship. Another relevant theory is the Tourism Crime Cycle proposed by Kandy (2003). According to that theory, there is a lower crime rate in places designated for families with children or for elderly people compared to places for recreation for younger people. Many studies confirmed the mutual relationship between tourism and crime. Moyo & Ziramba (2013) investigated the impact of crime indicators on tourist inflows to South Africa using monthly data from March 2003 to April 2011. And they used the bound tests ADRL (Autoregressive Distributed Lag) approach for the analysis. It has been confirmed that crime as a whole is not a significant determinant of inflows, but kidnappings, car hijackings, murder and sexual offences harm tourist inflows to the country, both in the short and the long run.

2.2 *Analysis of tourism related crime in Sri Lanka*

Even though Sri Lanka is well known as one of the popular tourism destinations all over the world, there is limited research on crimes affecting

tourists. The majority of existing crime studies in Sri Lanka focus on general criminal behaviour, lacking a specific focus on tourist-related incidents. Rathnayake and Wijesundara (2015) have investigated the tourists' perception on harassment in Sri Lanka with special reference to Mount Lavinia. The research implied that there is a significant positive association between awareness in social life and tourists' perception on harassment. Arachchi (2020) examined a study based on the effect of tourist's harassments on the destination image and repeat visit intention of tourists with a special reference to the Hikkaduwa beach site. He concluded that tourists who stayed at Hikkaduwa frequently faced harassment from individuals such as beach boys, pesters, and thieves. Out of different types of violent crimes, they have identified that only verbal harassment has a significant impact on the destination image of Sri Lanka. The above literature review suggests that there is a dearth of systematic datadriven approaches to find tourism crimerelevant studies, especially temporal analysis, in Sri Lanka. This research aims to fill that gap by conducting a data driven temporal analysis of tourist related violent crimes in Sri Lanka from 2020 to 2024.

3. METHODOLOGY

3.1 *Data collection and Scope*

Secondary data was collected from the Sri Lankan Tourist Police to identify crime records on complaints related to tourists. The dataset consists of monthly data spanning from 2020 to 2024. The raw dataset includes the number of crime cases for each type of violent crime,

categorised by month, covering the period.

3.2 Procedures and Techniques Used

Descriptive statistics were used to identify and summarise the distribution and frequency of various types of touristrelated crimes. Bar charts and line graphs are used to visualise the trend over time and the distributions by crime type.

The main objective of applying time series techniques is to capture both temporal and seasonal variations in different types of violent crimes, such as sexual crimes, aggravated assault, etc. The Augmented Dickey-Fuller (ADF) test was first applied to examine the stationarity of the time series for each crime type. After confirming stationarity, trend plots, autocorrelation and partial autocorrelation were performed in order to identify the long-term trend and the cyclic behaviour. ARIMA models have been fitted to crime data, which were stationary (Prashant, 2020). The models have been tested for diagnostic checking in order to find the most adequate model. Diagnostic checking consists of several steps, such as significance of the parameters, randomness of the residuals, parameter redundancy and normality of residuals. The best model is fitted using these techniques. Forecasting accuracy is tested using MAPE (Mean Absolute Percentage Error) during the model validation. R software was used for generating descriptive statistics, whereas Minitab was employed to analyse temporal patterns and trends.

4. RESULTS AND DISCUSSION

4.1 Descriptive Statistics Analysis

To examine the patterns and variability of the number of crimes from

2020-2024, descriptive statistics were calculated for each violent crime type: aggravated assault, robbery, sexual harassment and intimidation. The aim is to find which crime types frequently occur, how incidents vary over time and the distribution of each crime type.

Table 1: Summary of the Descriptive Statistics

Crime Type	Mean	Median	Maximu	Minimu	Standard deviation	Range
Violent	52.4	55	80	6	31.4	80
Aggravated Assault	11.8	14	17	1	6.30	16
Robbery	6.4	2	21	1	8.38	20
Sexual Harassments	32.6	38	61	3	21.6	58
Intimidation	1.75	1.5	61	3	21.6	58

(Source: Author compile)

Among specific violent crime types, sexual harassment had the highest average yearly counts.

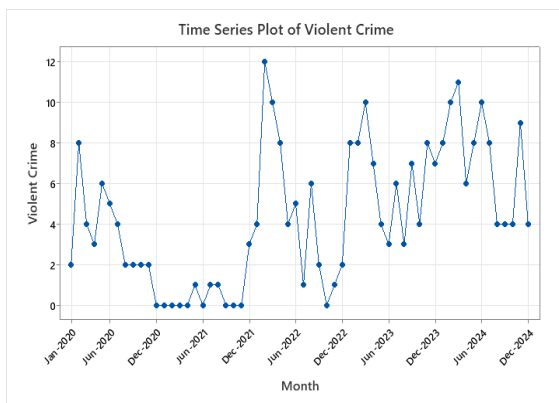
4.2 Time Series Analysis

The ADF test was used to check whether the data set is stationary. The null hypothesis (Ho) states that the series has a unit root. (i.e., it is not stationary), while the alternative hypothesis (H1) states that the series is stationary (i.e., does not have a unit root). If the test statistic is less than the 5% critical value ($t > 5\%$), there is not enough evidence to reject the null hypothesis and vice versa. The non-stationary robbery crime series has become stationary after firstorder differencing is done.

Table 2: ADF Test Results

Crime Type	Test Statistic	Critical Value	P value	Stationary or not
Violent	-3.47	-2.91	0.00	Stationary
Aggravated Assault	-5.41	-2.91	0.00	Stationary
Robbery	-1.57	-2.91	0.03	Non-stationary
Sexual Harassments	-4.53	-2.91	0.00	Stationary

(Source: Author compile)

**Figure 2:** Time Series Plot for Violent Crimes.

When differencing is combined with autoregressive and moving average components, the resulting model is referred to as a non-seasonal ARIMA model. The general form of the ARIMA(p,d,q) model can be expressed as $y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$ where y_t is the differenced series. This is an ARIMA (p, d, q) model, where p = order of the autoregressive part, d = degree of first differencing involved and q = order of the moving average part. ARIMA (0,0,1) (0,0,0)₁₂ and ARIMA (1,0,0) (0,0,0)₁₂ are two adequate models for the violent crimes. To identify the bestfitted model, MAPE is calculated.

ARIMA (1,0,0) (0,0,0)₁₂ has the highest accuracy value with 69.0%, which becomes the best fitted model, while the model ARIMA (0,0,1) (0,0,0)₁₂ has 64.9% accuracy. The model ARIMA (1,0,0) (0,0,0)₁₂ has very low correlation, which is -0.0186 suggesting there is no parameter redundancy between the AR1 term and the constant term. The modified box-pierce (Ljung-Box) chi-square statistic shows that all the p values are greater than 0.05 at lags 12, 24, 36 and 48 respectively. This indicates that the residuals are random. The normality of residuals is checked, and all the data are scattered approximately within the straight line, which indicates the normality. Since parameters are significant, residuals are random, normally distributed, and there is no parameter redundancy, the model ARIMA (1,0,0) (0,0,0)₁₂ is adequate. This model is used to forecast crime incidents for the year 2025. The forecast indicated the year 2025, the occurrence of violent crimes will be approximately about 4 in every month.

5. CONCLUSION AND RECOMMENDATIONS

In this study, forecasting models were developed using the Box-Jenkins ARIMA approach, which is considered to be an appropriate method for modelling and forecasting violent crimes.

Table 3: ARIMA models for different violent crimes

Crime Type	Model
Violent crimes	ARIMA (1,0,0) (0,0,0) ₁₂
Aggravated Assault	ARIMA (1,0,0) (0,0,0) ₁₂
Robbery	ARIMA (1,1,0) (0,0,0) ₁₂
Sexual Harassments	ARIMA (1,0,0) (0,0,0) ₁₂

(Source: Author compile)

Forecasted values for 2025 suggest an average of 1 aggravated assault, 3 robberies, and 2 no. of sexual crimes per month affecting tourists. This study forecasts each violent crime type separately, providing clearer patterns that help policymakers prioritise specific offences and allocate resources more effectively. The above models can be used to forecast future crimes affecting tourists in Sri Lanka. Having an appropriate model for forecasting will aid in for better planning and devising policies for the tourism industry. For crime types where ARIMA models were fitted, significant discrepancies were observed between actual and forecasted values during the validation phase. One of the contributing factors for this issue is that the time frame of the data set is expanded from 2020 to 2024, and it is a period significantly impacted by the COVID-19 pandemic (Chinazzi et al., 2020). It is recommended to extend the dataset to include more historical data to 8 -10 years, which would enhance the validity of the results.

As a significant limitation of this study, it can be noted that the sample size is relatively small. The results would be much more accurate if a considerable sample size were used in the analysis, and a good comparison could be built if the analysis were conducted for a longer time period. Future research can incorporate with alternative approaches, such as, seasonal decomposition models, state-space models, or modern machine learning techniques, to better capture the complex patterns in crime data.

REFERENCES

- Arachchi, R.S.S.W. (2020). Effect of tourists' harassments to the destination image and repeat visit intention of tourists: a case study of Sri Lanaka. *Journal of Management*, 15(1), 48.
- Chinazzi, M., Davis, J.T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Piontti, A.P.Y., Mu, K., Rossi, L., Sun, K., Viboud, C., Xiong, X., Yu, H., Halloran, M.E., Longini, I.M., & Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, 368(6489), 395 - 400.
- Cohen, L. Felson, M. 1979. Social Change and Crime Rate Trends: A Routine Activities Approach. *American Sociological Review*, 44(4), 588 - 608.
- Farhan, M., & Susanti, H. (2024). *The effect of crime on tourism in Indonesia*. UI Scholars Hub.
- Kandy, J. 2003 Crime. In: J. Jenkins, J. Pigram (eds.), *Encyclopedia of Leisure and Outdoor Recreation*. Taylor & Francis, 86-88.
- Moyo, B., & Ziramba, E. (2013). The impact of crime on inbound tourism to South Africa. *African Security Review*, 22(1), 4 - 18.
- Nagendrakumar, N., SLIIT Business School, Lokeshwara, A.A., Gunawardana, S.A.D.C.K., Kodikara, U.P., Rajapaksha, R.W.N.H., & Rathnayake, K.R.M.C.S. (2021). Modelling and Forecasting Tourist Arrivals in Sri Lanka. In *SLIIT Business Review* (pp. 95 - 120).
- Rathnayake, R., Wijesundara, W., & Department of Public Administration, Faculty of Management, Uva Wellassa University, Badulla, Sri Lanka. (2015). Tourist perception on harassments in Sri Lanka (With special reference to Mount Lavinia). *International Journal of Scientific and Research Publications*, 1.



A Comprehensive Analysis of Trends and Patterns in Sri Lankan Tourism: Pre-Post COVID Growth and Visitor Demographics

Sameera LN¹, Habaragoda BS²

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka*
ns.liyanwala@gmail.com¹

ABSTRACT

This study investigates the evolving trends and demographic patterns in Sri Lankan tourism, focusing on growth dynamics before and after the COVID-19 pandemic and changes in visitor profiles. Tourism is a vital pillar of Sri Lanka's economy, contributing significantly to the foreign-exchange earnings and employment. However, the sector's high sensitivity to global crises, such as the COVID-19 pandemic, has reshaped both arrival patterns and traveler characteristics. The objectives of this study are twofold: first, to analyze visitor demographics, including age, gender, regional origins, top source countries, purpose of visit, duration of stay, and occupation profile using descriptive statistical techniques; and second, to compare tourism growth trends before and after the pandemic using quadratic regression models. Secondary data spanning 2000–2024 were obtained from the Sri Lanka Tourism Development Authority (SLTDA) and analyzed using R and Excel. Findings indicate that post-COVID tourism is experiencing accelerated recovery, driven by pentup demand and government-led revival initiatives, contrasting with the steady but slower pre-pandemic growth. The demographic analysis revealed an increase in older and female travelers, diversification of source markets, and emerging trends in travel purpose and duration of stay. These results highlight the resilience of Sri Lanka's tourism industry and underscore the importance of adaptive marketing strategies, demographic-focused policy planning, and sustainable tourism development to strengthen future sectoral performance.

KEYWORDS: Post-COVID Recovery, Tourism growth, Visitor Demographics

1. INTRODUCTION

Tourism is a key driver of Sri Lanka's economy, contributing significantly to foreign exchange earnings, employment, and infrastructure development. The country's diverse attractions, from natural and cultural heritage to wellness tourism, position it as a leading destination in South Asia. However, the sector is highly vulnerable to disruptions such as pandemics, political instability, and global economic fluctuations. The COVID-19 pandemic caused an unprecedented collapse in international travel, severely affecting Sri Lanka's tourist arrivals and market composition.

Understanding how the sector has recovered and how visitor characteristics have changed is essential for effective policy and marketing decisions. Examining visitor profiles before and after COVID-19 is particularly important because demographic shifts shape travel motivations, spending behaviour, and destination preferences. Identifying these changes enables stakeholders to design targeted promotions and refine tourism products to maintain global competitiveness.

This study addresses this need by analyzing demographic patterns and modelling tourism growth trends across the pre- and post-COVID periods. Using descriptive statistics and quadratic regression, it provides empirical evidence

of both structural and temporal changes. The study integrates demographic profiling with trend modelling, offering a comprehensive understanding of tourism dynamics to support data-driven recovery planning and long-term development. Accordingly, the study is guided by the following objectives:

Objective 1: To analyze visitor demographic patterns, including age, gender, regional origin, purpose of visit, length of stay, and occupation, using descriptive statistical techniques.

Objective 2: To compare pre-COVID and post-COVID tourism growth using quadratic regression models to identify differences in recovery trajectories and long-term trends.

2. LITERATURE REVIEW

Tourism is widely recognized as one of the most crisis-sensitive economic sectors, with recovery patterns varying depending on destination characteristics and the type of disruption. The COVID-19 pandemic caused an unprecedented collapse in global travel, highlighting the importance of resilience and diversification in tourism policy (WTTC, 2023). According to Ritchie and Jiang (2019), adaptive crisis-management systems and market reorientation are central to ensuring recovery. In Sri Lanka, measures such as visa liberalization, safety certification programs, and international promotional campaigns have significantly contributed to post-pandemic revival (First Capital, 2025).

Global tourism trends have also shifted due to demographic changes. Research identifies a growing segment of older and female travelers with increasing disposable income and demand for safe

and wellness-oriented experiences (Bak & Szczecińska, 2020; Richards, 2018). These changes require more inclusive and flexible tourism products. In the Sri Lankan context, Rajapakse (2024) has reported similar trends, noting the rising participation of older and female visitors. Market diversification has also become essential, as Sri Lanka's historical reliance on the Asia-Pacific and European markets exposes the industry to regional shocks.

Empirical evidence shows that sustainable tourism recovery relies on data-driven analysis of macro trends and traveller behaviour (Ritchie & Jiang, 2019; WTTC, 2023). Combining demographic profiling with quantitative trend modelling offers a more complete understanding of tourism dynamics. This study adds to the existing literature by examining demographic shifts alongside comparative pre- and post-pandemic growth patterns. The dual analytical approach provides evidence to guide policymaking, strengthen crisis preparedness, and enhance the long-term competitiveness and resilience of Sri Lanka's tourism sector.

3. METHODOLOGY

The current research employed a quantitative approach using secondary data from the Sri Lanka Tourism Development Authority (SLTDA). The dataset includes 25 annual observations of tourist arrivals from 2000 to 2024 and 15 years of monthly demographic records from 2010 to 2024, totaling 180 monthly data points. These consist of annual arrival volumes supported by monthly demographic statistics.

The variables analyzed include total tourist arrivals; age groups (3 - 19, 20 - 29, 30 - 39, 40 - 49, 50 - 59, 60+); gender (male, female); regional origin (America, Africa, Asia & Pacific, Europe, Middle East); top ten source markets (India, UK, Maldives, Germany, Australia, France, USA, Russia, China, Canada); purpose of visit (Pleasure, Business, Visiting Friends & Relations, Convention & Meeting, Religious & Cultural); occupational categories (Businessmen, Professionals, Executives, Scientists & Technicians, Educationists, Other Occupations, No Occupation); and duration of stay (1 - 3, 4 - 7, 8 - 14, 15 - 21, 22 - 30, 31+ nights).

To achieve Objective 1, descriptive statistics and visualizations (line and bar charts) were used to examine changes in visitor composition. This analysis identified shifts such as the increase of senior travelers, narrowing of gender gaps, and the growth of market diversification.

For Objective 2, quadratic regression modelling was employed to compare pre-COVID and post-COVID growth trends in tourist arrivals using the model equation is:

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 \quad (1)$$

where Y_t represents annual tourist arrivals and t denotes time (years since the base year). Model performance was assessed using R^2 , p -values, and residual diagnostics.

Preprocessing and visualization of data were carried out by Excel, while regression estimation and validation were performed using R software. Ethical compliance was ensured solely through the use of publicly available government sources of data. The methodology framework ensures replicability,

transparency, and consonance with the twin aims of demographic profiling and growth trend comparison of the study.

4. RESULTS AND DISCUSSION

4.1 Visitor Demographic Analysis (2010 – 2024)

The demographic analysis of inbound tourists to Sri Lanka from 2010 to 2024 reveals noticeable shifts across key characteristics, as shown in Figures from 1 to 7.

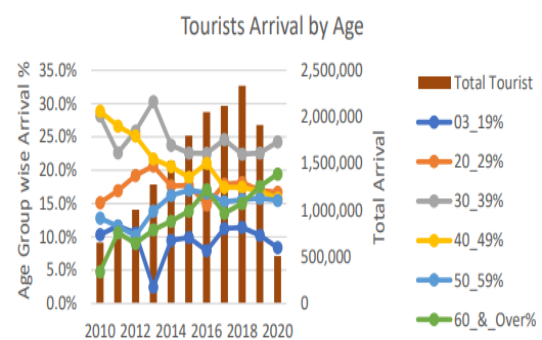


Figure 1: Age Profile (2010 - 2020)

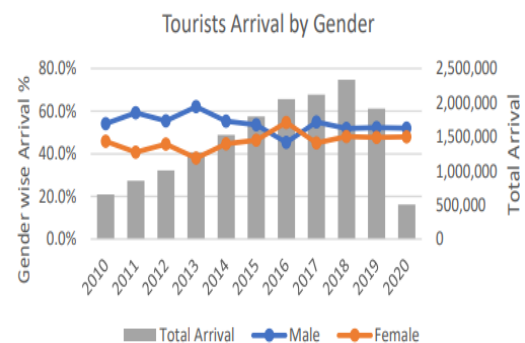


Figure 2: Gender Distribution (2010 - 2020)

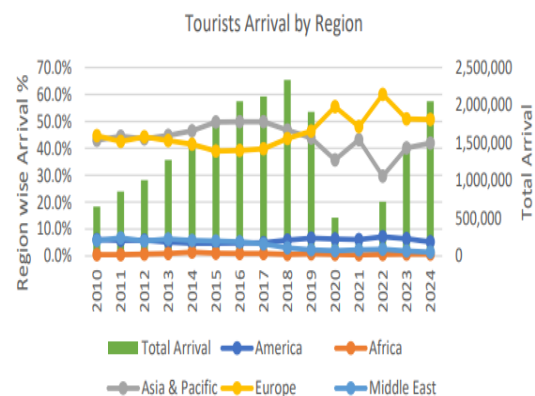


Figure 3: Regional Distribution (2010 - 2024)

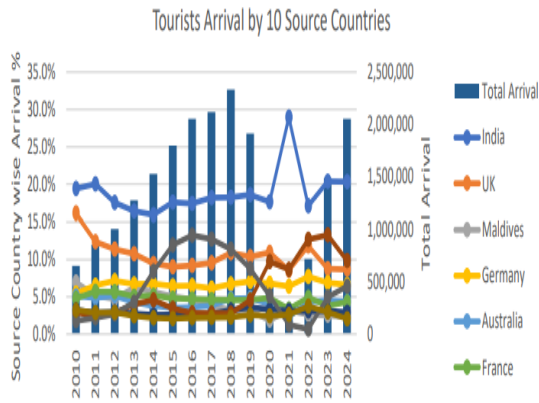


Figure 4: Top 10 Countries (2010 - 2024)



Figure 5: Purpose of Visit (2010 - 2020)



Figure 6: Duration of Stay (2010 - 2020)

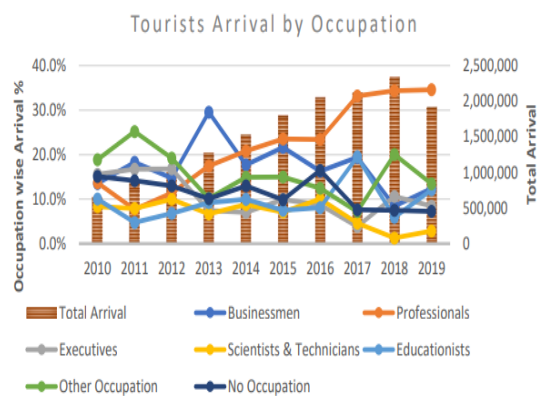


Figure 7: Occupation Profile (2010 - 2019)

According to Figure 1, the highest arrivals are in the 30 - 49 age group, and arrivals within the 60+ group have steadily grown, showing growing interest among older visitors. Figure 2 shows that female arrivals have been higher than male arrivals since 2016, suggesting increasing safety and welcome. Figure 3 shows that Europe and Asia remain the prime regions, and Figure 4 shows India, the UK, China, Germany, and France as Sri Lanka's top source countries. Figure 5 shows that leisure remains the prime reason for visit, with a moderate bounce back in business and visiting friends and classes. Figure 6 shows that the length of stay, which fell during COVID - 19, has come back to settle at 10 - 12 days. professionals and retirees from an increasing share of visitors shows a trend towards experience-based, higher-income tourists (Figure 7). Together, descriptive analysis reflect Sri Lanka's post COVID diversification of tourism and potential for enhanced targeted promotion and sustainable product delivery.

4.2 Tourism Growth Trends: Before and After COVID-19

Tourist arrivals in Sri Lanka increased steadily from 336,000 in 2000 to over 2.3 million in 2018, reflecting strong postwar recovery and destination promotion. As shown in Figure 8, growth was stable until 2018, followed by a slight drop in 2019 due to the Easter Attacks. The quadratic model indicated significant longterm growth with good fit ($R^2 = 0.82$).

Table 1: Pre-COVID Growth Model Summary

Equation	R ²	Trend
$Y_t = 35,725 - 249t + 4.10t^2$	0.82	Steady growth

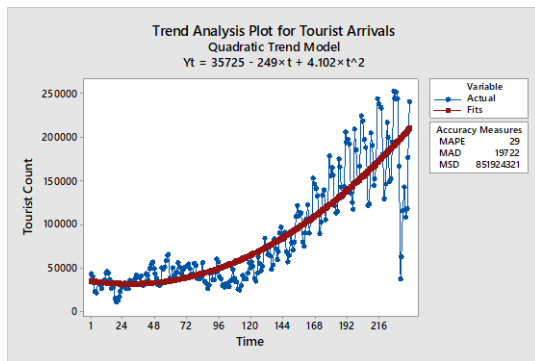


Figure 8: Trend Analysis for Pre COVID 19 Growth

After the 2020 collapse, arrivals rebounded from 194,000 in 2021 to over 1.4 million by 2024. Figure 9 shows the rapid post-pandemic recovery driven by global travel resurgence and national revival programs. The regression model produced a strong fit ($R^2 = 0.91$), confirming accelerated growth.

Table 2: Post-COVID Growth Model Summary

Equation	R^2	Trend
$Y_t = 44,176 + 3,647t + 14.9t^2$	0.91	Rapid recovery

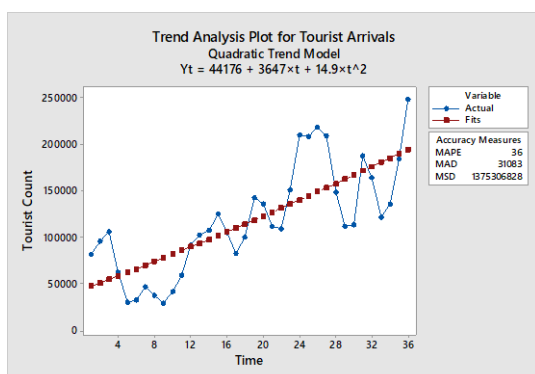


Figure 9: Trend Analysis for Post COVID 19 Growth

Table 3 illustrates that pre-COVID growth was gradual and stabilizing; the post COVID - 19 period demonstrates rapid acceleration and a substantially higher baseline of arrivals. Figure 10 presents two growth curves representing Sri Lanka's tourist arrivals before and after COVID - 19. The pre-COVID curve

shows a slow and steady increase, reflecting gradual growth that stabilized over time. In contrast, the post-COVID curve starts at a higher level and rises sharply, illustrating a strong and accelerating recovery.

Table 3: Comparative analysis before and after COVID-19

Aspect	Before COVID	After COVID
Quadratic Growth	+4.1 (slow)	+14.9 (fast)
Starting Point (t=0)	35,725	44,176
5-Year Projection	~35,000	~62,669
Recovery Status	Gradual	Strong & fast

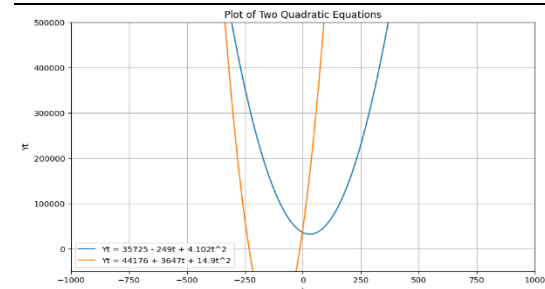


Figure 10: Quadratic growth models - before & after COVID - 19

5. CONCLUSION

This study provided a comprehensive examination of Sri Lanka's tourism sector by analyzing visitor demographics and growth trends before and after the COVID-19 pandemic. The demographic results revealed key structural shifts, including a rising share of senior and female travelers and greater diversification of source markets. These changes highlight the need for tourism offerings that respond to evolving traveler expectations, particularly in the areas of safety, wellness, accessibility, and culturally authentic experiences.

The growth trend comparison showed that while pre-pandemic tourism followed a steady long-term upward path, the post-

pandemic period has experienced a faster and more pronounced recovery. This rebound reflects the effectiveness of national revival measures and the global resurgence of leisure travel. At the same time, the regression analysis indicates some volatility in post-COVID growth, underscoring the need for strengthened resilience planning.

Based on these findings, several recommendations are proposed. Sri Lanka should enhance inclusive tourism infrastructure to accommodate the increasing presence of senior and female travelers. Diversification in regional arrivals suggests opportunities for targeted promotions in Middle Eastern and Asia-Pacific markets. The rise in medium- and long-stay visitors supports extending visa options and promoting long-stay packages. Finally, observed volatility highlights the importance of crisis preparedness and market diversification to ensure long-term stability.

REFERENCES

- Bąk, I., & Szczecińska, B. (2020). Global demographic trends and effects on tourism. *European Research Studies Journal*, 23(4), 571–585.
- First Capital Holdings PLC. (2025). *First Capital Equity Fund Annual Report 2024/25*. First Capital.
- Rajapakse, V., Jayasuriya, N., Senadheera, S., & Senevirathne, N. (2025). An investigation on the crisis management practices of boutique hotels and villas in Sri Lanka: Lessons from the COVID-19 era. *ASEAN Journal on Hospitality and Tourism*, 23(1), 48 - 65.
- Richards, G. (2018). Cultural tourism: A review of recent research and trends. *Journal of Hospitality and Tourism Management*, 36, 12 - 21.
- Ritchie, B. W., & Jiang, Y. (2019). A review of research on tourism risk, crisis,

and disaster management: Launching the annals of tourism research curated collection on tourism risk, crisis and disaster management. *Annals of Tourism Research*, 79, 102812.

- World Travel and Tourism Council. (2023). *Economic Impact Report 2023*. WTTC.



Effect of Social Media Engagement and Trust in shaping Online Purchase Intention: A Study of Sri Lankan Consumers using MOA Framework

Karunathilaka IPKA¹, Wattegama EJ²

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka*
kasunkarunathilaka2k@gmail.com¹

ABSTRACT

The growing engagement with social media has significantly reshaped consumers' interactions with brands and formation of their online purchase intentions. In developing markets such as Sri Lanka, social media presents considerable opportunities for e-commerce growth; however, consumer participation remains constrained by low trust, limited digital competence, and high perceived risk. Grounded in the Motivation–Opportunity–Ability (MOA) framework, this study examines how consumers' motivation, opportunity, and ability influence social media engagement and, subsequently, overall online purchase intentions, without limiting the analysis to a specific product category. The study also investigates the moderating role of trust in online vendor advertisements. Data were collected from a convenience sample of 337 Sri Lankan social media users with purchasing power and analyzed using SPSS and Hayes' PROCESS macro. The results reveal that consumers' motivation and ability significantly enhance their social media engagement, which in turn positively influences online purchase intentions. Although trust in online vendor advertisements does not moderate this relationship, it exerts a direct effect on consumers' online purchase intentions. The findings emphasize the importance of psychological readiness and digital competence of the consumers, rather than mere access, in shaping their social media engagement and online purchase intentions. The study offers practical implications for marketers and policymakers seeking to strengthen trust-building mechanisms and digital empowerment initiatives to support e-commerce growth in Sri Lankan market.

KEYWORDS: MOA Framework, Purchase Intention, Social Media Engagement, Trust

1. INTRODUCTION

Social media has become a fundamental channel for consumer–brand interaction, providing unprecedented access to information, community, and commerce. The online social media platforms, such as, Facebook, Instagram, TikTok, and so on are not merely spaces for entertainment but are central to how modern consumers evaluate and purchase products. In the Sri Lankan context, although the number of active social media users exceeds 08 million, online purchase adoption remains limited due to concerns, such as, transaction security and the reliability of online vendors (Commonwealth Secretariat, 2020). To

understand why some social media users convert their engagement into their purchasing behavior for some products while others do not, this study adopts the Motivation, Opportunity, and Ability (MOA) framework, which was introduced by MacInnis and Jaworski (1989). This MOA model posits that an individual's action depends on his / her willingness (motivation), capacity (ability), and enabling environment (opportunity). In digital contexts, this framework helps explain how consumers' internal drives and external conditions shape their online engagement. However, consumers' social media engagement alone may not guarantee their purchase intentions unless it is accompanied by trust, which is a

critical factor in e-commerce (Pavlou, 2003). Therefore, the present study integrates trust into the MOA framework to examine how consumers' motivation, opportunity, and ability influence their social media engagement, how such engagement mediates the relationship between the MOA factors and online purchase intentions, and how trust in online vendor advertisements moderates this relationship. Moreover, the study aims to validate the MOA framework in the Sri Lankan context and to contribute to filling the literature gap in the domain of e-commerce research.

2. LITERATURE REVIEW

The MOA framework offers a holistic view of consumer behavior, emphasizing that an individual must be motivated, capable, and have the opportunity to act (MacInnis, 1991). In social media engagement and internet-use environments, motivation reflects the psychological interest or willingness to engage with social media. According to Brodie et al. (2011), motivation includes both hedonic motives (e.g., enjoyment) and utilitarian motives (e.g., information seeking). Opportunity represents situational enablers such as internet access, time availability, and platform usability (Briliana & Mursito, 2017). Meanwhile, ability encompasses digital literacy, familiarity with technology, and prior experience that enable users to engage meaningfully with social media. Social media engagement refers to the cognitive, emotional, and behavioral participation of consumers in online communities and brand-related content (Dessart, 2017). It has been widely recognized as a mediating mechanism linking antecedent factors such as

motivation and ability to behavioral outcomes, including loyalty, advocacy, and purchase intention (Hollebeek et al., 2014). Trust also plays a pivotal role in e-commerce by shaping online purchase intentions and reducing consumers' perceived risk and uncertainty when purchasing products online (Mou & Benyoucef, 2021). In social commerce, trust relates to beliefs in the reliability of social media platforms, peer-based information sources, such as, 3Rs (reviews, ratings, and recommendations) online vendors and their advertisements, and payment systems. In this study, trust specifically refers to consumers' confidence in online vendors, including their reliability, honesty in advertising, advertising practices, platform credibility, and payment methods. By incorporating trust into the MOA framework, this study enables a deeper understanding of how confidence facilitates the conversion of social media engagement into online purchase intention. In the Sri Lankan context, the interaction between motivation and opportunity operates differently due to structural challenges such as the digital divide, limited connectivity, and affordability constraints, which restrict equal access to online platforms (Fernando & Samaranayake, 2022).

Based on this theoretical foundation, a conceptual framework (Figure 1) and corresponding hypotheses were developed to achieve the study objectives.

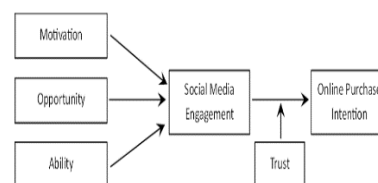


Figure 1: Conceptual Framework

H₁: Motivation positively influences social media engagement.

H₂: Opportunity positively influences social media engagement.

H₃: Ability positively influences social media engagement.

H₄: Social media engagement positively influences the online purchase intention.

H_{5a}: Engagement mediates the effect of motivation on the online purchase intention.

H_{5b}: Engagement mediates the effect of opportunity on the online purchase intention.

H_{5c}: Engagement mediates the effect of ability on the online purchase intention.

H₆: Trust on the advertisements that are posted by the online vendors on social media platforms moderates the relationship between social media engagement and online purchase intention.

3. METHODOLOGY

A quantitative cross-sectional survey approach was employed, and the sample comprised of conveniently selected 337 no. of Sri Lankan individuals who were social media users, were above 18 years old, and had independent purchasing power. A structured questionnaire was used to collect data on six constructs: motivation, opportunity, ability, social media engagement, trust, and online purchase intention using a five-point Likert scale. The questions were adopted from Briliana and Mursito (2017), Hollebeek et al. (2014), and Pavlou (2003). The questionnaires were distributed to the respondents via social media platforms.

The reliability testing yielded the Cronbach's alpha values between 0.70

and 0.89 for all six constructs, indicating satisfactory internal consistency. The responses were analyzed using SPSS 28.0 and Hayes' PROCESS macro tool (Model 3 for mediation and Model 4 for moderation).

4. RESULTS AND DISCUSSION

4.1 MOA Factors and Social Media Engagement - Model 01

The multiple regression analysis was carried out to assess the impact of motivation, opportunity, and ability on social media engagement. The overall model was significant ($F(3,333) = 79.51$, $p < 0.001$, $R^2 = 0.417$) and showed that 41.7% of the variance in social media engagement was explained by MOA factors.

Table 1: Regression Output of Model 01

Predictor	β	t-value	Sig.
Motivation	0.300	4.680	0.000
Opportunity	-0.153	-2.065	0.04
Ability	0.532	7.915	0.000

(Independent Variable = Social Media Engagement)

According to Table 1, the results demonstrated that only the motivation and ability have significantly enhanced the social media engagement of the consumers (Accepted H₁ and H₂), whereas the opportunity was not (Rejected H₃). The unexpected negative coefficient for opportunity suggests that while greater Internet access and time availability exist, users may divert their attention to other online activities, such as, entertainment, social networking, or alternative business platforms. This aligns with prior research showing that increased digital accessibility can fragment attention and reduce meaningful engagement with a single platform

(Rideout et al., 2010; Ophir et al., 2009; Kumar & Pansari, 2016; Luqman et al., 2017).

4.2 Social Media Engagement on Online Purchase Intention - Model 02

The influence of social media engagement on forming the online purchase intention was tested using the simple regression analysis. The results indicated a strong and statistically significant effect, and the model explained 63.7% of the variance in the online purchase intention ($R^2 = 0.637$, $F = 588.41$, $p < 0.001$).

Table 2: Regression Output of Model 02

Predictor	β	t-value	Sig.
Motivation	0.798	24.249	0.000

According to Table 2, the result has supported H_4 and underscored the critical role of social media engagement in shaping consumers' willingness to purchase products via online.

4.3 Mediation Effect of Social Media Engagement - Model 03

The mediating role of social media engagement was examined using Hayes's PROCESS Macro with 5000 bootstrap samples. The indirect effects were tested through a bias-corrected 95% confidence interval (CI).

Table 3: Summary of Output of Model 03

Path	Sig.	Indirect Effect	95% CI (LL-UL)
Motivation → Purchase Intention	0.146	0.536	[0.429, 0.653]
Opportunity → Purchase Intention	0.614	0.509	[0.425, 0.590]
Ability → Purchase Intention	0.000	0.533	[0.462, 0.650]

According to the results, as shown in Table 3, the mediation analysis revealed distinct patterns for Motivation, Opportunity, and Ability in predicting Online Purchase Intention through Social Media Engagement. Motivation demonstrated a significant indirect effect on Online Purchase Intention via Social Media Engagement, H_{5a} accepted, indicating that its influence operates entirely through engagement mechanisms. Similarly, Opportunity exerted a significant indirect influence on Online Purchase Intention through Social Media Engagement. Thus, H_{5b} was supported. In contrast, Ability exhibited significant indirect and direct effects, with the indirect pathway through Social Media Engagement complementing its direct influence on Online Purchase Intention. H_{5c} accepted.

4.4 Moderation Effect of Trust

The moderating effect of trust was tested using Hayes's PROCESS tool.

Table 4: Summary of Output of Model 04

Predictor	β	T	Sig.
Engagement	0.635	17.227	0.000
Trust	0.287	7.785	0.000
Engagement × Trust	-0.348	-1.628	0.104

According to Table 4, the results showed that the interaction term (Social Media Engagement × Trust) was not statistically significant. Thus, H_6 was not rejected. That means the trust on the vendor who published advertisements on social media, the credibility of the vendor platform has not moderated the relationship between the social media engagement and the online purchase intention. However, this trust itself

exerted a strong direct influence on the online purchase intention.

5. CONCLUSION

This study has explored how MOA factors of a consumer influence his / her social media engagement and how this engagement, in turn, shapes his / her online purchase intention, with trust as a potential moderator. Drawing on the responses from 337 no. of Sri Lankan consumers who were social media users, over 19 years old and had the purchasing power, data were collected using convenience sampling. The results showed that the motivation and ability significantly enhanced the social media engagement, while the opportunity was unexpectedly reduced it. The social media engagement itself was found to be a powerful predictor of online purchase intention and a key mediator for linking MOA factors to the online purchase intention. Then, the trust of the advertisements that are posted by the online vendors via social media platforms was not a moderator, but it exerted a strong direct influence on the online purchase intention, which was underscoring its importance in the Sri Lankan e-commerce landscape.

The aforementioned results of the study have contributed to the theory by extending the MOA framework to the context of social commerce, confirming that social media engagement acts as the bridge between psychological drivers and behavioral outcomes. As managerial implications, these findings have suggested that marketers and businesses must go beyond merely providing access to platforms. They should focus on creating motivating and interactive content, improving consumers' digital

skills, and most importantly, fostering trust through transparency, credibility, and secure transactions.

While this study provides valuable insights, its cross-sectional design limits the ability to capture behavioral changes over time. Hence the future researchers should adopt a longitudinal approach and expand the model with variables, such as perceived enjoyment, cultural influence, or brand loyalty, to deepen understanding. Moreover, as this study was not conducted for a specific product category, future research could apply the model to particular product types to examine whether consumer engagement and purchase intentions differ across categories. Additionally, since Opportunity and Trust did not yield the expected outcomes in this study, future studies could explore these findings further by using different samples or contexts to identify possible reasons for these variations. Future research should further examine random sample method to improve representativeness and reduce self-selection bias. As the present study relied on voluntary participation via social media, the findings are primarily generalizable to digitally active populations with similar characteristics.

In conclusion, the findings of the study demonstrate that online purchase behavior of consumers in emerging markets like Sri Lanka is driven less by opportunity and more by consumers' motivation and ability to engage in social media. Moreover, although the trust on the advertisements that are posted by the online vendors via social media platforms was not moderated the relationship between the social media engagement and the online purchase intention, there was a direct impact by it on forming the online

purchase intention. Hence, while strengthening the factors of motivation, opportunity, and ability for social media engagement of the consumers, it will be essential to convert those social media engagements into sustainable e-commerce growth in Sri Lanka.

model. *International Journal of Electronic Commerce*, 7(3), 101 - 134.

REFERENCES

- Binney, W. (2006). The nature and influence of motivation within the MOA framework: Implications for social marketing. *Journal of Nonprofit & Voluntary Sector Marketing*, 11(4), 271 - 283.
- Briliana, R. & Mursito, N. (2017). How motivation, ability, and opportunity can drive online shopping through social media. *International Journal of Business, Economics and Law*, 13(2), 52 - 57.
- Brodie, R.J., Hollebeek, L.D., Jurić, B. & Ilić, A. (2011). Customer engagement: Conceptual domain, fundamental propositions, and implications for research. *Journal of Service Research*, 14(3), 252 - 271.
- Commonwealth Secretariat. (2020). *E-commerce readiness assessment report: Sri Lanka*. London: Commonwealth Secretariat.
- Dessart, L. (2017). Social media engagement: A model of antecedents and relational outcomes. *Journal of Marketing Management*, 33(5 - 6), 375 - 399.
- Hollebeek, L.D., Glynn, M.S. & Brodie, R.J. (2014). Consumer brand engagement in social media: Conceptualization, scale development, and validation. *Journal of Interactive Marketing*, 28(2), 149 - 165.
- MacInnis, D.J. & Jaworski, B.J. (1989). Information processing from advertisements: Toward an integrative framework. *Journal of Marketing*, 53(4), 1 - 23.
- Mou, J. & Benyoucef, M. (2021). Consumer behavior in social commerce: Results from a meta-analysis. *Technological Forecasting and Social Change*, 167, 120734.
- Pavlou, P.A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance



Exploring Employee Perceptions of the Psychological Work Environment and Creative Performance in a Sri Lankan Private Bank

Wickramasinghe WADLT¹, Jothirathne PAAU²

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka*
lakminitheekshana@gmail.com¹

ABSTRACT

In the competitive and highly regulated banking industry, promoting employee creativity is essential for driving innovation and achieving sustained success. The study investigates employees' perceptions of the psychological work environment and its influence on creative performance within a leading private bank in Sri Lanka. The primary objective was to identify the key psychological factors influencing employee creativity and assess their relative importance. Grounded in Self-Determination Theory and Amabile's Componential Theory, the research focused on two psychological dimensions: work-life balance and intrinsic motivation. A cross-sectional quantitative study was conducted using a standardized online questionnaire to collect data from 212 bank employees. Data were analyzed using regression analysis. The findings revealed that psychological factors are the strongest predictors of creative performance. The study highlights that a psychologically supportive work environment, characterized by intrinsic motivation and a healthy work-life balance, is central to enhance employee creativity in the Sri Lankan banking sector. The results emphasize the need for bank management to prioritize autonomy-supportive leadership, implement policies to minimize burnout, and develop psychological safety to unlock employees' creative potential. Future research should employ longitudinal designs and larger and more diverse samples to establish causal relationships and validate these findings across other financial institutions.

KEYWORDS: Employee Creativity, Intrinsic Motivation, Psychological Work Environment, Sri Lankan Banking Sector, Work-Life Balance

1. INTRODUCTION

The psychological work environment encompasses the mental and emotional conditions that shape how employees think, feel, and behave within their workplace. It plays a vital role in determining individual well-being, motivation, and creative performance. Among the various psychological elements that influence employee behavior, work-life balance and intrinsic motivation are particularly significant.

Work-life balance is the ability of an employee to manage their personal and professional responsibilities effectively. When employees achieve a healthy work-life balance, it reduces stress levels, enhances satisfaction, and creates a positive emotional environment. The

internal urge to complete things out of curiosity, fulfillment, and personal development is known as intrinsic motivation.

The employees who are intrinsically motivated tend to demonstrate higher engagement, persistence, and creative problem-solving capabilities. In the Banking sector, the development of these psychological factors is crucial for sustaining creativity. Banks operate within highly structured, rule-bound, and performance-driven environments, which can sometimes restrict creative expression. Therefore, understanding about psychological work environment factors influences on the employees' creative performance which that is essential for managers for aiming at

enhancing innovation within such regulated contexts.

2. LITERATURE REVIEW

The study is grounded in two theories; Amabile's Componential Theory of Creativity (Amabile, 1983) and Self-Determination Theory (Deci & Ryan, 1985). While Amabile's model provides a general framework attributing to workplace environment being a critical component to creativity, it is Self-Determination Theory that offers the specific, psychological "why" behind such a phenomenon, particularly within a structured setting like banking. For a Sri Lankan private bank, understanding and nurturing the psychological work environment is not secondary; it is the primary strategy for unlocking creative performance (Amabile & Pratt, 2016).

Amabile's theory is that creativity results from the confluence of domain-relevant skills, creativity-relevant processes, and task motivation. However, its broad scope requires a sharper lens for the study. This is provided by Self-Determination Theory, which specifies that motivation and well-being depend on fulfilling three basic psychological needs: autonomy, competence, and relatedness (Adams et al., 2017).

The banking industry's built-in form, hierarchical, regulated, goal-oriented, systematically frustrates these requirements (Fu et al., 2022) and according to the findings of those authors the psychological factors have been identified as the strong predictor. Self-Determination Theory thereby provides the explanatory force necessary to leap from correlation to causation. Creativity is suppressed because the setting causes employees to feel controlled, inadequate,

and isolated. Amabile identifies the workplace as the key, and Self-Determination Theory reveals the psychological mechanisms that make workplace as a key factor.

Work-Life Balance is a Cognitive Resource. Today's research shifts work-life balance from a welfare issue to a strategic one with regard to cognitive ability (Sonnentag, 2018). Creativity demands high mental energy, which is depleted by chronic overloads and the inability to detach psychologically from work (Gagné, M et al., 2022). An overworked banker is cognitively exhausted, incapable of divergent thinking when creativity is required. This study reveals a more foundational issue that is basic manageable hours. The high mean score for work-life balance items in our descriptive statistics suggests that simply achieving a reasonable workload is a significant hurdle in banking, beyond the need for detachment. This contrasts with the previous studies in creative industries and highlights a sector-specific entry point for intervention.

Intrinsic Motivation depends on the presence of Psychological Safety. Although intrinsic motivation serves as the driving force behind. According to Lee et al., (2020), it can be easily undermined within banking environments that are heavily influenced by extrinsic pressures and rigid hierarchies. Psychological Safety defined as, the shared belief that individual can express new idea or take interpersonal risks without fear of negative consequences (Edmondson, 2018) is therefore essential for sustaining intrinsic motivation and enhancing creative behavior. It is especially crucial in Sri Lanka's cultural context, where high power distance may

naturally discourage challenging superiors (Newman et al., 2017). Intrinsic motivation here is not just about enjoying a task: it is about feeling safe enough to find satisfaction in proposing change.

The literature reveals a clear pathway that the banking environment inhibits creativity by frustrating core psychological needs, hence undermining the motivation required for creative output. Whereas previous research establishes broad connections, there is a clear gap in studying this particular psychological pathway in a Sri Lankan private bank. The study addresses this by investigating how work-life balance and intrinsic motivation serve as the fundamental mechanisms that either unlock or suppress creative performance, offering a human-centric blueprint for innovation.

3. METHODOLOGY

The study adopted a quantitative, cross-sectional and descriptive research design to examine how psychological work environment factors influence employee creativity in a leading commercial bank in Sri Lanka. A survey method was employed by using a structured 5-point Likert scale questionnaire to collect data on employee perception of psychological work environment factors and creative performance.

The target population consisted of managerial, operational, and technical staff of the bank. The total members were 4,428 employees (Sampath Bank PLC, 2024). Using Krejcie and Morgan's (1970) formula, a sample size of 354 respondents was determined in order to ensure a 95% level of confidence with 5% margin of error. Due to accessibility and

practical constraints, convenience sampling was employed.

Data were collected through an online self-administered questionnaire divided into three sections as demographic details, work environment assessment, and creativity evaluation. A pilot study with 10 respondents confirmed reliability, achieving a Cronbach's alpha value greater than 0.7.

The independent variable was the psychological work environment, represented by intrinsic motivation and work-life balance, while employee creative performance served as the dependent variable. The hypothesis of the study is as follows;

H1: Positive psychological work conditions (Intrinsic motivation, work-life balance) significantly predict employees' creative performance.

The findings are expected to demonstrate that promoting intrinsic motivation and supporting psychological well-being in structured, high – pressure work environments like banking can significantly enhance employees' creative performance.

4. RESULTS AND DISCUSSION

The regression analysis indicates a strong and statistically significant positive relationship between psychological work environment factors and employee creativity ($\beta = 0.458$, $t = 7.476$, $p < 0.001$). The model explains 21% of the variance in creative performance, confirming that psychological conditions are a major determinant of employee creativity within the banking sector.

Table 1: The effect of Psychological Work Environment Factors on Employee Creativity

Predictor	B	SE	β	p
Constant	2.307	0.215	-	<.001
Psychological Factors	1.026	0.137	0.458	<.001

Note: $R^2 = 0.210$

B = unstandardized coefficient, SE = standard error, β = standardized coefficient

The results support the hypothesis, showing that employees' creative performance is greatly improved by favorable psychological work environments, especially those that promote intrinsic motivation and work-life balance.

According to the results, psychological components had the strongest predictive influence, even after adjusting for other environmental factors, which highlights their critical role in promoting creative behavior.

The above findings support the fundamental ideas of Self-Determination Theory (Deci & Ryan, 1985), which emphasizes that when workers feel that they are competent, autonomous, and connected, then their creativity would flourish. Promoting psychological safety, intrinsic motivation, and emotional well-being are essential for improving creative outcomes in the banking sector, which is characterized by strict structures and high-performance pressures.

In conclusion, the study establishes that improving psychological conditions within the workplace by promoting work-life balance, reducing stress, and encouraging autonomy can substantially increase employees' creative outcomes. Hence, managerial strategies in the banking sector should move beyond task efficiency to include psychological well-

being as a strategic enabler of innovation and sustained competitive advantage.

5. CONCLUSION

This study highlights that psychological work environment factors particularly intrinsic motivation and work-life balance play a decisive role in developing employee creativity within Sri Lanka's banking sector. In an industry often characterized by rigid hierarchies, performance-driven cultures, and limited autonomy, the findings affirm that creativity thrives not merely through structural or material resources, but through an environment that nurtures psychological safety, trust, and emotional well-being.

The empirical results demonstrate that when employees feel valued, supported, and intrinsically motivated, they are more inclined to engage in innovative thinking and problem-solving. This aligns with the Self-Determination Theory (Deci & Ryan, 1985), which emphasizes autonomy, competence, and relatedness as fundamental drivers of intrinsic motivation and creativity.

Therefore, banks should give top priority to creating strategies that improve the psychological well-being of their workforce. These strategies should include leadership practices that foster transparency and trust, flexible work schedules, and recognition programs that increase intrinsic motivation. By strengthening these psychological dimensions, banking institutions can unlock their employees' creative potential and sustain innovation within a traditionally conservative and highly regulated industry.

REFERENCES

- Adams, N., Little, T.D., & Ryan, R.M. (2017). Self-determination theory. In *Development of self-determination through the life-course* (pp. 47-54). Dordrecht: Springer Netherlands.
- Amabile, T.M. (1983). The social psychology of creativity: a componential conceptualization. *Journal of personality and social psychology*, 45(2), 357.
- Amabile, T.M., & Pratt, M.G. (2016). The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in organizational behavior*, 36, 157 - 183.
- Deci, E.L., & Ryan, R.M. (1985). The general causality orientations scale: *Self-determination in personality*. *Journal of research in personality*, 19(2), 109 - 134.
- Edmondson, A.C. (2018). *The fearless organization: Creating psychological safety in the workplace for learning, innovation, and growth*. John Wiley & Sons.
- Fu, Q., Cherian, J., Rehman, K.U., Samad, S., Khan, M. A., Athar Ali, M., & Miculescu, A. (2022). Enhancing employee creativity in the banking sector: a transformational leadership Framework. *Sustainability*, 14(8), 4643.
- Gagné, M., Parker, S. K., Griffin, M.A., Dunlop, P.D., Knight, C., Klonek, F.E., & Parent-Rocheleau, X. (2022). Understanding and shaping the future of work with self-determination theory. *Nature Reviews Psychology*, 1(7), 378 - 392.
- Krejcie, R.V., & Morgan, D.W. (1970). Determining Sample Size for Research Activities. *Educational and Psychological Measurement*, 30(3), 607 - 610.
- Lee, A., Legood, A., Hughes, D., Tian, A. W., Newman, A., & Knight, C. (2020). Leadership, creativity and innovation: A meta-analytic review. *European Journal of Work and Organizational Psychology*, 29(1), 1 - 35.
- Newman, A., Donohue, R., & Eva, N. (2017). Psychological safety: A systematic review of the literature. *Human resource management review*, 27(3), 521 - 535.
- Sonnentag, S. (2018). The recovery paradox: Portraying the complex interplay between job stressors, lack of recovery, and poor well-being. *Research in Organizational Behavior*, 38, 169 - 185.
- Sampath Bank PLC. (2025, March 10). Annual report 2024. Retrieved from <https://www.sampath.lk/DigitalAnnualReport2024/>



Bridging Efficiency Gaps in Hospitality SMEs through a Computerized Payroll Management System: A Case Study

Kanchana GMSD¹, Wattegama EJ²

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka*
kanchanadinuk123@gmail.com¹

ABSTRACT

This paper looks at the design, implementation, and evaluation of a computerized payroll management system created for a purposely selected restaurant chain in Sri Lanka. The selected business firm is a small and medium sized company which is functioning under the hospitality sector and often struggle with maintaining effective payroll management therein. This is due to their heavy reliance in manual systems and limited digital resources. The main goal of this study was to automate payroll operations to improve accuracy, cut processing time, and support better decision-making. A mixed-methods approach was used collecting both qualitative and quantitative data to assess system performance. The system was built with the MERN stack (MongoDB, Express.js, React, and Node.js). It includes key modules such as attendance, leave, salary computation, EPF/ETF processing, and report generation. The results showed that payroll processing time decreased by 62%. Calculation errors fell from 8.5% to 1.2%. User satisfaction has also improved with a System Usability Scale (SUS) score of 88 out of 100. By addressing major gaps in existing research, especially the lack of empirical evidence on payroll automation in small and medium-sized businesses and obstacles in developing economies this study adds to the growing knowledge of digital transformation in hospitality businesses.

KEYWORDS: Automation, Efficiency, Hospitality, MERN Stack, Payroll System

1. INTRODUCTION

Payroll management is a vital administrative function in any organization, ensuring accurate and timely employee compensation while maintaining compliance with statutory requirements such as EPF and ETF contributions. However, in hospitality-sector small and medium-sized enterprises (SMEs) in developing countries like Sri Lanka, payroll activities remain heavily dependent on manual processes-most commonly handwritten records, basic spreadsheets, or outsourced accounting services (Silva & Fernando, 2023). These outdated practices frequently lead to calculation errors, inconsistencies in attendance tracking, delayed salary disbursements, data redundancy, and higher operational costs, all of which negatively impact employee morale and

organizational productivity. Although global literature highlights the transformative potential of payroll automation for reducing administrative burden and improving accuracy (Ahmed et al., 2023; Chennupati et al., 2023), most of existing researches were centered on technologically advanced economies or large corporate environments. Consequently, little empirical evidence exists on how automated payroll systems perform within hospitality SMEs operating under resource constraints, infrastructural limitations, and low levels of digital readiness-challenges that are particularly prevalent in developing-country contexts (Silva & Fernando, 2023). Furthermore, prior studies rarely examine the usability, user acceptance, and socio-technical difficulties encountered during system adoption, creating a significant knowledge gap

regarding how SMEs transition from manual to fully automated payroll processes.

This study addresses these gaps by presenting a real-world, evidence-based evaluation of a computerized payroll management system developed for a hospitality SME in Sri Lanka. The research demonstrates how a MERN-stack-based solution can enhance efficiency, reduce processing time, minimize human error, and improve transparency in payroll operations. By integrating core functions such as attendance tracking, leave management, salary computation, EPF/ETF processing, and report generation, the system directly responds to the operational challenges faced by SMEs that lack sophisticated HR technologies. The significance of the study lies in its ability to provide measurable performance outcomes-such as reductions in processing time and error rate, along with improved user satisfaction-thereby offering empirical insight into the effectiveness of payroll automation in environments where digital transformation is still emerging. The purpose of this research is therefore to design, implement, and evaluate a computerized payroll management system tailored to the needs of hospitality SMEs, and to contribute to the academic and practical understanding of how digital solutions can bridge efficiency gaps in developing-country business settings. By providing context-specific evidence and highlighting both technical and human factors influencing adoption, this study offers a replicable framework that can guide other SMEs seeking to modernize their payroll operations.

2. LITERATURE REVIEW

The automation in payroll management has shown great promise to streamline operations and improve compliance accuracy (Chennupati et al., 2023). Ahmed et al. (2023) found that computerized payroll systems can cut administrative work by over 60%. They highlighted usability and transparency as key factors in user adoption. Similarly, Tobing et al. (2022) pointed out inefficiencies in Indonesian organizations due to decentralized payroll systems and suggested that digital integration could help. Aina and Odun-Ayo (2025) proposed cloud-based structures for small and medium-sized enterprises (SMEs), showing that scalability and accessibility could address infrastructure issues. However, there is little research on hospitality SMEs in developing economies, particularly regarding empirical validation. Lakshmirevathi et al. (2024) looked into payroll modernization, stressing the importance of digital literacy and commitment from management. Phiri and Mogaji (2024) studied barriers to HR digital transformation in African SMEs and uncovered cultural and infrastructure challenges. In Sri Lanka, Silva and Fernando (2023) found that SMEs frequently struggle with digital adoption because of budget constraints and limited IT skills. This body of literature highlights that while there is significant global progress in payroll automation, the specific challenges and effects in developing areas remain less explored.

The present study aims at addressing this by empirically evaluating the efficiency, accuracy, and usability of a

MERN-based payroll system designed for the hospitality industry.

3. METHODOLOGY

This study had used a mixed-methods approach that combined both quantitative and qualitative techniques to ensure a thorough evaluation. Data were collected from the management of the Restaurant and 45 employees. The system was developed using the System Development Life Cycle (SDLC) model. The design included six modules: employee management, attendance, leave tracking (Figure 1 & Figure 2), payroll computation (Figure 3, 4, & Figure 5), EPF / ETF automation (Figure 6 & Figure 7), and report generation. Testing used both white-box and black-box methods to confirm functionality and accuracy. During the three months before and after implementation, we collected quantitative metrics such as processing time, calculation error rate, and cost by using questionnaire. We used the System Usability Scale (SUS) to measure user satisfaction-qualitative data from interviews offered insights into barriers to adoption and the benefits of the system.

4. RESULTS AND DISCUSSION

In the implemented system at the selected restaurant chain, the employee attendance was recorded through the use of barcode identification cards integrated with barcode readers, ensuring precise and reliable attendance tracking. This automation effectively eliminated the inaccuracies and inconsistencies that were previously associated with manual or spreadsheet-based methods. The system automatically processed all payroll-related computations that include salary generation, leave deductions, and

allowance adjustments, thereby significantly reducing the human error and enhancing the overall efficiency of payroll operations. Furthermore, the developed computerized platform provided comprehensive management of employee attendance, leave, and half-day records in real time. This functionality not only minimized administrative overhead but also streamlined payroll processing by consolidating all relevant employee data within a single, centralized system. Employees were granted secure, role-based access to their individual payroll and attendance information, which fostered greater transparency and accountability within the organization. By ensuring timely, accurate, and fully automated salary disbursement, the system contributed directly to improved employee satisfaction and motivation. Consequently, the implementation of this computerized payroll management system enhanced overall workforce productivity and operational effectiveness within the selected restaurant.

The results showed that payroll processing time was reduced from 8.5 hours to 3.2 hours per week, representing a 62% improvement (Figure 8). Calculation errors dropped from 8.5% to 1.2% (Figure 9), and external accounting costs fell by 45% (Figure 10). User satisfaction reached an SUS score of 88 out of 100 (Figure 11), indicating excellent usability. Feedback indicated increased transparency and a reduced manual workload, but also highlighted challenges such as initial user resistance and reliance on stable internet connectivity. These findings support the conclusions of Ahmed et al. (2023) and Chennupati et al. (2023), who found similar improvements after payroll

digitization. However, this study adds to the existing literature by validating these effects in a developing-country SME context, filling a clear research gap in

regional studies. It also shows that system success relies on organizational readiness, user training, and compliance with local regulatory standards.

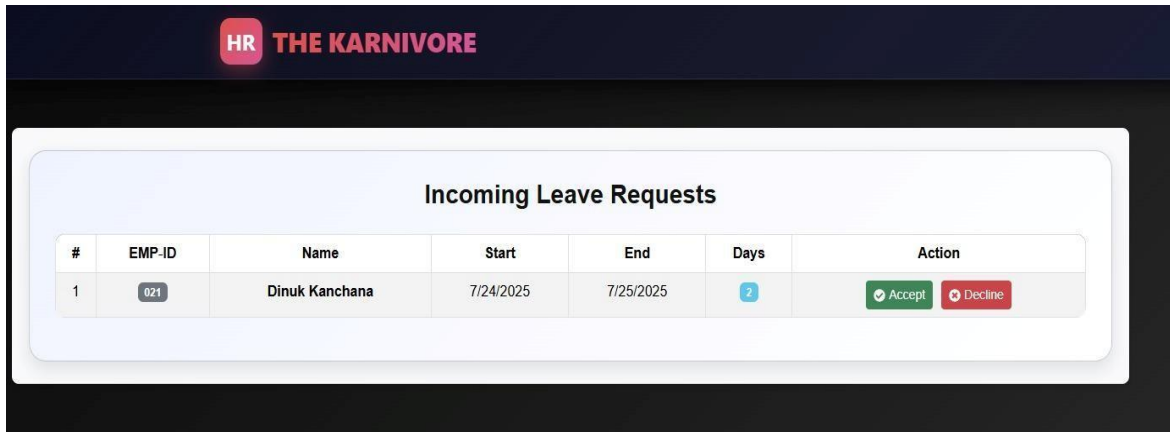


Figure 1: Leave Handling Interface

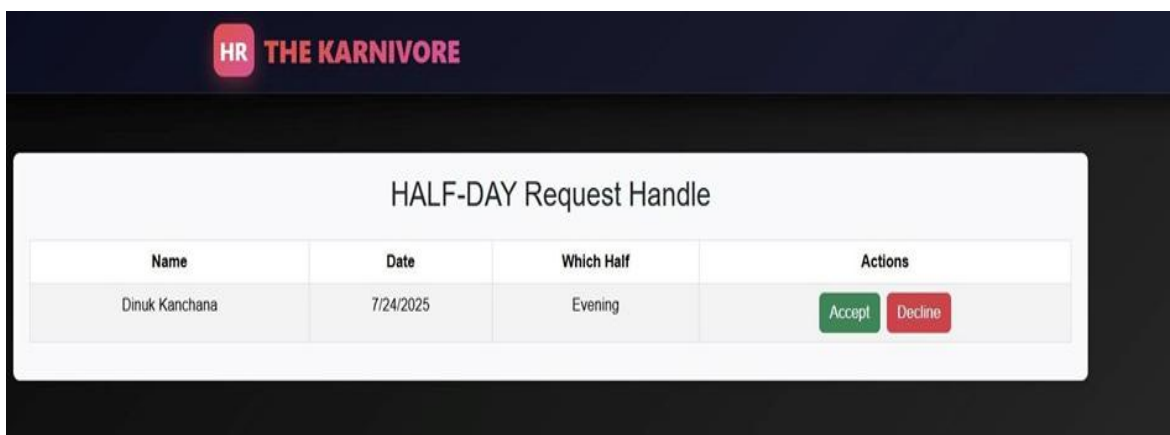


Figure 2: Half Day Handling Interface

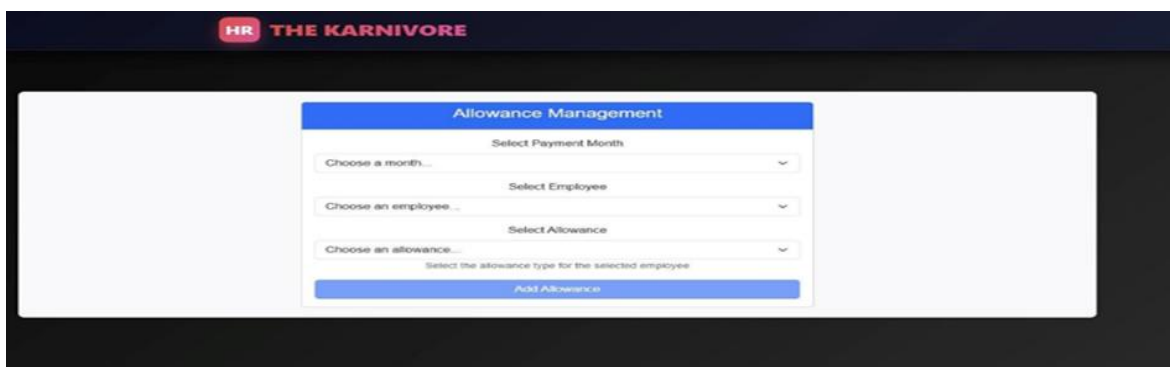


Figure 3: Allowances Handling Interface

BRIDGING EFFICIENCY GAPS IN HOSPITALITY SMES THROUGH A COMPUTERIZED PAYROLL MANAGEMENT SYSTEM: A CASE STUDY

#	Name	Role	Requested Salary	Status / Actions
1	Dinuk Kanchana	Headchef	Rs 10000	Accept Decline

Figure 4: Advance Handling Interface

Service Charge Management

Select Payment Month:

Select Employee:

Service Charge Amount:

Figure 5: Service charges Handling Interface

EPF Payment Management

Select Permanent Employee:

Year:

Month:

Contribution Details:

- Basic Salary: 0.00
- Employee Contribution (8%): 0.00
- Employer Contribution (12%): 0.00

Figure 6: EPF Handling Interface

ETF Payment Management

Select Permanent Employee:

Month:

Year:

Figure 7: ETF Handling Interface

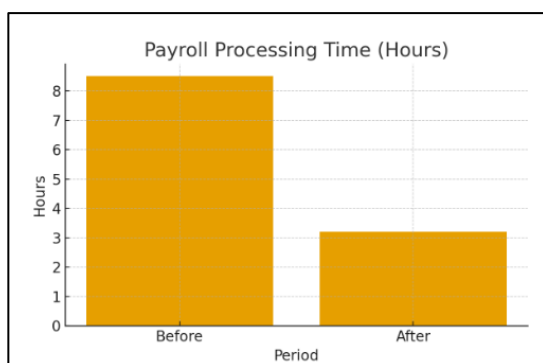


Figure 8: Comparison of Payroll Processing Time Before & After System Implementation

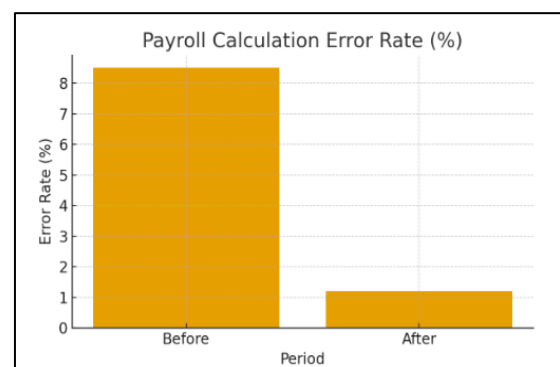


Figure 9: Comparison of Payroll Calculation Error Rates Before & After System Implementation

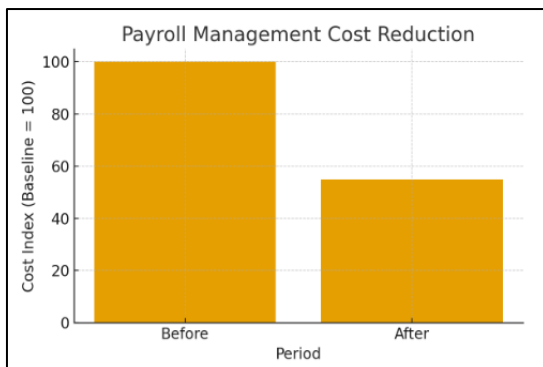


Figure 10: Cost Reduction in Payroll Management Before & After System Implementation

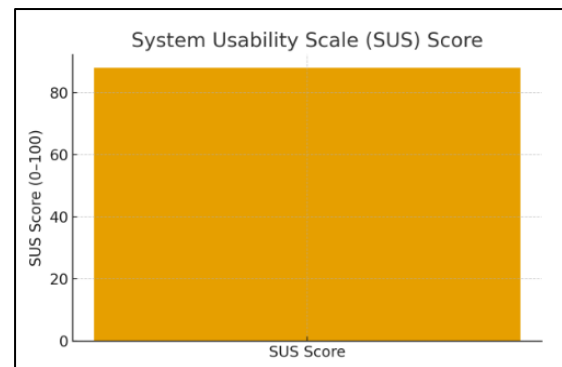


Figure 11: System Usability Scale (SUS) Score of the Computerized Payroll System

5. CONCLUSION

The computerized payroll management system which was implemented at the selected restaurant chain made the daily operation smoother, more accurate, and cost-effective. By tackling some important gaps like the lack of real-world evidence in hospitality businesses, overlooking people and technology challenges, and not focusing much on how easy systems are to use, this study helps giving better understand for how payroll systems are adopted in developing countries. Future researchers can look at adding features like fingerprint-based attendance, mobile access, or even predictive tools to help manage staff more efficiently. These findings not only provide a practical model that other small businesses can follow, but they also contribute to our academic understanding of how digital tools are transforming the way payroll is managed.

REFERENCES

- Ahmed, A.M., Nadheef, C., & Ahmad, A. M. (2023). Web-based payroll management system: Design, implementation, and evaluation. *Journal of Electrical Systems and Information Technology*, 10, 17.
- Aina, A., & Odun-Ayo, I. (2025). Development of a cloud-based payroll management system. arXiv preprint arXiv:2502.16321.
- Chennupati, P.S., Surya, G.V., & Karthik, T. (2023). Automated payroll system. *International Journal of Communication and Information Technology*, 4(1), 64 - 70.
- Lakshmirevathi, K., Varalakshmi, T., & Omprakash, K. (2024). Payroll management systems – compensation modern payroll management. *International Research Journal of Advanced Engineering and Management*, 2(5).
- Phiri, K., & Mogaji, E. (2024). Digital HR transformation in developing economies: Challenges and opportunities. *Journal of Business Transformation*, 6(2), 55 –70.
- Silva, D., & Fernando, R. (2023). Digital adoption barriers among SMEs in Sri Lanka. *South Asian Journal of Management*, 30(4), 112 - 127.
- Tobing, K.S.L., Lantana, D.A., & Digidowiseiso, K. (2022). The implementation of payroll accounting information systems in Indonesia: A systematic literature review. *International Journal of Economics, Management, Business, and Information Systems*, 3(2).



Impact of the Logistic Challenges on enhancing the Operational Performance: A Case Study

Rasangi NC¹, Wattagama EJ²

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka*
chethanarasangi24@gmail.com¹

ABSTRACT

The logistics service provides the structural and procedural backbone that enables smooth operations of a firm which is engaging in the supply and maintenance of cargo-handling equipment and other infrastructure, pilotage services, docking, and shipping. Aim of this study was identifying the most affected logistics challenges on operational performance and mitigates logistical inefficiencies to improve port operational efficiency of purposively selected firm of Sri Lanka Ports Authority. The variables of logistics challenges, such as, container freight station infrastructure, berth congestion, yard capacity stress and the operational performance, such as, imports, exports, and transshipments were measured using the data collected from primary and secondary sources of Sri Lanka Ports Authority. Descriptive statistics, comparison, and multiple linear regression analysis were used to test the study hypotheses. Findings revealed that significant variations in the logistic challenges influenced on the import operations of Ports Authority. The study recommends that; by minimizing container dwell time, optimizing berth allocation to expand capacity, upgrading port systems with modern technologies, and digitalizing documentation processes with staff trainings, enhance the operational performance of the Ports Authority. A feasibility assessment indicated that minimizing dwell time, optimizing the berth, and enhancing digitalisation were the most viable measures across technical, financial, operational, and organizational dimensions. As managerial implication, the study emphasized the critical need to address inefficiencies, particularly in the import operation to enhance the overall performance of Sri Lanka Port Authority.

KEYWORDS: Logistic Challenges, Operational Performance, Ports Authority

1. INTRODUCTION

This study examines how logistical challenges affect the operational performance of purposively selected firm of Sri Lanka Ports Authority while focusing on the operational performance as imports, exports, and transshipments. Generally, the port is a vital organization to do global trading for a country but such organizations are facing an increase pressure to handle larger volumes efficiently. The Ports Authority of Colombo experiences some logistics issues, such as, berth congestion, yard capacity limits, outdated systems, poor road access, and documentation delays and all of which hinder its operational

performance. The study aim at identifying key logistics constraints, measuring the impacts of those on operational performance, and determining which type of operation was mostly affected. The significance of this study lies in offering practical recommendations to relevant authorities to improve efficiency, support digitalization, and strengthen the competitiveness of Sri Lanka Ports Authority while contributing to limited empirical research on port logistics in emerging economies like Sri Lanka.

2. LITERATURE REVIEW

Port logistic is crucial for international trade, ensuring efficient cargo handling,

storage, documentation, and inland connectivity that support global supply chains. It includes customs clearance and multimodal transport essential for smooth cargo flow and economic growth. Port operations which cover imports, exports, and transshipments determine its overall efficiency (Val'kova et al., 2021). However, the ports face issues, such as, congestion from limited berth and yard capacity (Gidado, 2015; Bichou & Gray, 2004), regulatory bottlenecks (Notteboom & Winkelmanns, 2001), and poor inland transport connectivity (Rathnayake & Kavirathna, 2023). The manual documentation and outdated systems further reduce efficiency (Nguyen et al., 2016; Sarkar et al., 2023). The efficient logistics improve throughput and reliability while deficiencies increase costs and reduce competitiveness (Loh & Van Thai, 2014). Integrating port-centric logistics and digital innovations enhances the operational performance (Mangan et al., 2008; Notteboom & Rodrigue, 2022). Accordingly, present study highlights the need for empirical studies on logistics challenges of the port in the Sri Lankan context. Figure 1 shows the conceptual framework that was established to elaborate the study objective.

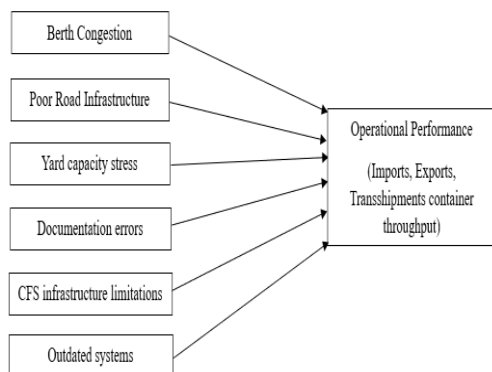


Figure 1: Conceptual Framework

3. METHODOLOGY

3.1 Research Design

This study adopted a quantitative research design with the explanatory approach to examine how logistical challenges affect operational performance in port operations in Sri Lanka. Table 1 shows the details of the variables that were selected for the study.

Table 1: Details of the Variables

Variable Name	Indicators
Road Infrastructure	Perceptions of the Road conditions, Traffic congestion and Regular maintenance
Documentation errors	Perceptions of the Error rate in documentation, Manual processing, and Resubmit documents
Systems efficiency	Perception on System failures, Outdated systems, and Automation level
Berth Congestion	Vessel turnaround time (days)
Yard capacity stress	Yard turnover ratio
CFS infrastructure	CFS lifting capacity (tons)
Operational performance	Imports, Exports, Transshipments (Container throughput)

3.2 Data Collection

This study used both primary and secondary sources to collect data to test its hypotheses. The primary data were collected for measuring the variables of Road infrastructure, Documentation errors, and System efficiency by interviewing and sending a questionnaire among randomly selected 60 no. of

stakeholders of the Port Authority; the staff of operations division, logistic division, navigation section, and inland transportation providers. All the question items were established to get the perceptions for measuring aforementioned variables in a 5-point Likert scale. A pilot test was carried out using a convenience sample and it was confirmed the acceptable reliability (Cronbach's $\alpha > 0.70$) of the question items to be used in this selected study context. The secondary data were collected for the past five years (2020 - 2024) from the annual reports, operational reports, and statistics of the Port Authority for measuring the variables of Berth congestion, Yard capacity stress, Container freight station (CFS) lifting capacity, Imports, Exports, and Transshipments.

3.3 Data Analysis / Interpretation Techniques

The data was analyzed using SPSS and R software. First, descriptive comparison was used to analyze the primary variables by calculating mean and standard deviation values to identify the most influential factors on operational efficiency. Correlation analysis was applied to examine the strength and direction of relationships between variables. Multiple linear regression models were then used to analyze the secondary data and Stepwise regression and the Akaike Information Criterion (AIC) were used to identify the most

adequate models. The model comparison using adjusted R^2 values help to identify the operational area that was most affected by logistical challenges.

4. RESULTS AND DISCUSSION

4.1 Primary Data Analysis

The primary data were analyzed using the composite mean scores that were calculated for each variable. The overall mean and standard deviation across 60 no. of responses were then computed. The higher mean values were indicated stronger perceived impact and the standard deviations were reflected the response consistency. The results are shown in Table 2.

Table 2: Descriptive statistics

Variable	Mean	SD	Result
Road Infrastructure	2.94	0.55	Neutral
Documentation Errors	3.48	0.67	High
System Efficiency	3.84	0.52	High

4.2 Correlation Analysis

Correlation analysis was performed and results are shown in Table 3. The strength and direction of the relationship between logistical challenges and port operations (imports, exports, transshipments) are indicating whether these challenges are positively or negatively associated with port performance, but without proving causality.

Table 3: Results of the Correlation Test

Independent Variable	Imports (r)	Exports (r)	Transshipments (r)	Interpretation
Berth congestion	-0.826**	-0.772**	-0.708**	Strong negative with all
Yard capacity stress	0.877**	0.820**	0.707**	Strong positive with all
Road infrastructure	0.144 (ns)	0.128 (ns)	0.745**	No effect on imports/exports Strong positive with transshipments
System efficiency	0.864**	0.711**	0.316*	Strong positive with imports/exports Weak positive with transshipments
CFS infrastructure	0.847**	0.753**	0.714**	Strong positive with all
Documentation errors	-0.841**	-0.649**	-0.380**	Strong negative with imports/exports Moderate negative with transshipments

*Significant at 5% level; **Significant at 1% level; ns = not significant

4.3 Multiple Regression Analysis

The results are shown in Table 4, 5, and 6.

Table 4: Stepwise Regression results on Imports

Variable	p- value	Result
(Intercept)	0.23813	Not significant
Berth congestion	0.10779	Not significant
Yard capacity stress	0.00275	Significant
CFS Infrastructure	0.15041	Not significant

Table 5: Stepwise Regression results on Exports

Variable	p- value	Result
(Intercept)	0.572055	Not significant

Berth congestion	0.103087	Not significant
Yard capacity stress	0.000141	Significant

Table 6: Stepwise Regression results on Transshipments

Variable	p- value	Result
(Intercept)	0.9282	Not significant
Berth congestion	0.0336	Significant
Yard capacity stress	0.0191	Significant

4.4 Identify the most influenced dependent variable

To determine which dependent variable is mostly influenced by the independent variables, the adjusted R^2 values of the regression models were compared. Adjusted R^2 was used instead of R^2 since it accounts for differing numbers of predictors and penalizes

variables that do not improve model explanatory power, making it a more reliable measure for model comparison. The results are shown in Table 7.

Table 7: R² Value Table of Three Models

Model	Adjusted R ² value
Imports model	0.7834
Exports model	0.6761
Transshipment model	0.5321

5. CONCLUSION

The study concludes that logistical challenges have a measurable and significant impact on the operational performance of Ports Authority with internal factors that are demonstrating a greater influence than external ones. Among primary data variables examined, the system efficiency and yard capacity stress were emerged as the most critical determinants of the operational performance and have derived as strongly influencing on the import and export operations. The results of the regression analysis confirmed that yard capacity stress had the highest statistical significance which was explaining as most of the variation in operational outcomes while berth congestion and documentation errors were found to negatively affect port throughput by causing operational delays and administrative inefficiencies. In contrast, road infrastructure exhibited a weaker impact and it was suggesting that while it contributes to logistical flow, internal port systems play a more direct role in determining operational effectiveness.

The results implied that the predictors have the strongest impact on imports, a moderate impact on exports and a relatively lower impact on transshipments. Accordingly, the relevant authorities should focus on the strategies

to eliminate the inefficiencies that are related to Yard capacity stress, Documentation errors, Berth congestion, and System efficiency and improve the operational performances of Ports Authority of Sri Lanka. Moreover, the study contributes to the body of knowledge by proposing that minimizing container dwell time, optimizing berth allocation, and digitalizing port processes are central theoretical constructs that provide a conceptual foundation for future empirical research aimed at modeling the causal relationships between operational strategies and port productivity outcomes.

REFERENCES

- Bichou, K. & Gray, R. (2004). A logistics and supply chain management approach to port performance measurement. *Maritime Policy & Management*, 31(1), 47 - 67. <https://doi.org/10.1080/0308883032000174454>
- Gidado, U. (2015). Consequences of port congestion on logistics and supply chain in African ports. *Developing Country Studies*, 5(6), 160 - 167. <https://core.ac.uk/download/pdf/234682294.pdf>
- Loh, H.S. & Van Thai, V. (2014). Managing port-related supply chain disruptions: A conceptual paper. *The Asian Journal of Shipping and Logistics*, 30(1), 97 - 116. <https://doi.org/10.1016/j.ajsl.2014.03.002>
- Mangan, J., Lalwani, C. & Fynes, B. (2008). Port-centric logistics. *The International Journal of Logistics Management*, 19(1), 29 - 41. <https://doi.org/10.1108/09574090810872587>
- Nguyen, M.D., Kim, S.J. & Jeong, J.S. (2016). Coordination challenges in port logistics and CFS operations. *Transport Policy*, 45, 145 - 156. <https://doi.org/10.5394/KINPR.2016.40.5.305>
- Notteboom, T. & Rodrigue, J.P. (2022). Port congestion and shipping network dynamics. *Transport Reviews*, 42(3), 289 - 306. <https://doi.org/10.1080/01441647.2021.1986005>

- Notteboom, T. & Winkelmanns, W. (2001). Structural changes in logistics: How will port authorities face the challenge? *Maritime Policy & Management*, 28(1), 71 - 89. <https://doi.org/10.1080/03088830119197>
- Rathnayake, P.G.B.P.S. & Kavirathna, C.A. (2023). Feasibility of establishing an integrated railway-based containerised cargo transport network between the Port of Colombo and free trade zones in Sri Lanka. *Journal of South Asian Logistics and Transport*, 3(2), 103 - 127. <https://doi.org/10.4038/jsalt.v3i2.72>
- Sarkar, B.D., Shankar, R., & Kar, A.K. (2023). Port logistic issues and challenges in the Industry 4.0 era for emerging economies: An India perspective. *Benchmarking: An International Journal*, 30(1), 50 - 74. <https://doi.org/10.1108/BIJ-08-2021-0499>
- Val'kova, S., Stepanets, V. & Kiseleva, E. (2021). Comparative evaluation of efficiency and justification of choosing option of cargo handling in port. *Vestnik of Astrakhan State Technical University. Series: Marine Engineering and Technologies*, 3, 107 - 115. <https://doi.org/10.24143/2073-1574-2021-3-107-115>



Impact of Climate Change on the Perceived Income of Street Food Vendors in Sri Lanka

Karunaratne BMD¹, Wategama EJ¹

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka*
maleeshadewindi74@gmail.com¹

ABSTRACT

Street food vending plays a vital role in the Sri Lankan informal economy, offering affordable food and livelihoods to low-income communities. However, the vendors face increasing disruption from climate-related factors, such as extreme heat, heavy rainfall, and unpredictable weather patterns. This study aimed to examine how those factors influence the perceived income of street food vendors and identify their adaptive strategies. Using a cross-sectional design, data were collected from 50 vendors across all nine provinces through structured interviews using a questionnaire, in which the questions were rated on a 5-point Likert scale. The variables included in the model were perceptions of extreme heat, rainfall intensity, weather unpredictability, and overall perceived income. The data were analyzed using SPSS 27 version, employing descriptive statistics, reliability testing, and multiple regression analysis. The results showed that weather unpredictability was the only statistically significant predictor in declining the perceived income, while extreme heat and rainfall intensity, though reported as operational challenges, were not significant in the model. The street food vendors have indicated that unpredictable weather is a limiting factor in their planning, which is leading to food wastage and a reduction in customer patronage. However, for the heat and rainfall, common adaptive practices, such as adjusting work hours, using temporary shelters, and relocating stalls, help reduce the negative impact on the income. These findings reveal the urgent need for policy interventions to have forecasting systems for providing vendor-specific early warning mechanisms, provide climate-resilient vending infrastructure, enhance vendors' adaptive capacity through training, and provide access to microfinance for buying protective equipment and improved public shelter facilities.

KEYWORDS: Perceived income, Street food vendors, Weather unpredictability

1. INTRODUCTION

Street food vending represents a major component of Sri Lanka's informal economy, which supplies affordable food and employment to marginalized groups. Climate change, which is manifested in rising temperatures, erratic rainfall, and unpredictable weather, disrupts such vendors' daily operations and income. Previous studies have highlighted informal workers' vulnerability to environmental shifts due to a lack of infrastructure and institutional support (For example, Bhowmik & Saha, 2012; Karunaratna & Tjandra, 2021; Samaraweera et al., 2024). The present

study investigates the impact of climate factors on the vendors' perceived income and identifies adaptation strategies to enhance their resilience.

2. LITERATURE REVIEW

Street food vending has been identified as an informal business sector that could be used to engage the marginalized community groups in the economic activities of a country like Sri Lanka (Bhowmik, 2005; Roever & Skinner, 2016). Despite its socio-economic importance, vendors remain highly vulnerable to climate risks due to their dependence on outdoor workspaces and lack of institutional protection (Lund

& Naidoo, 2016). In Sri Lanka, street food vending sustains thousands of urban livelihoods but remains under-researched regarding how climate impacts their daily income (Karunaratna & Tjandra, 2021).

Globally, the studies have shown that extreme heat, rainfall, and weather unpredictability have threatened the vendors' income by disrupting operations, reducing customer turnout, and increasing food spoilage (Dodman et al., 2023; Sverdlik et al., 2024). For example, the evidences from India and Africa demonstrate that unregulated street vending is highly sensitive to changing weather conditions which is resulting in lost income and poor health outcomes among workers (Bhowmik & Saha, 2012; Onyango, 2012). The vendors' adaptive capacity is further limited by the absence of climate-resilient infrastructure, access to finance, and inclusion in policy frameworks (Sverdlik et al., 2024).

In Sri Lanka, climate changes, such as rising temperatures, irregular rainfall, and increasing weather variability, have intensified challenges for outdoor workers, particularly street food vendors who rely on the daily income (Samaraweera et al., 2024). Some studies have reported that unseasonal rains and heatwaves cause early stall closures, damage to goods, and reduced sales. Yet, despite their contribution to food security and employment, vendors are rarely included in formal adaptation strategies (Karunaratna & Tjandra, 2021).

These findings suggest that climate change disproportionately affects informal street vendors through operational and economic disruptions. Key influencing factors, such as extreme heat, rainfall intensity, and weather unpredictability, form the conceptual

foundation for this study, focusing on their impact on vendors' perceived income and adaptive responses in Sri Lanka.

Based on the aforementioned theoretical foundation, a conceptual framework was drawn (Figure 1) together with the following hypotheses;

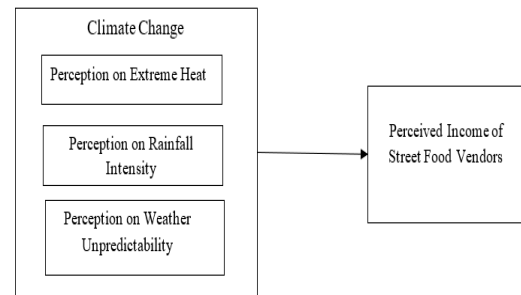


Figure 1: Conceptual Framework

H₁: The perception of the extreme heat has a significant negative effect on the perceived income of street food vendors.

H₂: The perception of the rainfall intensity has a significant negative effect on the perceived income of street food vendors.

H₃: The perception of the weather unpredictability has a significant negative effect on the perceived income of street food vendors.

3. METHODOLOGY

This study has adopted a quantitative research design, using a structured, partly adapted and self-developed questionnaire to examine the effects of climate change on street food vendors in Sri Lanka. Since no prior research had addressed this specific context, the instrument was based on Ghosh's "Heat Havoc" (2024) study and modified to reflect the Sri Lankan context. Data were collected through face-to-face interviews from a purposively selected 50 no. of vendors across nine provinces using

a 5-point Likert scale to quantify perceptions of extreme heat, rainfall intensity, and weather unpredictability. The responses were analyzed using SPSS for descriptive and multiple regression analysis to identify significant relationships. Ethical standards were maintained through informed consent, anonymity, and voluntary participation. Although the sample size of 50 is relatively small for quantitative analysis, it meets the minimum requirement for regression models and is acknowledged as a limitation regarding generalizability.

4. RESULTS AND DISCUSSION

A total of 50 valid responses from street food vendors across all nine provinces were analyzed. A pilot test was conducted to ensure the reliability and internal consistency of the self-developed questionnaire. Cronbach's Alpha values for all constructs: Extreme Heat ($\alpha = 0.772$), Rainfall Intensity ($\alpha = 0.703$), Weather Unpredictability ($\alpha = 0.709$), and Perceived Income ($\alpha = 0.732$) have exceeded 0.7 and confirmed the instrument's reliability for statistical analysis. The multiple regression model examined the effects of extreme heat, rainfall intensity, and weather unpredictability on perceived income. The results are shown in Table 1.

Table 1: Model Summary

Model	R	R ²	Adjusted R ²	Std. Error of the Estimate
1	0.870 ^a	0.757	0.741	0.613

a. Predictors: (Constant), Weather Unpredictability Score, Extreme Heat Score, Rainfall Intensity Score

According to Table 1, it was shown that the model explained 75.7% of the variance in perceived income, which indicates strong explanatory power.

Table 2: Results of Regression Analysis

Variable	B	Sig.
Extreme Heat	-0.008	0.927
Rainfall Intensity	0.202	0.287
Weather Unpredictability	0.803	0.000

According to Table 2, the results showed that weather unpredictability has a significant positive association with perceived income impact ($p < 0.001$), confirming it as the strongest driver of income instability. The negative coefficient for extreme heat ($B = -0.008$), though insignificant, aligns with expectations that higher temperatures reduce customer turnout and productivity. However, the vendors' adaptive practices, such as working during cooler hours and using shaded areas, likely reduced their measurable effect. Rainfall intensity, though positive, was not significant, possibly reflecting adaptive behaviour such as selling warm foods or using mobile stalls during rain.

These findings indicate that Sri Lankan vendors are resilient to predictable conditions like heat and rain but remain highly vulnerable to unpredictable weather changes, echoing findings by Dodman et al. (2023) and Samaraweera et al. (2024). Targeted weather communication and micro-adaptive support are essential to enhance their climate resilience.

5. CONCLUSION

The study found that weather unpredictability has significantly affected

the perceived income of street food vendors in Sri Lanka, while extreme heat (negative but insignificant) and rainfall intensity (positive but insignificant) had no notable impact. These results suggest that vendors have developed localized adaptations, such as shifting working hours, using shaded or waterproof stalls, and relocating, that help reduce the direct financial impacts of predictable weather. However, sudden and erratic weather changes remain a major source of income instability, aligning with findings by Dodman et al. (2023) and Heal and Park (2013).

Policy interventions are essential, including localized weather forecasting, climate-resilient vending infrastructure, and financial safety nets for informal workers. Vendor-driven strategies, such as portable shelters and flexible working hours, should also be encouraged.

While the study's sample size and reliance on self-reported perceptions have limited the generalizability, its findings highlight the urgent need for integrated policies and further longitudinal research linking objective climate and income data to better support Sri Lankan Street food vendors against increasing climate variability. Future studies should incorporate larger and more diverse samples to improve generalizability and strengthen the statistical power of the findings.

REFERENCES

- Bhowmik, S.K. (2005). *Street Vendors in Asia: A Review*, 40(22). <http://www.jstor.org/stable/4416705>
- Bhowmik, S.K., & Saha, D. (2012). *Street Vending in Ten Cities in India*.
- Dodman, D., Sverdlik, A., Agarwal, S., Kadungure, A., Kothiwala, K., Machemedze, R., & Verma, S. (2023). Climate change and informal workers: Towards an agenda for research and practice. *Urban Climate*, 48. <https://doi.org/10.1016/j.uclim.2022.101401>
- Karunaratna, A.C. & Tjandra, N.C. (2021). A Glance into the Life of Street Vendors in Sri Lanka: An Overview. *Sri Lanka Journal of Marketing*, 7(2), 169 - 193. <https://doi.org/10.4038/sljmuok.v7i2.69>
- Lund, F., & Naidoo, R. (2016). The changed world of work. In *New Solutions* (Vol. 26, Issue 2, pp. 145–154). SAGE Publications Ltd. <https://doi.org/10.1177/1048291116653053>
- Onyango, J.O. (2012). *Dynamics of Street Vending Phenomenon in the Kisumu Municipality, Kenya*. <https://www.researchgate.net/publication/381096422>
- Samaraweera, W.G.R.L., Dharmadasa, R.A.P.I.S., Kumara, P.H.T., & Bandara, A.S.G.S. (2024). Evidence of Climate Change Impacts in Sri Lanka - A Review of Literature. *Sri Lanka Journal of Economic Research*, 11(2), 69 - 94. <https://doi.org/10.4038/sljer.v11i2.205>
- Sverdlik, A., Kothiwala, K., Kadungure, A., Agarwal, S., Machemedze, R., Verma, S., & Loewenson, R. (2024). Understanding the interplay of occupational, public health, and climate-related risks for informal workers: A new framework with findings from Zimbabwe and India. *Social Science and Medicine*, 348. <https://doi.org/10.1016/j.socscimed.2024.116750>



Factors Influencing Student Engagement in Learning Management System: A Case Study

Somarathna WNM¹, Dharmawansa AD¹

^{1,2}Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka
manoj64chathuranga@gmail.com¹

ABSTRACT

Learning Management Systems (LMSs) are essential in higher education, yet achieving sustained student engagement remains a major challenge. Understanding the factors that influence engagement is critical for optimizing LMS effectiveness and improving learning outcomes. This study examined the impact of four factors; Perceived Ease of Use, Perceived Usefulness, Content Interactivity, and Institutional Support with LMS Policy on student engagement with the existing LMS at a higher education institution. A quantitative case study was conducted using structured questionnaires administered to 203 students. Data were analysed through multiple linear regression, which revealed that the model explained 71.1% of the variance in student engagement ($p < 0.001$). All four predictors were significant, with Content Interactivity emerging as the strongest contributor ($\beta = 0.300$), followed by Perceived Usefulness ($\beta = 0.284$), Institutional Support & LMS Policy ($\beta = 0.224$), and Perceived Ease of Use ($\beta = 0.138$). These findings indicate that engagement is driven not only by ease of use but also by the perceived value of the system, supportive institutional policies, and, most importantly, interactive learning content. The study recommends that higher education institutions enhance content interactivity, emphasize system usefulness, and strengthen support mechanisms to improve student engagement and maximize LMS effectiveness.

KEYWORDS: Content Interactivity, Institutional Support, Learning Management System (LMS), Perceived Ease of Use, Perceived Usefulness, Student Engagement.

1. INTRODUCTION

Learning Management Systems (LMS) are widely used in higher education to improve accessibility, flexibility, and student-centered learning. While LMS platforms can enhance engagement and academic performance, many institutions face challenges with low student participation. This issue is particularly relevant in Sri Lanka, where private higher education institutions are adopting paid LMS platforms to meet global standards. Institutional feedback suggests that the LMS is under-utilised and students mainly access lecture slides or assignments. This highlights the need to identify factors influencing engagement and develop strategies to enhance active and collaborative learning.

2. LITERATURE REVIEW

The Technology Acceptance Model (TAM), proposed by Davis (1989), is a widely used framework explaining how users adopt digital systems. TAM emphasises two core constructs—Perceived Ease of Use (PEOU) and Perceived Usefulness (PU)—which influence behavioral intention and actual system use. In Learning Management Systems (LMS), these constructs shape students' perceptions of usability and academic value (Venkatesh & Davis, 2000). Thus, TAM provides a strong foundation for examining how PEOU, PU, content interactivity, and institutional support affect LMS engagement.

2.1 *Perceived Ease of Use (PEU) and Student Engagement*

Perceived Ease of Use (PEU) plays a critical role in student engagement with LMSs. User-friendly, intuitive systems encourage frequent use, while complex designs reduce motivation. Studies show that clear navigation, responsive layouts, and the absence of technical barriers significantly increase student participation (Panergayo & Aliazas, 2021; Buil et al., 2019). Thus, PEU lowers barriers to learning and positively influences engagement.

2.2 *Perceived Usefulness (PU) and Student Engagement*

Perceived Usefulness (PU) reflects how students view an LMS's contribution to their learning performance and productivity. Studies show that when students perceive academic benefits such as flexible learning, timely feedback, and better resource access, they are more likely to adopt and continue using the system. PU not only drives initial acceptance but also deeper and sustained engagement, often influencing behavioral intention more strongly than ease of use (Mamun & Lawrie, 2023; Panergayo & Aliazas, 2021; Siregar, 2022). This highlights that student engagement is primarily shaped by the perceived academic value of the LMS.

2.3 *Content Interactivity and Student Engagement*

Interactive LMS features such as multimedia lessons, gamified quizzes, forums, and collaborative tools foster active participation and deeper learning. Studies show that interactivity enhances

academic performance, satisfaction, and self-directed learning while sustaining attention and motivating students to engage longer (Suhadi & Mustaffa, 2021; Mamun & Lawrie, 2023; Fernando et al., 2021). Thus, content interactivity is a key factor that significantly boosts student engagement beyond passive use.

2.4 *Institutional Support and LMS Policy and Student Engagement*

Institutional support, including training, technical assistance, mentoring, and infrastructure, is vital for sustaining LMS engagement. Clear policies further strengthen adoption by providing consistent guidelines and embedding LMS use into academic practices. Studies show that tailored onboarding and mentoring reduce adoption barriers, while integrating LMS activities into assessments promotes active use (Ahmad et al., 2021; Buil et al., 2019). Thus, institutional support and policy are positively associated with student engagement.

The literature indicates that LMS engagement is influenced by both technological and institutional factors. Perceived Ease of Use reduces access barriers, Perceived Usefulness emphasizes academic value, Content Interactivity fosters motivation and participation, and Institutional Support ensures continuity. Together, these factors interact dynamically, forming a solid foundation for the study's hypotheses that all four variables significantly predict student engagement with LMS platforms.

3. METHODOLOGY

3.1 Conceptual Framework

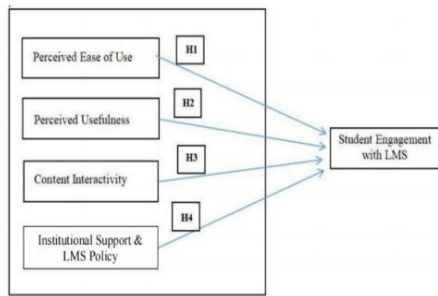


Figure 03 : Conceptual Framework

This study adopted a quantitative, explanatory research design to investigate the factors influencing student engagement with the Learning Management System (LMS) at higher education Academy. A case study approach was selected as it allows for in-depth analysis within a specific institutional context. The explanatory design was appropriate, as the study sought to test hypothesized relationships between independent variables (Perceived Ease of Use, Perceived Usefulness, Content Interactivity, and Institutional Support & LMS Policy) and the dependent variable (Student Engagement with LMS).

The research model of this study was developed to investigate the factors that influence student engagement with the Learning Management System (LMS) at Selected Academy. Based on the review of prior empirical studies on LMS adoption and engagement, four independent variables were identified as potential predictors of student engagement:

1. Perceived Ease of Use (PEOU)
2. Perceived Usefulness (PU)
3. Content Interactivity (CI)
4. Institutional Support & LMS Policy (ISLMSP)

Data were collected through an online questionnaire using a five-point Likert scale, with constructs adapted from validated studies. Responses from students were analyzed in SPSS. Descriptive statistics summarized demographics, Pearson's correlation tested initial relationships, and multiple regression assessed the effects of Perceived Ease of Use, Perceived Usefulness, Content Interactivity, and Institutional Support & LMS Policy on Student Engagement. Standardized beta coefficients measured predictor strength, with significance set at $p < 0.05$. This approach identified the key factors influencing engagement and tested the research hypotheses.

4. DATA ANALYSIS AND RESULTS

Table 1: Multiple Linear Regressions

R-value	R-Square	Adjusted R Square	Standard Error
0.843 ^a	0.711	0.705	0.4534

Table 2: Coefficient Table

Variable	Beta Value	T Value	P Value	Results
PEOU	0.138	2.257	0.012	supported
PU	0.284	3.926	0.000	supported
CI	0.300	4.696	0.000	supported
ISLMSP	0.224	3.802	0.000	supported

4.1 Explanatory Text

The regression analysis showed that all four factors significantly influenced LMS engagement. Content Interactivity

was the strongest predictor, underscoring the importance of quizzes, multimedia, and forums in motivating students. Perceived Usefulness also had a strong impact, while Institutional Support and Policy highlighted the need for training and supportive structures. Perceived Ease of Use had the weakest but still significant effect, showing that simple navigation supports engagement. Overall, student engagement is driven by interactivity, perceived value, and institutional support, with interactive features being the key driver.

4.2 Testing of Hypotheses

H1: Perceived Ease of Use has a significant positive effect on student engagement with the LMS.

The analysis shows that Perceived Ease of Use has a positive and statistically significant effect on student engagement ($\beta = 0.138$, $p = 0.012$). This means that when the LMS is user-friendly and easy to navigate, students are more likely to engage. However, the effect is relatively weak compared to other predictors, suggesting that ease of use mainly lowers barriers to access rather than driving strong or sustained engagement. H1 is supported.

H2: Perceived Usefulness positively influences student engagement with the LMS.

Perceived Usefulness has a strong, statistically significant positive effect on student engagement ($\beta = 0.284$, $p < 0.001$). Students are more likely to engage with the LMS when they believe it enhances their learning outcomes, efficiency, and academic performance. This indicates that usefulness is a central

motivator, with a stronger influence than ease of use. H2 is strongly supported.

H3: Content Interactivity significantly enhances student engagement with the LMS

Content Interactivity is the strongest predictor of student engagement ($\beta = 0.300$, $p < 0.001$). Interactive elements such as quizzes, gamified activities, multimedia lessons, and discussion forums significantly enhance participation by making the learning process dynamic and engaging. This finding shows that interactivity is essential not just an add on for sustaining meaningful engagement. H3 is very strongly supported.

H4: Institutional Support & LMS Policy positively affect student engagement with the LMS.

Institutional Support & LMS Policy has a significant positive impact on student engagement ($\beta = 0.224$, $p < 0.001$). When institutions provide training, technical assistance, clear policies, and incentives, students are more likely to engage consistently with the LMS. This shows that engagement depends not only on system design but also on the supportive organizational environment. H4 is supported.

5. DISCUSSION

The regression analysis confirmed that all four factors; Perceived Ease of Use, Perceived Usefulness, Content Interactivity, and Institutional Support & LMS Policy significantly enhance student engagement with the LMS. Content Interactivity was the strongest predictor, emphasizing the importance of quizzes, gamified activities, and discussion

forums. Perceived Usefulness also had a strong effect, as students are more engaged when they recognize clear academic benefits. Institutional Support reduced adoption barriers through training and policies, while Ease of Use had a weaker but supportive role. Overall, engagement is most effectively improved through interactivity, usefulness, and institutional backing.

6. CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

The study found that Perceived Ease of Use, Perceived Usefulness, Content Interactivity, and Institutional Support all significantly influenced LMS engagement. Content Interactivity was the strongest predictor, followed by Usefulness and Institutional Support, while Ease of Use showed a weaker effect. All research objectives were met, and all hypotheses were supported.

6.2 Recommendations

1. Enhance Interactive Content:

Institutions should prioritize the development of interactive and engaging learning materials, such as multimedia lessons, modified activities, and discussion forums, to stimulate active student participation.

2. Emphasize Perceived Usefulness:

Course designers and instructors should ensure that LMS content is directly linked to learning outcomes, demonstrating clear academic value to encourage stronger engagement.

3. Strengthen Institutional Support:

Institutions should provide consistent technical support, training sessions, and awareness programs to build student confidence in using the LMS. Clear

policies should be established to encourage continuous usage.

4. Improve User Experience:

Although ease of use had a weaker effect, ensuring a user-friendly interface and simplified navigation remains important for minimizing barriers and frustrations. Continuous system updates and usability testing are recommended.

6.3 Future Research Directions

Further studies could expand the model by including additional factors such as motivation, peer collaboration, or instructor involvement. Conducting comparative studies across different institutions or educational contexts would also enhance the generalizability of the findings.

REFERENCES

- Ahmad, A., et al. (2021). Bridging support and technology: Exploring the role of institutional support in shaping LMS adoption practices among vocational higher education academicians. *International Journal of Research and Innovation in Social Science*
- Buil, I., Martínez, E., & Matute, J. (2019). Transformational leadership and employee performance: The role of identification, engagement, and proactive personality. *International Journal of Educational Technology in Higher Education*, 16(1), 120
- Fernando, W. M. A. P. S., Sisindra, K. S.P., Liyanage, N.G.L.S.J., & Nawastheen, F.M. (2021). Factors affecting student engagement in online courses in Bachelor of Education programmes.
- Suhadi, A.Y., & Mustaffa, N. (2021). Online learning makes student perform better: A quantitative study of interactivity during class and academic performance among USIM students. *Al-i'lam - Journal of Contemporary Islamic Communication and Media*, 1(1), 21 - 35.
- Mamun, M.A., & Lawrie, G. (2023). Student-content interactions: Exploring

behavioural engagement with self-regulated inquiry-based online learning modules. *Smart Learning Environments*, 10(1), 1 - 15.

- Siregar, H.S. (2022). Perceived usefulness and perceived ease of use of online learning for Islamic Religious Education teachers. *Jurnal Pendidikan Islam*, 8(2), 233 - 245.
- Panergayo, A.A., & Aliazas, J.V. (2021). Students' Behavioral Intention to Use Learning Management System: The Mediating Role of Perceived Usefulness and Ease of Use. *International Journal of Information and Education Technology*, 11(11), 538 - 545.



Evaluating the Efficiency of Customs Clearance Procedures in Sri Lanka's Import and Export Industry

Krishnarathna BVNR¹, Jothirathne PAAU²

^{1,2} Department of Industrial Management Faculty of Applied Sciences, Wayamba University of Sri Lanka
nisanduniroshani@gmail.com

ABSTRACT

The efficiency of customs clearance procedures plays a pivotal role in shaping the competitiveness of international trade, particularly in developing economies for instance Sri Lanka. This study aims to evaluate the operational effectiveness of customs clearance processes within the country's import and export sector, focusing on delays, documentation challenges, and stakeholder experiences. A structured questionnaire was distributed among key stakeholders, including customs officers, importers, exporters, and clearing agents, generating quantitative data for analysis. Descriptive statistics and Chi-square tests were applied using SPSS to identify relationships between procedural bottlenecks, documentation issues, and the time required for clearance. Findings revealed that delays were primarily attributed to incomplete documentation, lack of automation in certain stages, and coordination gaps among stakeholders. Moreover, the study highlights the positive impact of digitalization initiatives while emphasizing the need for policy reforms and enhanced inter-agency collaboration. By integrating empirical evidence with stakeholder perspectives, this research provides actionable insights for streamlining customs operations and fostering trade facilitation in Sri Lanka.

KEYWORDS: Customs clearance, Digitalization, Documentation challenges, Policy reforms, Trade facilitation

1. INTRODUCTION

International trade is a key contributor to Sri Lanka's economy, with imports ensuring the availability of essential goods and exports generating foreign exchange earnings. The efficiency of customs clearance procedures including documentation, HS code classification, valuation, scanning, and destuffing is critical for smooth trade flows. However, delays in these processes increase costs, extend delivery times, and weaken competitiveness. Despite efforts to modernize customs operations through digital platforms like ASYCUDA stakeholders including importers, exporters and clearing agents continue to encounter delays, particularly in documentation, HS code disputes, valuation, and scanning. These challenges

highlight the need for a systematic evaluation of the customs clearance system to identify bottlenecks and assess the effectiveness of digitalization in enhancing efficiency.

This study aims to evaluate the operational efficiency of Sri Lanka's customs clearance process by examining key sources of delay and assessing impact of digital tools on process performance improving efficiency. Despite ongoing modernization efforts such as ASYCUDA and partial digitalization, Sri Lanka's customs clearance process continues to face delays, documentation issues, HS code disputes, and coordination challenges. However, existing studies mainly focus on macro-level trade facilitation and policy analysis, with limited empirical evidence from

stakeholders who directly engage in day-to-day clearance activities. There is a clear gap in understanding the operational bottlenecks experienced by importers, exporters, wharf clerks, and customs officers themselves.

Therefore, this study fills this gap by providing an evidence-based assessment of operational inefficiencies using stakeholder data, statistical analysis, and evaluation of digital system performance.

2. LITERATURE REVIEW

Sri Lanka's trade system faces inefficiencies that hinder competitiveness, including excessive paperwork, documentation delays, valuation disputes, scanning bottlenecks, and weak inter-agency coordination (Jayaratne, n.d.). Digital systems such as ASYCUDA and Electronic Data Interchange have been partially adopted, while the National Single Window remains incomplete (World Bank, 2020). These issues increase costs, reduce transparency, and disadvantage smaller traders and foreign investors.

Aligning customs procedures with international standards like the Revised Kyoto Convention and adopting advanced technologies such as blockchain and machine learning can improve efficiency (WCO, 2019; WTO, 2021). Existing studies often overlook perspectives and experiences of key stakeholders involved in daily operations. This study addresses this gap by integrating insights from importers, exporters, and customs officers to propose practical reforms that enhance automation, coordination, and trade facilitation in Sri Lanka.

3. METHODOLOGY

This study adopts a quantitative research design to investigate delays and operational challenges in Sri Lanka's customs clearance process. The focus is on documentation errors, valuation disputes, scanning and destuffing delays, and the use of digital systems such as ASYCUDA. The conceptual framework links independent variables (stakeholder role, experience, use of automation) with dependent variables (frequency of delays, procedural challenges, satisfaction with ASYCUDA), assuming that inefficiencies and system performance directly influence clearance speed and associated transaction cost.

The primary data were collected from 40 stakeholders engaged in import-export operations, including importers, exporters, wharf clerks, freight forwarders, and customs officials. A purposive sampling approach was used to ensure that respondents possessed direct experience with customs procedures. Data were collected via a structured questionnaire administered online and in person, including close-ended questions for quantitative analysis and open-ended questions for qualitative insights. The questionnaire was pre-tested with a small group of trade professionals to confirm clarity, consistency, and content validity prior to full-scale data collection.

Descriptive statistics were employed to summarise frequency of delays, automation usage, and stakeholder experiences. Cross-tabulations and Chi-square tests examined relationships between variables, while thematic coding captured qualitative suggestions and causes of delays. Participation was voluntary. Participation was voluntary

and respondents were informed about the academic purpose and confidentiality of the research. All responses were anonymized, and informed consent was obtained prior to participation. While previous research discusses general trade facilitation challenges in Sri Lanka, very few studies systematically examine how specific procedural issues such as documentation errors, HS code disputes, and ASYCUDA failures directly affect clearance time at the operational level. Moreover, stakeholder experiences are rarely integrated into empirical analysis.

This study addresses these gaps by combining quantitative data from frontline stakeholders with statistical analysis to determine how documentation challenges, digital system issues, and inter-agency processes influence clearance efficiency.

4. RESULTS & DISCUSSION

4.1 Descriptive Statistics

The survey results indicate that missing or incomplete documents are the most significant cause of delays in customs clearance, with 50% of respondents reporting this issue. Other major causes include manual data entry delays (43%), lack of knowledge on requirements (23%), and wrong HS codes (20%). Missing digital copies were less of an issue, reported by only 15% of respondents.

Table 4.1 - Descriptive Statistic for document delay

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Missing or incomplete documents	40	0	1	.50	.506
Wrong HS code	40	0	1	.20	.405
Missing digital copies	40	0	1	.15	.362
Manual data entry delays	40	0	1	.43	.501
Lack of knowledge on requirements	40	0	1	.23	.423
Valid N (listwise)	40				

Documentation Delay

Minimum and Maximum values represent the binary coding used for analysis (0 = No, 1 = Yes). For each variable, respondents indicated whether they experienced a specific issue. These values allow SPSS to compute descriptive statistics.

The primary delays in the clearance process are Long Room delays (58%) and documentation issues (48%). ASYCUDA/digital scanning issues affected 32% of respondents, while valuation disputes and destuffing delays were less common, each affecting 15-18% of the respondents.

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Documentation issues	40	0	1	.48	.506
Long Room delay	40	0	1	.58	.501
Valuation disputes	40	0	1	.15	.362
ASYCUDA or digital problems	40	0	1	.32	.474
Scanning delay	40	0	1	.18	.385
Destuffing delay	40	0	1	.15	.362
Valid N (listwise)	40				

Delays

A Chi-Square test revealed a significant association between respondents' roles and their experience with ASYCUDA-related issues ($\chi^2 = 12.111$, $p = 0.002$). This suggests that different roles within the customs process experience varying levels of system problems, emphasizing the importance of role-specific solutions.

Chi-Square Tests			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	12.111 ^a	2	.002
Likelihood Ratio	12.664	2	.002
N of Valid Cases	40		

a. 2 cells (33.3%) have expected count less than 5. The minimum expected count is .85.

Chi Square Test

4.2 Frequencies

The analysis of 40 valid survey responses revealed several recurring challenges in Sri Lanka's customs

clearance process. The majority of participants (69.2%) frequently experienced delays due to incomplete or incorrect documentation, while 23.1% reported no delays. Although 64.1% of respondents felt confident in identifying correct HS codes, 35.9% lacked confidence, indicating knowledge gaps. HS code errors were a universal issue, with all participants reporting extra checks or penalties. Similarly, scanning procedures and destuffing operations caused delays for all valid respondents. Technical issues with the ASYCUDA system were also widespread, affecting all participants who responded to this question. These findings highlight that documentation inefficiencies, procedural errors, and system limitations are prevalent across the customs clearance process, emphasizing the need for targeted training, process improvements, and enhanced digital system reliability to increase operational efficiency.

4.3 Cross-tabulation

The cross-tabulation analysis explored the relationship between participants' experience in customs clearance and the documentation issues they encounter. Staff members frequently reported missing or incomplete documents, while wharf clerks noted delays in manual data entry. Exporters and importers experienced fewer issues, although HS code errors were observed among some importers. Respondents with less than one year of experience most often faced missing or incomplete documents, whereas those with over three years of experience reported manual entry delays and missing digital copies. Pearson Chi-Square tests indicated no statistically significant associations at the 0.05 level,

though Likelihood Ratios suggested potential links. Many cells had low expected counts, limiting reliability. Overall, documentation challenges were observed across all roles and experience levels, highlighting the need for targeted training and process improvements to enhance efficiency in customs operations.

5. DISCUSSION

The findings indicate that inefficiencies in Sri Lanka's customs clearance process are systemic, cutting across roles and levels of experience. Documentation errors, HS code disputes, and technical failures in ASYCUDA are the most persistent challenges. These outcomes are consistent with previous research identifying fragmented procedures, weak interagency coordination, and limited digital integration as key barriers to trade facilitation (De Mel et al., 2011; Edirisinghe & Jayathilake, 2013; World Bank, 2020).

The prevalence of HS code-related issues highlights insufficient technical knowledge and training, echoing international studies that emphasize the importance of capacity building for customs efficiency (WCO, 2019; WTO, 2021). The absence of significant statistical relationships between experience and documentation issues implies that even experienced personnel are constrained by structural inefficiencies rather than individual skill deficiencies. This highlights the need for systemic reforms rather than ad hoc operational fixes.

The results further demonstrate that digital platforms like ASYCUDA have not yet delivered their full potential due to technical limitations and inconsistent user

adoption. Enhancing the platform's stability, interoperability, and user training could substantially reduce delays. In broader terms, the findings reaffirm that efficiency improvements require both technological upgrades and institutional alignment across agencies.

6. CONCLUSION AND RECOMMENDATIONS

This study found that Sri Lanka's customs clearance faces persistent delays due to documentation issues, HS code errors, scanning and destuffing inefficiencies, and technical problems with ASYCUDA. These challenges affect all stakeholders, regardless of experience, highlighting systemic and knowledge-based gaps. To improve efficiency, it is recommended to standardize and digitize documentation, provide HS code training and reference tools, upgrade ASYCUDA reliability, streamline scanning and destuffing processes, enhance inter-agency coordination, and implement continuous monitoring through surveys and performance indicators.

Adopting these measures can reduce delays, lower operational costs, and strengthen the competitiveness of Sri Lanka's import-export sector.

REFERENCES

- Aggarwal, A. (2006). Performance of export processing zones: A comparative analysis of India, Sri Lanka, and Bangladesh. *Journal of Instafalg Institute*. Retrieved from https://www.academia.edu/68073628/Performance_of_export_processing_zones_A_comparative_analysis_of_India_Sri_Lanka_and_Bangladesh
- Edirisinghe, L. (2013). Cross-border logistics performance in Sri Lanka: The way forward. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2965973
- Edirisinghe, L., & Jayathilake, S. (2013). Frontier logistics performance in Sri Lanka: The roleplay of the customs. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssm.2965967>
- Ilangasekara, A., & Premaratne, W. (2018). Underperforming dimension of Logistics Performance Index (LPI) in Sri Lanka. *International Journal of Advanced Research*, 6(11), 585 - 599. <https://doi.org/10.21474/ijar01/8034>
- Jayaratne, S.M. (n.d.). Enhancing export competitiveness through trade facilitation: The experience of Sri Lanka.
- Jayaratne, S., Premaratne, D., & Wijayasiri, J. (2016). Trade and transportation audit: Sri Lanka Institute of Policy Studies. Retrieved from <https://www.ips.lk/wp-content/uploads/2017/01/Trade-Transportation-Audit.pdf>
- Judith, S. (2024). Enhancing export supply chains: A comparative analysis of Sri Lankan exports to Finland. Retrieved from https://www.theseus.fi/bitstream/handle/10024/861345/Judith_Salomi.pdf?sequence=2
- De Mel, D., Jayaratne, S., & Premaratne, D. (2025). Improving import-export procedures and processes in Sri Lanka. *Econstor*, <http://hdl.handle.net/10419/64258>



Governance, Risk and Compliance (GRC) Readiness Assessment Factors for Sri Lankan Banking and Finance Sector: A Systematic Literature Review

Rathnayaka RMSO¹, Wickramarachchi APR²

^{1,2} *Department of Industrial Management, Faculty of Science, University of Kelaniya*
rathnay-im20094@stu.kln.ac.lk

ABSTRACT

Governance, Risk, and Compliance (GRC) frameworks are vital for improving governance, risk management, compliance, and trust in banking. In Sri Lanka, economic volatility, Central Bank directives, digital growth, and cybersecurity threats underscore GRC readiness needs post-2022 crisis. This Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)-guided systematic literature review identifies tailored factors, addressing gaps in localized models influenced by hierarchical cultures and socio-economic constraints. Databases like IEEE Xplore, Scopus, and Google Scholar are searched for 2010–2025 peer-reviewed studies, yielding 51 sources. Thematic synthesis highlighted six factors: Governance Structure and Leadership (e.g., board oversight), Risk Management Practices, Compliance and Regulatory Alignment, Technological Capability and Infrastructure, Human Resources and Capacity, and Financial Stability and Resources. Challenges include functional silos, legacy systems, cultural barriers, and Central Bank of Sri Lanka (CBSL)-specific regulations. Findings provide a foundation for a customized GRC readiness model to mitigate risks, ensure compliance, and enhance resilience.

KEYWORDS: Compliance, Governance, Readiness, Risk Management

1. INTRODUCTION

Governance, Risk, and Compliance (GRC) frameworks strengthen governance, the structures ensuring accountability and ethical oversight, while enhancing risk management, the process of identifying, and mitigating threats, and compliance, adherence to regulations and standards. These build stakeholder trust in banking sector (Shahim, Batenburg, & Versendaal, 2012; Makaš, 2023). In Sri Lanka, economic instability, Central Bank of Sri Lanka (CBSL) directives, digital banking expansion, and cybersecurity threats heighten GRC readiness, the preparedness to implement such frameworks amid cultural and socio-economic factors shaping adoption (Wanniarachchige & Suzuki, 2010; Ranawaka, Khatibi, Azam, & Tham, 2023). This Systematic

Literature Review (SLR) reveals key GRC readiness factors, challenges, and indicators for a context-specific model in Sri Lanka's post-crisis digital landscape (Barafort, Mesquida, & Mas, 2018; Ranawaka et al., 2023).

2. METHODOLOGY

2.1 Search Strategy and Databases

A structured search strategy had used to find relevant studies on GRC readiness and implementation in banking/finance, adaptable to Sri Lanka. Searches across databases like IEEE Xplore, SpringerLink, ScienceDirect, Scopus, Web of Science, DOAJ, Google Scholar, and regional repositories, employed keywords like "Governance, Risk and Compliance" OR "GRC," "GRC readiness" OR "adoption," "banking sector," and "Sri Lanka" OR "developing countries," with Boolean operators.

Filters included 2010–2025 publications, English language, and peer-reviewed types. Policy sources like Central Bank of Sri Lanka (CBSL) guidelines have included for contextual insights. The last search occurred in August 2025, yielding 228 records after initial identification.

2.2 Inclusion and Exclusion Criteria

Inclusion criteria focused on peer-reviewed articles, conference proceedings, and reputable reports from 2010–2025 in English, emphasizing GRC readiness or implementation in banking/finance or adaptable from other sectors. Geographic priorities included Sri Lanka, South Asia, emerging economies, and global transferable studies. Exclusion criteria have eliminated non-empirical work, unrelated studies, non-English publications, pre-2010 literature, and those without full-text access, ensuring relevance to modern GRC developments.

2.3 PRISMA Screening Process

Following PRISMA guidelines, 210 records underwent multi-stage screening using Zotero. As shown in Figure 1, Stage 1 retained 105 studies relevant to GRC. Stage 2 excluded non-empirical or irrelevant works, leaving 62. Stage 3 confirmed alignment with criteria, resulting in 511 included studies.

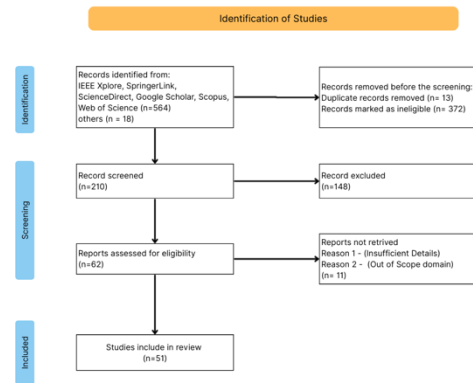


Figure 1: PRISMA Screening Process

2.4 Data Synthesis

A thematic synthesis of 51 studies found six key GRC categories: governance, risk management, compliance, technology, human resources, and financial stability, analyzed within Sri Lanka’s regulatory and economic context.

3. RESULTS

Through the thematic synthesis, six key GRC readiness factors have derived: Governance and Leadership, Risk Management, Compliance and Regulation, Technology and Infrastructure, Human Resources, and Financial Resources. The table 1 below outlines these factors and their indicators influencing GRC readiness in the banking and finance sector.

Table 2: Factors and Indicators GRC Readiness

Factor	Indicator
Governance Structure and Leadership	Board oversight and independence, Organizational culture and leadership support, Ethical leadership and transparency, Stakeholder engagement and communication, Governance structure effectiveness
Risk Management Practices	Risk management effectiveness, Credit and liquidity risk management, Operational and third-party risk management, Systemic risk and stress testing
Compliance and Regulatory Alignment	Regulatory compliance and alignment, Compliance function and reporting, Anti-money laundering (AML) and terrorist financing, Audit and consumer protection, Regulatory adaptation
Technological Capability and Infrastructure	Technological maturity and integration, Cybersecurity and resilience, Data management and interoperability, Technology adoption
Human Resources and Capacity	Human skills, training, and awareness, Internal expertise and capacity building, Employee awareness and training programs, Resistance to change and cultural barriers
Financial Stability and Resources	Capital adequacy and financial stability, Resource constraints and availability, Process maturity and optimization, Continuous improvement and proactive planning

4. DISCUSSION

Robust governance, risk management, compliance, technology, human resources, and financial stability collectively drive GRC readiness in Sri Lanka's banking sector. Strong governance ensures strategic alignment, and ethical conduct, supported by independent oversight and transparency (Shahim et al., 2012; Kulmie & Ibrahim, 2024; Makaš, 2023; Ranawaka et al., 2023; Suša Vugec, Spremić, & Bach, 2017). Risk management guided by ISO 31000 and CBSL frameworks enhances visibility, resilience, and strategic alignment through stress testing and NPL monitoring (Barafort et al., 2018; Gozman

& Currie, 2015; Central Bank of Sri Lanka [CBSL], 2014; Losiewicz-Dniestrzanska, 2015; Oyeniyi, Igwe, Ofodile, Ewim, & Olorunyomi, 2021; Victor, Moronkunbi, Oyedeji, Victor, & Samuel, 2024). Compliance with AML/CFT and COSO/ISO 31000 strengthens transparency and competitiveness under CBSL guidance (Shekhar, 2025; Pricillia, 2021; Ayodele & Adelaja, 2024). Technological infrastructure enhances automation and resilience through IT-security integration and AI/blockchain adoption aligned with COBIT and ISO 27001 (De Smet & Mayer, 2016; Victor et al., 2024; Govindji, Peko, & Sundaram, 2018, p16;

Chen et al., 2022; Jayasundara & Wickramarachchi, 2024; Barafort et al., 2018). Human resources improve readiness via certification and awareness initiatives addressing resistance and capability gaps (Victor et al., 2024; Makaš, 2023; Shekhar, 2025; Al-Tamimi et al., 2016; Rathakrishnan & Baskar, 2024). Financial stability underpins GRC through Basel III-aligned capital adequacy and efficient allocation (Makaš, 2023; Al-Tamimi et al., 2016; Ahmad & Akter, 2019; Victor et al., 2024; Karthick, Prabhakaran, Banu, & Senthil Kumar, 2023; Nissen & Marekfa, 2013).

Implementation challenges include functional silos, legacy systems, and cultural barriers that hinder integration and collaboration (Victor et al., 2024; Ranawaka et al., 2023; Shekhar, 2025; Makaš, 2023). Leadership commitment and training are key to overcoming these obstacles.

Sri Lanka's regulations feature CBSL directives on governance/risk committees (Ranawaka et al., 2023) and cybersecurity aligned with ISO/IEC 27001 (Fairouz & Wickramasinghe, 2019). Enforcement issues, include uneven implementation, demand capacity building (Fairouz & Wickramasinghe, 2019; Wanniarachchige & Suzuki, 2010).

5. CONCLUSION

This SLR has found six key GRC readiness factors—governance, risk management, compliance, technology, human resources, and financial stability—specific to Sri Lanka's banking sector. Supported by indicators like AML training and capital adequacy, these factors address challenges such as silos and legacy systems. Aligned with CBSL and global standards, the study provides a

contextual foundation for an integrated GRC readiness model to strengthen resilience and compliance.

REFERENCES

- Ahmad, S., & Akter, R. (2019). A study on the implementation of Basel III: Bangladesh perspective. *Academy of Strategic Management Journal*, 18(6).
- Al-Tamimi, H.A.H., Warsame, M.H., & Duqi, A. (2016). Readiness of the UAE banks for the implementation of Basel III. *International Journal of Economics and Business Research*, 11(4), 303-316. doi:10.1504/IJEER.2016.077017
- Ayodele, O.O., & Adelaja, A.O. (2024). Cybersecurity threats in financial institutions: Implications for governance, risk management, and compliance. *Journal of Economics, Finance and Management Studies*, 7(8), 3736-3744. doi:10.47191/jefms/v7-i8-12
- Barafort, B., Mesquida, A.L., & Mas, A. (2018). ISO 31000-based integrated risk management process assessment model for IT organizations. *Journal of Software: Evolution and Process*, 31(1). doi:10.1002/smr.1984
- Central Bank of Sri Lanka. (2014). Guidelines on stress testing of licensed commercial banks and licensed specialized banks. Retrieved from https://www.cbsl.gov.lk/sites/default/files/cbslweb_documents/laws/cdg/bsd_Guidelines_Stress_testing_LCBs_LSBs_0.pdf
- Chen, X., Chang-Richards, A., Ling, F. Y.Y., Yiu, T.W., Pelosi, A., & Yang, N. (2022). Developing a readiness model and a self-assessment tool for adopting digital technologies in construction organizations. *Building Research & Information*, 241 - 256. doi:10.1080/09613218.2022.2136130
- De Smet, D., & Mayer, N. (2016). Integration of IT governance and security risk management: A systematic literature review. Proceedings of the International Conference on Information Society (i-Society), Dublin, Ireland. doi:10.1109/i-Society.2016.7854200
- Fairouz, H.M.M., & Wickramasinghe, C.N. (2019). Development of an information security framework for the deployment of electronic voting systems in Sri Lanka. *International Journal of*

GOVERNANCE, RISK AND COMPLIANCE (GRC) READINESS ASSESSMENT FACTORS FOR SRI
LANKAN BANKING AND FINANCE SECTOR: A SYSTEMATIC LITERATURE
REVIEW

- Information and Education Technology*, 9(11), 808 - 813. doi:10.18178/ijiet. 2019. 9.11.1307
- Govindji, S., Peko, G., & Sundaram, D. (2018). A context adaptive framework for IT governance, risk, compliance and security. doi:10.1007/978-3-319-77818-1_2
 - Gozman, D., & Currie, W. (2015). Managing governance, risk and compliance for post-crisis regulatory change: A model of IS capabilities for financial organizations. Proceedings of the 48th Hawaii International Conference on System Sciences, Kauai, HI, USA. doi:10.1109/HICSS.2015.555
 - Jayasundara, H., & Wickramarachchi, A. P.R. (2024). Critical factors affecting digital resilience in the Sri Lankan banking sector: A systematic review. *Journal of Science of the University of Kelaniya*, 17(2), 129-145. doi:10.4038/josuk.v17i2.8126
 - Karthick, V., Prabhakaran, J., Banu, P., & Senthil Kumar, U.S. (2023). Systematic literature review on GRC - A study on best practices and implementation strategy in GRC. *Samdarshi*, 16(4).
 - Kulmie, D.A., & Ibrahim, M.S. (2024). Bank corporate governance: Shield against fraud. *Journal of Ecohumanism*, 3(3), 1917 - 1932. doi:10.62754/joe.v3i3.3582
 - Losiewicz-Dniestrzanska, E. (2015). Monitoring of compliance risk in the bank. *Procedia Economics and Finance*, 26, 800 - 805. doi:10.1016/S2212-5671(15)00846-1
 - Makaš, A. (2023). Governance, risk and compliance frameworks applicability in the organizations. *International Journal of Science and Research Archive*, 10(2), 716 - 724. doi:10.30574/ijstra.2023.10.2.1024
 - Nissen, V., & Marekfi, W. (2013). Towards a research agenda for strategic governance, risk and compliance (GRC) management. doi:10.1109/CBI.2013.9
 - Oyeniyi, L.D., Igwe, A.N., Ofodile, O.C., Ewim, C.P.M., & Olorunyomi, T.D. (2021). Optimizing risk management frameworks in banking: Strategies to enhance compliance and profitability amid regulatory challenges. *International Journal of Science and Technology Research Archive*, 1(2), 113 - 119. doi: 10.53771/ijstra.2021.1.2.0052
 - Pricillia, N. (2021). Comparison of Indonesian banking regulation for integrated governance, risk management, compliance with its ISO counterparts. *RSF Conference Series Business Management and Social Sciences*, 1(5), 84 - 96. doi:10.31098/bmss.v1i5.455
 - Ranawaka, P., Khatibi, A., Azam, S. M. F., & Tham, J. (2023). An overview of corporate governance laws and regulations applicable to the banking industry in Sri Lanka. *International Journal of Research Publications*, 130(1). doi:10.47119/IJRP1001301820235306
 - Rathakrishnan, S., & Baskar, T. (2024). Fortifying financial integrity: Insights into fraud detection and prevention strategies across various financial companies in Sri Lanka from the perspectives of accountants and internal auditors in an analytical review. *International Journal of Research and Innovation in Social Science*, 8(6), 2168-2181. doi:10.47772/IJRISS.2024.806163
 - Shahim, A., Batenburg, R., & Versendaal, J. (2012). Governance, risk and compliance: A strategic alignment perspective applied to two case studies. *IFIP Advances in Information and Communication Technology*, 202 - 212. doi:10.1007/978-3-642-33332-3_19
 - Shekhar, C. (2025). A critical study of governance, risk, and compliance (GRC) control frameworks in India: A comparative perspective. *International Journal of Research Publication and Reviews*, 6(2), 806 - 813. doi:10.55248/gengpi.6.0225.0732
 - Suša Vugec, D., Spremić, M., & Bach, M. P. (2017). IT governance adoption in banking and insurance sector: Longitudinal case study of COBIT use. doi:10.18421/IJQR11.03-13
 - Victor, A.A., Moronkunbi, M. A., Oyedeji, O. C., Victor, P. O., & Samuel, S. A. (2024). The role of IT governance risk and compliance (IT GRC) in modern organizations. doi:10.51583/IJLTEMAS.2024.130607
 - Wanniarachchige, M.K., & Suzuki, Y. (2010). Bank competition and efficiency: The case of Sri Lanka. *Asia Pacific World*, 1(1), 117 - 131. doi:10.3167/apw.2010010108



A Study on Skill Gaps and Training Needs among Operational Employees in the Apparel Industry

Piyamantha NJ¹, Pallegedara A²

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka*
janath.priyamantha.92@gmail.com

ABSTRACT

The apparel industry plays a vital role in Sri Lanka's economy by generating exports and employment. With the increasing adoption of automation and lean manufacturing systems, operational employees are now required to possess more advanced technical and soft skills. However, a persistent mismatch exists between job requirements and employee competencies. This study investigates the prevailing skill gaps and training needs among operational employees in selected apparel manufacturing firms. A structured questionnaire was administered to 85 operational employees across Production, Quality, Technical, Engineering, Cutting, Embroidery, and Warehouse divisions. Quantitative data were analyzed using descriptive statistics, correlation, and regression analysis via SPSS. The findings reveal moderate levels of training effectiveness, with training relevance, methodology, and employee professionalism significantly reducing skill gaps. Conversely, training frequency and satisfaction showed no statistically significant impact. The study emphasizes the importance of practical, job-oriented, and blended learning-based training programs to enhance workforce capabilities. Recommendations are provided to guide apparel sector organizations in developing strategically aligned and department-specific training interventions.

Keywords: Skill gaps, Training needs, Apparel industry, Operational employees, Continuous learning

1. INTRODUCTION

The apparel sector is one of Sri Lanka's most significant contributors to export earnings and employment generation. Operational employees, who perform day-to-day production, quality, and support activities, directly influence efficiency, accuracy, and delivery performance. With technological advancements such as semi-automation, digital quality systems, and lean manufacturing practices, employees are increasingly required to handle sophisticated tools, problem-solving tasks, and continuous improvement activities.

However, despite industry-level investments in training, many organizations continue to report challenges such as low productivity, high

rework rates, machine-handling issues, and communication barriers. These issues often stem from unmet training needs and widening skill gaps. Existing literature highlights notable deficiencies in training relevance, practical exposure, and structured development pathways within Sri Lankan apparel firms. Yet, empirical evidence addressing department-level training effectiveness remains limited.

This study addresses this gap by examining the impact of training frequency, satisfaction, relevance, methodology, and employee professionalism on reducing skill gaps among operational employees.

2. LITERATURE REVIEW

2.1. Skill Gaps in the Apparel sector

A skill gap occurs when the competencies required for a job exceed

those possessed by employees (Mason, 2011). In the apparel sector, technological advancements demand updated technical knowledge, analytical thinking, and adaptability (Ekanayake & Wickramasinghe, 2020). Skill gaps negatively influence productivity, quality, and innovation capacity.

2.2. Training Needs Analysis (TNA)

Training Needs Analysis identifies knowledge and performance deficiencies and helps organizations design targeted interventions (Goldstein & Ford, 2002). Effective TNA improves employee satisfaction and performance as training directly aligns with job requirements (Noe, 2017).

2.3 Training and Organizational Learning Culture

A continuous learning culture supports upskilling, professional development, and adaptability (Marsick & Watkins, 2003). However, Sri Lankan apparel firms often face challenges related to outdated training materials and lack of practical exposure (Wickramasinghe & Perera, 2014).

Proposed Hypotheses

H1: Training satisfaction significantly reduces skill gaps.

H2: Training relevance significantly reduces skill gaps.

H3: Training frequency significantly reduces skill gaps.

H4: Training methodology significantly reduces skill gaps.

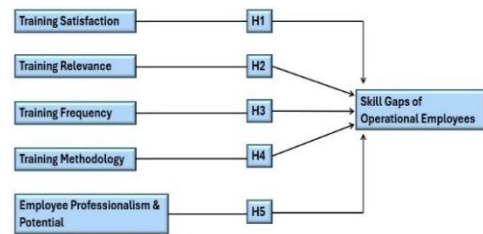
H5: Employee professionalism significantly reduces skill gaps.

3. METHODOLOGY

A quantitative research design was used to examine relationships between training-related factors and skill gaps among operational employees in the

apparel sector. Data were collected from 85 operational employees representing Production, Quality, Technical, Cutting, Engineering, Embroidery, Finished Goods Warehouse, and Raw Material Warehouse divisions. A structured questionnaire using five-point Likert scales measured all variables.

Figure 1: Conceptual Framework



All variables: training frequency, satisfaction, relevance, methodology, professionalism, and skill gaps were operationalized using established scales. Training frequency and satisfaction were adopted from Noe (2017) and Dessler (2020), training relevance from Goldstein & Ford (2002), methodology from Salas et al. (2012), and professionalism from Wickramasinghe & Perera (2014). Skill gap items were adapted from industry competency frameworks. Data were analyzed using descriptive statistics, correlation, and multiple regression in SPSS. Participation was voluntary and anonymous, with all ethical requirements met.

Table 1: Summary of Variables and Measurement Scales

Variable	Measurement Scale	Source
Training Frequency	Likert scale (1–5)	Adapted from Noe (2017)
Training Satisfaction	Likert scale (1–5)	Dessler (2020)

Training Relevance	Likert scale (1–5)	Goldstein & Ford (2002)
Training Methodology	Likert scale (1–5)	Salas et al. (2012)
Employee Professionalism	Likert scale (1–5)	Wickramasinghe & Perera (2014)
Skill Gaps	Likert scale (1–5)	Adapted from industry guidelines

4. RESULTS

4.1 Descriptive Statistics

Descriptive results indicated moderate training effectiveness across all variables. Mean values for training frequency (3.41), satisfaction (3.38), relevance (3.43), methodology (3.76), professionalism (3.72), and skill gaps (3.42) show generally consistent perceptions among employees. Higher mean values for methodology and professionalism suggest a preference for practical, skill-based learning approaches. Standard deviations ranged from 0.77 to 0.98, indicating uniformity across departments. Overall, findings show that although training opportunities exist, they remain insufficient to fully meet evolving operational needs.

4.2 Correlation Analysis

Correlation analysis revealed positive and significant associations between all independent variables and skill gap reduction. The strongest relationships were observed for professionalism ($r = 0.681$), training methodology ($r = 0.636$), and relevance ($r = 0.563$). Training satisfaction ($r = 0.307$) and frequency ($r = 0.341$) showed weaker but positive relationships. These results emphasize that training quality, relevance, and practical delivery has a greater impact

on reducing skill gaps than frequency or satisfaction.

4.3 Regression Analysis

Regression analysis confirmed that the model explains 57.5% of the variance in skill gaps (Adjusted $R^2 = 0.548$). The model was statistically significant ($F(5,79) = 21.34$, $p < 0.001$), indicating that the selected predictors collectively influence skill gap reduction. Training relevance ($\beta = 0.267$, $p = 0.013$), methodology ($\beta = 0.328$, $p = 0.001$), and professionalism ($\beta = 0.389$, $p = 0.000$) were significant predictors.

4.4 Hypothesis Testing

Results showed that H2, H4, and H5 were supported, confirming that relevant training content, effective delivery methods, and employee professionalism significantly reduce skill gaps. Training frequency ($\beta = -0.061$, $p = 0.513$) and satisfaction ($\beta = -0.094$, $p = 0.344$) were not significant, leading to rejection of H1 and H3. These findings highlight that skill development depends more on the relevance and quality of training than on the number of sessions or satisfaction levels.

5. DISCUSSION

The findings reveal that training programs in the apparel sector are moderately effective but require improvements in content and delivery. The strong links between training relevance, methodology, and professionalism with skill gap reduction highlight the need for practical, job-oriented programs that address real operational challenges. Employees who received relevant and hands-on training displayed higher confidence and adaptability, leading to fewer performance deficiencies. The results

align with prior studies emphasizing that interactive and continuous training enhances performance and learning transfer (Salas et al., 2012). The role of professionalism also stands out as a behavioral complement to technical skills, helping employees apply their knowledge effectively in real work environments. This supports the view that training effectiveness depends more on relevance, methodology, and professionalism than on frequency or satisfaction. From a managerial perspective, organizations should identify department-specific training gaps and adopt blended learning approaches that combine classroom, on-the-job, and digital methods. Regular evaluations, performance feedback, and mentorship programs can further strengthen professionalism, engagement, and learning outcomes. By prioritizing these measures, apparel manufacturers can effectively reduce skill gaps and build a more competent, adaptable workforce.

6. CONCLUSION AND RECOMMENDATIONS

The study concludes that while apparel industry employees appreciate available training opportunities, current programs require greater relevance, methodological diversity, and professional focus to effectively reduce skill gaps. The findings confirm that training relevance, methodology, and professionalism are the most influential factors in improving employee competence. Training initiatives should therefore move beyond frequency-based measures and concentrate on job-specific, practical, and blended approaches that connect learning outcomes with operational needs. Organizations are encouraged to conduct regular Training

Needs Analyses (TNA) to identify department-specific deficiencies and design tailored programs accordingly. Applying blended learning that combines classroom sessions, on-the-job exposure, and digital modules can enhance both engagement and retention, while continuous evaluation and follow-up ensure learning transfer to performance. Promoting professionalism through mentorship, peer learning, and structured feedback further strengthens employee accountability and growth. By integrating these strategies, apparel manufacturing organizations can close skill gaps, improve productivity, and build a more agile and future-ready workforce capable of adapting to evolving industry demands.

REFERENCES

- Asian Development Bank. (2014). *Innovative strategies in technical and vocational education and training for accelerated human resource development in South Asia: Sri Lanka*. Asian Development Bank. <https://www.adb.org/publications>
- Becker, B. E., & Huselid, M. A. (2006). Strategic human resources management: Where do we go from here? *Journal of Management*, 32(6), 898 - 925. <https://doi.org/10.1177/0149206306293668>
- Dessler, G. (2020). *Human resource management* (16th ed.). Pearson Education.
- Ekanayake, E.M.S., & Wickramasinghe, V. (2020). Skills development in the Sri Lankan apparel sector: Challenges and strategies. *Colombo Business Journal*, 11(2), 45 - 67. <https://doi.org/10.4038/cbj.v11i2.52>
- Goldstein, I.L., & Ford, J.K. (2002). *Training in organizations: Needs assessment, development, and evaluation* (4th ed.). Wadsworth.
- International Labour Organization. (2021). *Skills for a resilient youth: Supporting youth in the garment sector*.

A STUDY ON SKILL GAPS AND TRAINING NEEDS AMONG OPERATIONAL EMPLOYEES IN
THE APPAREL INDUSTRY

International Labour Organization.
<https://www.ilo.org>

- Jin, B., & Moon, H. (2006). The diamond approach to the competitiveness of Korea's apparel industry: Michael Porter and beyond. *Journal of Fashion Marketing and Management*, 10(2), 195 - 208. <https://doi.org/10.1108/1361202 061 06 67513>
- Marsick, V.J., & Watkins, K.E. (2003). Demonstrating the value of an organization's learning culture: The dimensions of the learning organization questionnaire. *Advances in Developing Human Resources*, 5(2), 132 - 151. <https://doi.org/10.1177/1523422303005 002002>
- Mason, G. (2011). What is the value of higher education in the labour market? *OECD Education Working Papers*, 54. OECD. <https://doi.org/1 0.1787/5kgk9qz j9nnx-en>
- Noe, R.A. (2017). *Employee training and development* (7th ed.). McGraw-Hill Education.
- Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. *Psychological Science in the Public Interest*, 13(2), 74 - 101. <https://doi.org/10.1177/1529100 612436661>
- Schroeder, R.G., & Goldstein, S.M. (2018). *Operations management in the supply chain: Decisions and cases* (7th ed.). McGraw-Hill Education.
- Senge, P.M. (1990). *The fifth discipline: The art and practice of the learning organization*. Doubleday.
- Wickramasinghe, V., & Perera, L. (2014). Soft skills for the next generation workforce: A study among Sri Lankan undergraduates. *Education + Training*, 56(3), 275 - 294. <https:// doi.org/10. /ET- 07-2012-0079>



The Impact of Digital Skill Proficiency on Youth Employment Opportunities in Sri Lanka

Perera ISRG¹, Pallegedara A²

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

ABSTRACT

Digital literacy, defined as the ability to effectively use digital tools, access online platforms, and participate in the digital economy, has emerged as a key determinant of youth employability. In Sri Lanka, despite improvements in digital infrastructure, youth unemployment remains high, partly due to gaps in digital competencies. This study addresses the limited empirical evidence on the relationship between digital skill proficiency and youth employment by analyzing Labour Force Survey data from 2018 and 2023. Using Probit regression analysis, the findings reveal a positive and statistically significant effect of digital skills on employment outcomes, with the impact strengthening over time. The results indicate that digitally proficient youth are more likely to secure employment, adapt to changing workplace demands, and remain competitive in an increasingly digital labor market. These findings provide actionable insights for policymakers, educators, and employers to prioritize digital skill development as a strategy to enhance youth employability and address persistent labor market challenges.

KEYWORDS: Digital Literacy, Digital Skill Proficiency, Employability, Youth Unemployment

1. INTRODUCTION

In today's world, technology has emerged as a central element shaping economic growth and employment trends. Digital literacy defined as being able to utilize digital tools to their full potential, access online platforms, and participate in the digital economy is increasingly being considered a component of employability (Luthfia et al., 2021). The demand for a digitally literate workforce continues to grow with economies everywhere shifting to digital.

In Sri Lanka, youth unemployment remains a major challenge. Despite improvements in digital infrastructure and expanding access to technology, the youth unemployment rate was 23% in 2023, with new entrants to the labor market constituting the majority of those unemployed (Sri Lanka Labour Force Survey, n.d.). Analysis of the 2018 Labour Force Survey data revealed that

digitally literate youth were 3.61% more likely to be unemployed, highlighting a mismatch between existing skills and labor market requirements (Weerasiri & Samaraweera, 2021). This deficit in digital competencies restricts employment opportunities for youth and undermines the country's ability to remain competitive in a global digital economy (Weerasinghe & Samaraweera, 2021; Shita et al., 2018).

This study investigates the effect of digital skill proficiency on youth employment in Sri Lanka, addressing a gap in the empirical literature. By analyzing Labour Force Survey data from 2018 and 2023, the research evaluates how digital skills influence employment outcomes and provides evidence to guide policies, educational programs, and workforce development strategies aimed at enhancing youth employability.

2. LITERATURE REVIEW

Digital skill competency has become a key requirement for participation in the labour market during the digital age. Previous researches show that those who possess better ICT skills are more resilient to technology shifts and thus have better employability results (Peng, 2017; Matli & Ngoepe, 2019). It also shows that the digital skills help reduce the risk of job loss due to automation by enabling workers to transition into new jobs created (Stojanović et al., 2023). Thus, digital literacy is a crucial driver of career progression and labour market integration, particularly among the youth.

Global evidence indicates that youth are disproportionately exposed to unemployment due to skill imbalances and insufficient labour experience (Baah-Boateng, 2013). Even in Sri Lanka, youth labour market expectations and abilities remain an underlying cause of unemployment (Weerasiri & Samaraweera, 2021). Some studies find age has mixed effects while successive generations face greater risks of unemployment, employability increases with the accumulation of skills and experience (Qayyum & Siddiqui, 2007; Tegegne, 2019).

The limited access to digital sources and ICT education bars young people living in rural and estate areas from accessing similar economic opportunities (UNESCO, 2018; Luthfia et al., 2021). With the ongoing development of digital technologies, digitally unskilled people are increasingly at risk of being excluded from formal employment and digital labor markets. Reducing digital inequality has therefore become a high policy priority in

increasing young people's labour force participation.

3. METHODOLOGY

This study adopts a quantitative research design to empirically examine the impact of digital skill proficiency on youth employment outcomes in Sri Lanka. The analysis is based on secondary data obtained from the Labour Force Surveys (LFS) of 2018 and 2023, administered by the Department of Census and Statistics (DCS). The unit of analysis is the individual youth respondent, focusing on their digital skill proficiency and employment status. The study uses a cross-sectional comparative time horizon, evaluating two distinct years (2018 and 2023) to assess how the impact of digital skills on youth employability has evolved over time.

The LFS applies a stratified random sampling framework to ensure adequate representation across all districts and residential sectors urban, rural, and estate. From the total national household sampling frame of 25,750 units, respondents aged between 18 and 35 years were extracted to form the study sample, resulting in 18,752 responses for 2018 and 17,000 for 2023.

Digital skill proficiency serves as the primary independent variable and is constructed as a composite index representing multiple competency dimensions, including digital awareness, device operational skills, and functional task performance. This operationalization is guided by established frameworks from UNESCO (2018), Hargittai (2005), and van Deursen and van Dijk (2014). Youth employment status, dichotomously coded (1 = employed, 0 = unemployed), is considered the dependent variable.

H₀: There is no significant relationship between Digital skill proficiency and youth Employment status.

Data processing involved cleaning, coding adjustments, and consistency checks conducted using STATA 11 statistical software. Initially, descriptive analysis was performed to summarize sample characteristics and variable distributions. Thereafter, Probit regression modeling was employed to estimate the marginal effects of digital skill proficiency on employment likelihood while controlling for a set of demographic and socioeconomic characteristics, including age, gender, location, education level, and English literacy. The empirical model is specified as follows:

$$P(Y_i=1|X_i) = \Phi(\beta_0 + \beta_1 X_i + \gamma TC_i)$$

Y_i - Youth employment status (1 = employed, 0 = not employed)

X_i - Digital Skill Proficiency score of i

$\Phi(\cdot)$ - CDF of standard normal

C_i - Vector of control variables (age, gender, etc.)

β_1 - Marginal effect of digital skill proficiency

4. RESULTS AND DISCUSSION

The descriptive statistics from both the 2023 and 2018 Labour Force Survey datasets provide initial evidence supporting the hypothesis that digital skill proficiency is positively associated with youth employment opportunities in Sri Lanka. Across both years, youth with higher levels of education particularly O/L, A/L and degree holders exhibit higher digital skill proficiency and are also more likely to be employed,

indicating a complementary relationship between digital and formal skills. Employment outcomes improve consistently with age, suggesting that as youth accumulate skills and work readiness, their employability increases.

Table 1: Probit Regression Results

Year	Variable	OLS Coeff (p-value)	Probit Marginal Eff. (p- value)
2023	Digital skill proficiency	0.1532 *** (0.000)	0.1799 *** (0.000)
	Constant	-0.6521 *** (0.000)	—
2018	Digital skill proficiency	0.0704 ** (0.000)	0.0900 ** (0.000)
	Constant	-0.5644 ** (0.000)	—

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The regression results provide strong empirical evidence in favor of Hypothesis that digital skill proficiency greatly enhances the employment opportunities for young people in Sri Lanka. In 2023, digital skill proficiency exhibits a marginal effect of 0.1799 ($p < 0.01$), indicating that a one-unit increase in proficiency raises the probability of employment by approximately 18 percentage points. In comparison, the 2018 results show a smaller effect, with a one-unit increase in digital skill proficiency associated with a 9-percentage point higher probability of employment. Both the 2018 as well as the 2023 evidence reveal a positive and statistically significant correlation between digital competences and jobs, with the effect also becoming

significantly more powerful in the recent period. This increasing trend proves that with the more digitalized workplace comes a greater demand for technology-literate young people, making digital literacy an essential employability skill. In addition, the findings indicate that more capable young people are more apt to be employed, adapt to new workplaces, and compete in an emerging labor market. Therefore, encouraging digital skills development is required to strengthen the labor market opportunities of youth and reduce skills imbalances in Sri Lanka's growing economy.

5. CONCLUSION

The analysis confirms that digital skill proficiency is a key driver of youth employment in Sri Lanka. Between 2018 and 2023, the positive impact of digital literacy on employment increased, reflecting the rising demand for technology workers in the digital economy. Youth with higher digital competencies are better equipped to access job opportunities, adapt to workplace changes, and reduce the risk of unemployment. Bridging digital skill gaps through targeted education, training programs, and inclusive ICT access, especially in rural and underserved areas can enhance employability and economic participation. Strengthening digital literacy among young people is therefore essential for reducing unemployment, fostering equitable economic growth, and ensuring Sri Lanka remains competitive in a rapidly digitalizing global labor market.

This study is constrained by the use of secondary Labour Force Survey data, which limits the analysis to available

variables. Important influences on digital skill proficiency, such as training quality, personal motivation, and workplace digital requirements, were therefore not captured. Additionally, the lack of employer-level information and detailed job categories restricts deeper insights into employment contexts. Future research should use primary data and richer employment classifications to better examine how digital skills shape employment outcomes.

REFERENCES

- Baah-Boateng, W. (2013). Determinants of unemployment in Ghana. *African Development Review*.
- Department of Census and Statistics & Ministry of Finance, Economic Stabilization and National Policies. (2023). *Sri Lanka labour force survey annual report*
- Hargittai, E. (2005). Survey measures of web-oriented digital literacy. *Social Science Computer Review*, 23(3), 371 - 379.
- Luthfia, A., Wibowo, D., Widyakusumastuti, M.A., & Angeline, M. (2021). The role of digital literacy on online opportunity and online risk in Indonesian youth. *DOAJ: Directory of Open Access Journals*.
- Matli, W., & Ngoepe, M. (2019). Capitalizing on digital literacy skills for capacity development of people who are not in education, employment or training in South Africa. *African Journal of Science Technology Innovation and Development*, 12(2), 129 - 139.
- Peng, G. (2017). Do computer skills affect worker employment? An empirical study from CPS surveys. *Computers in Human Behavior*, 74, 26 - 34.
- Qayyum, W., & Siddiqui, R. (2007). Causes of youth unemployment in Pakistan [with comments]. *The Pakistan Development Review*, 46(4), 611 - 621.
- Shita, A., Dereje, M., & Debre Markos University. (2018). Determinants of urban youth unemployment: Evidence

- from East Gojjam Zone, Ethiopia. *International Journal of Economic Development*, 245 - 265.
- Stojanović, A., Milošević, I., & Arsić, S. (2023). The influence of digital literacy on the employment rate of the older population. In *XIX International May Conference on Strategic Management*. University of Belgrade, Technical Faculty in Bor, Department of Engineering Management.
 - Tegegne, T.K. (2019). Socioeconomic determinants of youth unemployment in Ethiopia: The case of Wolaita Sodo Town, Southern Ethiopia. *Developing Country Studies*, 9(10), 1 - 9.
 - UNESCO. (2018). *Digital literacy global framework*. United Nations Educational, Scientific and Cultural Organization.
 - van Deursen, A.J.A.M., & van Dijk, J.A. G. M. (2014). The digital divide shifts to differences in usage. *New Media & Society*, 16(3), 507 - 526.
 - Weerasiri, A., & Samaraweera, G. (2021). Factors influencing youth unemployment in Sri Lanka. *Asian Journal of Management Studies*, 1(1), 49 - 72.
 - Weerasinghe, T., & Samarasinghe, G. (2021). Digital inclusion and employment among youth in urban Sri Lanka. *Sri Lanka Journal of Social Sciences*, 44(2), 45 - 59.



Minimizing Overstocking through Improved Supplier Delivery Management: A Case Study

Vijayamini KPSM¹, Jothirathne PAAU²

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka*
sachithyavijayamini@gmail.com.lk

ABSTRACT

Overstocking is a major challenge in inventory management, particularly within the apparel sector where product lifecycles are short and lead times are critical. This study investigates how supplier delivery performances contribute to raw material overstocking at ABC Pvt Ltd. The research utilizes real operational data from 1616 delivery transactions across 15 suppliers. Data have analyzed using descriptive statistics, chi-square tests, one-way ANOVA and Turkey, post-hoc tests the results revealed that 97.1% of deliveries have delayed, confirming that supplier performance is a key factor influencing overstocking. ANOVA results revealed significant differences in mean delivery variation, while Turkey post hoc analysis identifies that three suppliers are affecting for early deliveries that may cause production disruptions. The study concludes that overstocking primarily arises from supplier behavior rather than the order size, emphasizing the need for vendor specific delivery control policies and improves scheduling communication.

KEYWORDS: Apparel Industry, Inventory Management, Overstocking, Supplier Delivery

1. INTRODUCTION

Efficient inventory management is essential in the apparel industry due to short product lifecycles and tight production schedules. At the selected company, irregular supplier deliveries (extremely prior or consistently late) have disrupted material flow, increased storage costs, and created production delays, contributing directly to overstocking (Mweshi, 2022). Earlier studies highlighted internal causes of excess inventory such as weak forecasting (Pertiwi & Imaroh, 2024) and imbalanced production processes (Kong, Ning & Kong, 2024). However, despite the importance of delivery reliability in supply-chain coordination theories (Chopra & Meindl, 2016) supplier delivery performance endures underexamined as a direct contributor to overstocking, particularly in Sri Lanka's apparel sector where empirical evidence is limited (Senanayake, 2010). This gap

restricts the industry's ability to develop targeted strategies for reducing inventory buildup. Therefore, this study aims to fill this research gap by: Identifying suppliers whose delivery practices contribute to overstocking and proposing strategies to minimize overstocking through improved delivery management.

2. LITERATURE REVIEW

Overstocking is negatively affecting profitability and process stability especially in the manufacturing sector. In the apparel industry, short product life cycles and tight production schedules make inventory control particularly challenging. Prior research shows that overstocking attributes to weak forecasting, unbalanced production processes, and unreliable supplier performance. Mweshi (2022) highlighted that excess inventory increases storage, capital costs while hiding underlying process inefficiencies. Similarly, Kong et al. (2024) have found that overstocking

often arises from imbalanced production stages, which create bottlenecks and congestion in apparel manufacturing systems. They proposed lean production and pull-based scheduling methods to reduce such waste. Forecasting accuracy has also been identified as a crucial factor. Chopra and Meindl (2016) have emphasized that forecasting methods must align with product characteristics and demand volatility to prevent excess stock. While much of the literature focuses on internal inefficiencies, external supply-side factors particularly supplier delivery performance remain under-researched. Irregular supplier deliveries, whether significantly early or late, directly disturb planned inventory levels. Early deliveries increase holding costs, while late deliveries cause production delays and emergency procurement. According to supply chain coordination and inventory control theories, delivery variability amplifies uncertainty within production systems and forces firms to uphold higher safety stocks. Lean production theory further suggests that improved supplier coordination can reduce waste and ensure smoother material flow. Despite these theoretical insights, empirical evidence quantifying supplier delivery variation as a direct driver of overstocking has been limited, especially in Sri Lanka's apparel sector. Addressing this gap, the present study investigates how supplier delivery performance influences inventory levels at the chosen company and proposes strategies to minimize overstocking.

3. METHODOLOGY

The study adopted a deductive research approach. Theoretical concepts of supply chain performance and

overstocking have been collected using literature. Empirical data have been collected to test these relationships. This approach allowed the study to confirm whether supplier delivery performance significantly contributes to overstocking in a real-world industrial context. The research employed a quantitative method, focusing on numerical data analysis to recognize delivery patterns and the identification of statistical relationships among variables. This method has been chosen for its suitability in producing goal-oriented, measurable, and replicable findings related to supplier behavior and inventory performance.

The study has used secondary data obtained from the chosen company. Data have been collected from the company's internal delivery tracking system, consisting of 1,615 supplier delivery records from 15 major vendors. Each record included information about the vendor's name, goods requested date, goods received date, and quantity delivered. From these details, a new computed variable "Delivery Variation (Days)" has been derived by subtracting the goods requested date from the goods received date. Before analysis, the dataset has been examined for missing values and inconsistencies. Several records displayed unrealistic delivery variations, such as extremely early or delayed dates, most probably due to data entry or system errors. To improve data accuracy, an outlier removal process has been applied, while controlling only deliveries occurring within 365 days of the scheduled delivery date. This cleaning reduced the dataset to 1,570 valid records, ensuring the reliability and realism of the analysis. The population for this study consisted of all supplier delivery transactions recorded at

the chosen company during the selected period. This included every completed delivery from the company's 15 key suppliers. The study analyzed the entire population. This approach ensured comprehensive coverage of all supplier performance data and withdrawn sampling bias. After data cleaning, a total of 1,570 valid delivery records have included in the analysis. Data have analyzed using IBM SPSS Statistics, Statistical techniques included frequency analysis, Chi-square tests, one-way ANOVA, Tukey post-hoc comparisons, and Pearson correlation analysis. These tools have applied to recognize delivery performance trends, search/recognize whether differences exist among suppliers, and evaluate how delivery behavior contributes to overstocking within the organization.

4. RESULTS & DISCUSSION

4.1. Frequency Analysis

A frequency analysis has carried out to examine the overall distribution of delivery performance among suppliers.

4.2 The Delivery Classification variable

The results showed that 99.9% (n = 1,569) of deliveries have classified as Late. Only 0.1% (n = 1) has recorded as Very Early. There were no deliveries that matched the exact planned date. This extreme imbalance manifests hat

suppliers consistently abortive to meet scheduled delivery dates, with approximately overall. Shipments arriving after the expected duration.



Figure 1: Supplier Distribution

Table 2: Delivery Classification Distribution

		Statistics			
		Delivery Variation (Days)	Delivery Classification	Supplier	Qty delivered (m)
N	Valid	1570	1570	1570	1570
	Missing	0	0	0	0

Table 2: Delivery Variation (Days)

Supplier				
Valid	Frequency (Days)	Percent	Valid Percent	Cumulative Percent
A	32	2.0	2.0	2.0
B	6	.4	.4	2.4
C	77	4.9	4.9	7.3
D	124	7.9	7.9	15.2
E	428	27.3	27.3	42.5
F	5	.3	.3	42.8
G	726	46.2	46.2	89.0
H	12	.8	.8	89.8
I	7	.4	.4	90.3
J	31	2.0	2.0	92.2
K	15	1.0	1.0	93.2
L	32	2.0	2.0	95.2
M	48	3.1	3.1	98.3
N	8	.5	.5	98.8
O	19	1.2	1.2	100.0
Total	1570	100.0	100.0	

a. Cross Tabulation

A Chi-square test of independence examined whether delivery performance differed significantly among suppliers.

Table 4: Cross Tabulation

Test	Value	df	Sig.
Pearson Co Relation	11.669	14	0.633

Because the significance value ($p = 0.633 > 0.05$) is not statistically significant, delivery delays appear uniform across most suppliers. Although statistical dependence has not claimed specifically, the predominance of late deliveries across all vendors suggests a system-wide scheduling inefficiency rather than isolated supplier issues.

Therefore, strengthened communication and delivery monitoring should mandate across the supplier network.

Table 5: Chi square Test

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	11.669 ^a	14	.633
Likelihood Ratio	5.085	14	.985
N of Valid Cases	1570		

b. Theoretical Implications

The findings of this study make several contributions to supply chain and inventory management theory. First, the empirical evidence reinforces supply chain coordination theory, which posits that delivery variability increases uncertainty and forces firms to hold higher levels of safety stock (Chopra & Meindl, 2016). The results showing that 99.9% of deliveries had explicated lately how persistent deviations between planned and actual lead times create systemic instability within a production environment. Second, the study adds empirical support to inventory control and lean production theories, which argue that unreliable supplier performance disrupts material flow and contributes to excess inventory (Hall, 2016). Prior literature has often highlighted internal causes of overstocking, such as forecasting errors or production imbalances. However, empirical studies quantifying supplier delivery performance as a direct contributor to overstocking had limited, particularly in the context of Sri Lanka's apparel industry. This study therefore fills an important research gap by showing that supply-side lead-time variability is a significant operational risk factor, even when order quantities are not the primary cause. Finally, the findings challenge the common assumption that overstocking has driven primarily, by internal inefficiencies. Instead, this study insinuates that system-wide lateness across suppliers can create unavoidable inventory accumulation, suggesting the need to expand theoretical models to incorporate external delivery reliability as a core constituent of inventory behavior.

c. Practical Implications

From a managerial perspective, the study illustrates actionable insights for improving supply-chain performance at ABC Pvt Ltd. The dominance of late deliveries indicates a need for stronger coordination mechanisms between the company and its suppliers. Implementing a digital delivery-tracking system would give procurement real-time visibility into expected vs. actual delivery dates, reducing scheduling uncertainty. The results also suggest that lead-time agreements must be reassessed, as current contractual timelines do not reflect actual supplier capabilities. By introducing realistic delivery windows, along with penalties for chronic delays, the company can create incentives that improve supplier accountability.

Internally, chronic lateness means procurement teams currently rely on buffer stock and emergency ordering; both of which increase costs and contribute to overstocking. Streamlining approval processes and enhancing forecasting communication with suppliers can reduce these inefficiencies. Finally, establishment of a supplier performance scorecard will allow ABC Pvt Ltd to observe punctuality, while exclusion of high-risk suppliers early, and take proactive corrective actions.

5. RECOMMENDATIONS

The findings of this study clearly confirm that supplier delivery performance at ABC Pvt Ltd has characterized by chronic lateness across considerably overall vendors, pointing to a system-wide coordination issue rather than isolated supplier inefficiency. Therefore, improvements should focus on strengthening scheduling, monitoring, and communication processes throughout

the supply chain. The company has to implement digital delivery-tracking and monitoring system that links purchase orders to both expected and actual delivery dates, providing automatic alerts for overdue consignments. Lead-time agreements have introduced and adjusted to reflect realistic supplier capabilities, with contractual delivery windows and penalty clauses introduced to ensure greater accountability. Enhanced coordination between the procurement team and key suppliers through regular review meetings and shared production forecasts is also essential to improve schedule adherence. Internally, the company should streamline approval and purchase workflows to reduce administrative delays and improve responsiveness. In addition, a supplier performance scorecard has to develop to measure punctuality, responsiveness, and reliability, reviewed periodically for the determination of consistent delays. For future research, the study recommends extending the analysis to a longer duration and incorporating variables such as supplier distance, shipment mode, and material category to better understand the factors influencing delivery performance and overstocking. By addressing these areas, ABC Pvt Ltd can strengthen supplier reliability, reduce unnecessary stock buildup, and enhance the overall efficiency of its inventory management system.

6. LIMITATIONS AND FUTURE WORKS

The dataset used in this study has derived from real operational records, which contained certain inconsistencies and extreme values. While these irregularities may affect statistical

precision, they also represent the actual data quality and process reliability within the company's supply chain system. Another limitation of the study is the short data collection period, which covered only a limited time frame.

To enhance the findings of this study and strengthen future analyses, the following directions have recommended: Expand the time limit of data collection, conduct comparative studies, include supplementary variables, and evaluate policy interventions.

REFERENCES

- Chopra, S., & Meindl, P. (2016). *Supply Chain Management: Strategy, Planning, and Operation*. Pearson.
- Hall, N.G., & Porteus, E.L. (2000). Customer and supplier lead-time variability in supply-chain inventory management. *Manufacturing & Service Operations Management*, 2(2), 159 - 172.
- Kong, L., Ning, Y., & Kong, X. (2024). Lean flow optimization and overstock reduction in apparel manufacturing. *International Journal of Production Research*, 62(4), 1120 - 1135.
- Mweshi, G.K. (2022). Effects of overstocking and stockouts on the manufacturing sector. *International Journal of Advances in Engineering and Management*, 4(9), 1054 - 1064.
- Pertiwi, L.M., & Imaroh, A. (2024). Adaptive forecasting methods for minimizing inventory imbalance in manufacturing industries. *Journal of Industrial Engineering and Management*, 15(1), 78 - 89.
- Senanayake, M.M., & Little, T. (2010). A review of apparel supply chain issues in South Asia. *International Journal of Advanced Manufacturing Technology*, 48(9 - 12), 1343 - 1355.



Investigating Factors Influencing Behavioral Intention toward the Adoption of a New ERP System in the Financial Management Sector of Sri Lanka's Automobile Industry; Systematic Literature Review

Ramesha V¹, Oruthotaarachchi CR²

^{1,2} *Department of Industrial Management, Faculty of Science, University of Kelaniya*
ramesha-im20097@stu.kln.ac.lk¹; chalani@kln.ac.lk²

ABSTRACT

This systematic literature review investigates the behavioral factors influencing the implementation of Enterprise Resource Planning (ERP) systems within the financial management sector of organizations in Sri Lanka's automobile industry. Using the PRISMA framework, the review synthesizes findings from 40 peer-reviewed studies published between 1989 and 2025, selected from databases such as Google Scholar, ScienceDirect, Emerald Insight, IEEE Xplore, ResearchGate etc. The review focuses on theoretical models including the Technology Acceptance Model (TAM), Theory of Planned Behavior (TPB), Unified Theory of Acceptance and Use of Technology (UTAUT). Key behavioral constructs such as perceived usefulness, perceived ease of use, social influence, resistance to change, and organizational support were found to significantly influence user intention toward ERP adoption. However, research specific to ERP implementation in the financial departments of Sri Lankan automobile organizations remains limited. This study addresses this gap and contributes valuable insights for future ERP implementation strategies in similar developing country contexts.

KEYWORDS: Automobile Industry, Behavioral Intention, ERP Adoption

1. INTRODUCTION

ERP systems enhance organizational efficiency by integrating key functions like finance, supply chain, and HR, offering benefits such as real-time data and improved decision-making. However, successful implementation depends on end-users' behavioral intention, influenced by factors like perceived usefulness, ease of use, organizational culture, and leadership support. In Sri Lanka, the financial management sector of the automobile industry remains significantly underexplored in existing research. This review addresses that gap by analyzing literature on behavioral intention toward ERP adoption, using models such as TAM, TPB, UTAUT etc.; with a focus on insights relevant to financial departments in Sri Lanka's automobile industry. This

review aims to systematically synthesize recent academic literature on behavioral intention toward ERP system adoption and to identify key research gaps, with a particular focus on the underexplored context of the Sri Lankan automobile industry.

2. METHODOLOGY

This review follows the PRISMA framework to ensure a structured process. The review was designed to examine behavioural factors influencing ERP adoption, with a special focus on Sri Lanka's automobile industry and its financial management sector. A systematic literature search was conducted across leading academic databases covering studies published between 2000 and 2025. A few key theoretical papers before 2000 were included because they introduced

foundational acceptance models.
Keywords used included:

"ERP adoption" OR "ERP implementation" OR "Enterprise Resource Planning"

AND "behavioral intention" OR "TAM" OR "UTAUT" OR "TPB"

AND "Sri Lanka" OR "automobile industry" OR "financial management"

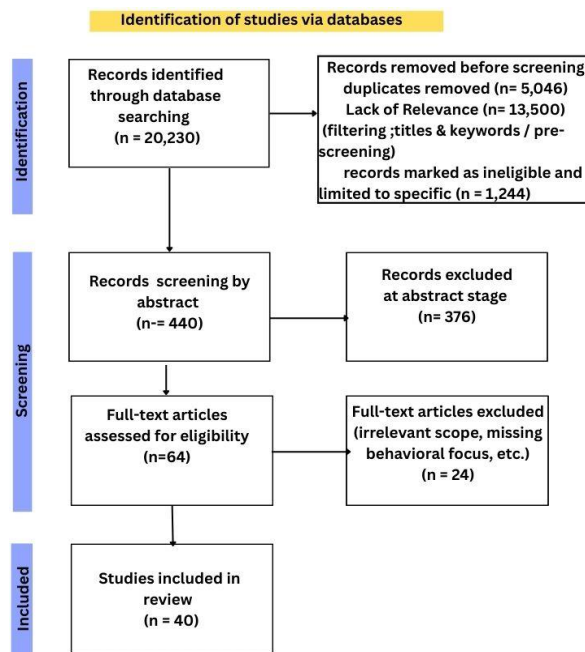


Figure 1: PRISMA Framework

Key details (Author and Year, Title, Objectives, Scope, Country/Region, Industry/Domain, Methodology, Key Findings, Behavioral Factors Identified, Relevance to Research, Research Gaps) were extracted into an Excel matrix for synthesis.

3. RESULTS

3.1 Thematic Synthesis

This section synthesizes the behavioral factors identified across the selected literature, structured around four key themes.

1. Technology Acceptance and Unified Adoption Models (TAM, TPB, UTAUT, and Related Extensions)

The Technology Acceptance Model (TAM) (Davis, 1989) and its extensions, including TAM2 (Venkatesh & Davis, 2000) and the Task-Technology Fit (TTF) model, have been widely applied to explain ERP system adoption through key determinants such as perceived usefulness (PU), perceived ease of use (PEOU), and task-technology fit. Likewise, the Theory of Planned Behavior (TPB) (Ajzen, 1991) and the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003) emphasize attitudes, performance expectancy, effort expectancy, social influence, and facilitating conditions as major predictors of behavioral intention. Empirical studies (e.g., Wu et al., 2013; Williams, 2015; Wang & Teo, 2022; Rawashdeh et al., 2020) consistently validate these constructs across diverse contexts. However, there remains limited research applying these theoretical models to ERP system adoption in Sri Lanka's automobile industry, particularly within the financial management sector, highlighting a significant research gap.

2. Factors Influencing Behavioral Intention

Theme 1: Perceived Usefulness and Perceived Ease of Use (TAM Constructs)

Perceived usefulness (PU) and perceived ease of use (PEOU) are widely recognized as key predictors of ERP adoption (Davis, 1989; Venkatesh & Davis, 2000). Studies show that users are more likely to accept ERP systems when they find them helpful in improving job performance and easy to navigate. Wu et

al. (2013) and Samarasinghe & Gunawardena (2019) confirmed that systems enhancing productivity and offering user-friendly interfaces strengthen behavioral intention, especially in financial departments where efficiency and accuracy are vital.

Theme 2: Organizational Factors

Organizational support, including leadership commitment, effective training, and clear communication, significantly influences ERP adoption. Research (Al-Mashari et al., 2003; Perera & Rajapakse, 2015) highlights that management involvement and tailored training enhance employee readiness and confidence. In Sri Lanka, open communication and phased rollouts have been shown to reduce resistance and promote acceptance (Samarasinghe & Rajapakse, 2020; Eampoonga & Leelasantitham, 2023).

Theme 3: System-Related Factors

System quality factors such as interface design, reliability, and compatibility with existing processes strongly impact user satisfaction and trust. Studies (Wu et al., 2013; Schniederjans & Yadav, 2013; Iqbal & Su, 2025) emphasize that reliable, intuitive, and well-integrated ERP systems enhance confidence, while poor design or performance discourages usage especially in finance-focused roles.

Theme 4: Attitudes, Fear of Change, and Motivation (UTAUT2 and TPB Constructs)

Personal attitudes, fear of change, and motivation also shape ERP acceptance. Employees resistant to change or anxious about automation often

show lower adoption levels (Kazmi, 2016; Samarasinghe & Gunawardena, 2019). Conversely, motivated and confident users with higher self-efficacy display stronger behavioral intention (Lekhuleni, 2019; Salih et al., 2022). Social influence and peer encouragement further affect attitudes, particularly in team-based cultures (Venkatesh et al., 2003; Williams, 2015).

3.2 Descriptive Analysis

The review of 40 studies on ERP adoption shows growing research interest since 2014, peaking in 2020, with most studies from Asia, though only a few focus on Sri Lanka. Research spans multiple industries, but the automobile sector especially financial management is largely unexamined. Common frameworks include TAM, TPB, UTAUT, TOE, and TTF, emphasizing user perceptions, organizational readiness, and system fit. Methodologically, studies are mostly quantitative, with limited qualitative insight, highlighting a clear gap in context-specific research on behavioral intention toward ERP adoption in Sri Lanka's automobile finance sector.

4. DISCUSSION

4.1 Research Gap

Despite extensive research using TAM, TPB, and UTAUT to study ERP adoption, gaps remain in Sri Lanka's automobile industry, particularly within financial management. Most studies focus on general ERP implementation or other industries, overlooking behavioral intention, user acceptance, and the impact of cultural and organizational factors in hierarchical business settings. Key aspects such as attitudes, motivation,

resistance to change, task-technology fit, and organizational readiness are underexplored, and few studies use qualitative or mixed-method approaches. This study addresses these gaps by examining behavioral intention among financial staff, integrating cultural, organizational, and system-related factors to better understand ERP adoption in this context.

5. CONCLUSION

This systematic literature review examined existing research on behavioral intention toward ERP system adoption, focusing on key influencing factors and theoretical frameworks such as TAM, TPB, and UTAUT. The review found that while global studies emphasize factors like perceived usefulness, ease of use, social influence, and facilitating conditions, there is limited research specific to Sri Lanka's automobile industry particularly its financial sector. Most studies use quantitative approaches, overlooking qualitative insights into employee attitudes and organizational culture. Therefore, future research should adopt context-specific and integrated models to better understand ERP adoption behavior and improve implementation strategies in developing country contexts.

REFERENCES

- Abuhassna, H., et al. (2023). Trends on using the Technology Acceptance Model (TAM) for online learning: A bibliometric and content analysis. *International Journal of Interactive Mobile Technologies (IJIM)*, 17(8), 1–20.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2).
- Al-Ghamdi, F. (2014). Change management strategies and processes for the successful ERP system implementation: A proposed model.
- Al-Mashari, M., Al-Mudimigh, A., & Zairi, M. (2003). Enterprise resource planning: A taxonomy of critical factors. *European Journal of Operational Research*.
- Amid, A., Moalagh, M., & Zare Ravasan, A. (2012). Identification and classification of ERP critical failure factors in Iranian industries. *Computers in Industry*.
- Buonanno, G., Faverio, P., Pigni, F., Ravarini, A., Sciuto, D., & Tagliavini, M. (2005). Factors affecting ERP system adoption: A comparative analysis between SMEs and large companies. *Journal of Enterprise Information Management*.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of IT. *MIS Quarterly*, 13(3), 319–340.
- Eampoonga, I., & Leelasantham, A. (2023). Overall success factors affecting the performances of hybrid cloud ERP: A case study of automobile industries in Thailand.
- Fadelelmoula, A.A. (2018). The effects of the critical success factors for ERP implementation on the comprehensive achievement of the crucial roles of information systems in the higher education sector.
- Farhadi, M., et al. (2020). Identifying and prioritizing the key factors of ERP system implementation in Saderat Bank.
- Fasileem, M., & Rajapakse, B. (2022). Underutilization of XYZ Cloud ERP system in apparel industry. *SCIRP*.
- Haddara, M., & Elragal, A. (2018). Evolution of ERP systems in the cloud: A study on system updates.
- Hawking, P., Stein, A., & Foster, S. (2004). Second wave ERP education. *Journal of Information Systems Education*.
- Iqbal, M.A., & Su, J. (2025). Technology adoption in the apparel industry of Bangladesh: A qualitative study. *Research Journal of Textile and Apparel*.
- Kaluarachchi, T.R., Weerabahu, W.M.S. K., & Nanayakkara, L.D.J.F. (2019). Factors influencing enterprise information systems adoption of SMEs: A case study on Sri Lankan manufacturing sector.
- Kazmi, S.M.H. (2016). Benefits and challenges of enterprise resource planning in Pakistani SMEs.
- Kashmeery, F. (2016). Investigating the

INVESTIGATING FACTORS INFLUENCING BEHAVIORAL INTENTION TOWARD THE
ADOPTION OF A NEW ERP SYSTEM IN THE FINANCIAL MANAGEMENT SECTOR OF SRI
LANKA'S AUTOMOBILE INDUSTRY; SYSTEMATIC LITERATURE REVIEW

- effects of power relations on the implementation of enterprise resource planning through the lens of actor network theory: The case of MESAIR. ProQuest.
- Kharuddin, S., et al. (2015). Effects of decision rationality on ERP adoption extensiveness and organizational performance. *Journal of Enterprise Information Management*.
 - Kotter, J.P. (1996). Leading change. Harvard Business School Press.
 - Ledlum, J.E. Jr. (2010). Enterprise resource planning (ERP) software package recommendation and assessing the readiness for change within a privately own company in Monrovia, Liberia, West Africa.
 - Lekhuleni, A. (2019). Exploring factors affecting the implementation of enterprise resources planning systems for South African small and medium scale constructions industry.
 - Lugonja, D. (2020). ERP implementation and interoperability growth as organization improvement agents in public administration further development. *Journal of Information and Organizational Sciences*.
 - Musaa, A., et al. (2024). Twenty years of TAM research in marketing: A bibliometric review.
 - Nijher, H.S. (2014). Exploring critical success factors of ERP implementation in United Nations types of organizations: Relationship between factors impacting user experience. Library and Archives Canada.
 - Perera, K., & Rajapakse, J. (2015). Challenges in ERP implementation in Sri Lankan organizations.
 - Rawashdeh, A.M., Elayan, M.B., Alhyasat, W., Dawood, M., & Shamout, A. (2020). Electronic human resources management perceived usefulness, perceived ease of use and continuance usage intention: The mediating role of user satisfaction in Jordanian hotels sector.
 - Saade, R.G., & Nijher, H. (2016). Critical success factors in enterprise resource planning implementation: A review of case studies. *Journal of Enterprise Information Management*, 29(3), 370–396.
 - Salih, S.H., Abdelsalam, S., Hamdan, M., Abdelmaboud, A., Hamzah, M., & Hilal, A. M. (2022). Critical success factors for ERP systems' post-implementations of SMEs in Saudi Arabia: A top management and vendors' perspective.
 - Samaranayake, P., & Samarasinghe, D. (2018). ERP adoption in Sri Lankan industries: A case from the automobile sector. *Sri Lanka Journal of Management*.
 - Samarasinghe, D., & Rajapakse, J. (2020). Cultural and behavioral influences on ERP acceptance in Sri Lankan organizations. *South Asian Journal of Business and Management Cases*, 9(2), 200–212.
 - Schniederjans, M., & Yadav, S. (2013). Successful ERP implementation: An integrative model. *Journal of Enterprise Information Management*, 26(2), 151–165.
 - Siddiqui, N., et al. (2023). A systematic review of ERP implementation strategies in the retail industry.
 - Venkatesh, V., & Davis, F.D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2).
 - Venkatesh, V., Morris, M.G., Davis, G.B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view (UTAUT).
 - Wang, Y., & Teo, T.S.H. (2022). Examining shopping intention and behavior in mobile payment adoption among middle-aged users using UTAUT. *Frontiers in Psychology*, 13, 830842.
 - Williams, M. D. (2015). The unified theory of acceptance and use of technology (UTAUT): A literature review.
 - Wu, M.Y., Wu, M.H., & Shih, Y.Y. (2013). A study of factors affecting behavioral intention to use ERP systems in high-tech industry using TAM and TTF.



Explainable AI for Voice and Broadband Customer Churn Prediction in the Sri Lankan Telecommunication Sector

Boralessa KT¹, Prabodanie RAR²

^{1,2} *Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

ABSTRACT

Customer churn has a major impact on revenue and long-term sustainability in the telecom sector. This research proposes an explainable artificial intelligence XAI framework for predicting customer churn in Sri Lanka, one of the leading telecommunications providers of voice and broadband services, by combining machine learning predictive ML algorithms with Shapley Additive Explanations (SHAP) for explaining the predictions. Data from 30,022 customer records over five months, were analyzed including customer demographic data, service usage data, billing and payment data, voice data, and fault report data. Three models, namely Random Forest, Logistic Regression, and XGBoost, were developed, and their performance was evaluated using three metrics: precision, recall, and F1 score. The XGBoost model outperformed the other two models. The hyperparameters of the XGBoost model were optimized using random search. The final model performed with a recall of 0.8, a precision of 0.67, and an F1 score of 0.73. Out of the features analyzed, the most important features for churn prediction were identified using SHAP analysis. They include unpaid balance, payment amount, total bill, service usage pattern, and number of faults. The results suggest that customers avoid paying their bills or using the services regularly due to dissatisfaction with payments and services. This causes them to leave. This research provides actionable insights for targeted retention strategies in the Sri Lankan telecommunications industry.

KEYWORDS: Churn Prediction, Explainable AI, Machine Learning, Shap, Telecommunication

1. INTRODUCTION

Customer churn poses a significant challenge to the telecom industry, directly impacting revenue and long-term business sustainability. With global churn rates ranging between 20% and 50% (Waweru, 2025), retaining profitable customers has become increasingly critical amid intense competition and service commoditization. Predicting churn behaviour enables telecom operators to implement targeted retention strategies, thereby reducing subscriber loss. Machine learning models, especially ensemble methods like XGBoost, have demonstrated superior accuracy in churn prediction due to their ability to handle complex, high-dimensional data and nonlinear relationships (Asif, 2025). However, many high-performing models operate as

black boxes, limiting actionable insights for decision-makers. Explainable Artificial Intelligence (XAI) techniques, such as SHAP (Shapley Additive Explanations), address this gap by providing interpretability on feature contributions to model predictions (ÖZKURT, 2024). This study focuses on churn prediction within the voice and broadband service segment of Sri Lankan telecommunication, leveraging machine learning with integrated explainability. By analyzing extensive customer data, including demographics, service usage, and billing, this research aims to enhance understanding of churn drivers and deliver interpretable, evidence-based retention recommendations tailored for telecom providers.

2. LITERATURE REVIEW

The customer churn prediction remains a focal point for operators seeking to reduce revenue loss and increase retention efficiency. The use of machine learning techniques has increasingly dominated churn prediction due to their superior ability to analyze complex, high-dimensional, and imbalanced telecom datasets (Fauzan, 2021; Alotaibi, 2024). Traditional models like Logistic Regression are valued for simplicity and interpretability, but often fall short in capturing nonlinear patterns in churn behaviour (Jain, 2020). Random Forests improve upon this by handling large feature sets and imbalances more effectively, often yielding better precision and recall metrics (Ahmed, 2017). Among ensemble methods, XGBoost has gained prominence due to its ability to regularize better, parallel processing capabilities, and robustness, frequently outperforming other models in churn prediction accuracy when paired with hyperparameter optimization strategies such as random search and genetic algorithms ((K & Y, 2022); (Saleh, 2024)) while predictive performance is critical, the interpretability of models has emerged as equally vital for operational decision-making. Explainable AI (XAI) methods like SHAP provide both local and global interpretability by quantifying feature contributions to individual churn predictions, enabling actionable insights (Lundberg & Lee, 2017). Comparative research underscores SHAP's superiority over techniques like LIME and Partial Dependence Plots in offering consistent, reliable explanations (Huang, 2020). Integration of SHAP with XGBoost models has revealed pivotal churn drivers such as contract types,

monthly charges, and technical support presence (ÖZKURT, 2024). In the Sri Lankan telecommunication context, key churn factors include pricing, service quality, and customer support factors that influence subscriber retention, especially within bundled services (Samarasinghe, 2016). This body of work collectively highlights the effectiveness of combining high-performing predictive models with explainability techniques to develop transparent, data-driven retention strategies in telecom.

3. METHODOLOGY

For this study, data relevant to customer churn were obtained from a leading telecommunications service provider in Sri Lanka. Extensive preprocessing of extracted data was required to prepare the data for machine learning models. Three types of machine learning models, Random Forest, Logistic Regression, and Gradient Boosting, were trained and tested. Then hyperparameters of the best-performing model were optimized using random search. SHAP (Shapley Additive explanations) analysis was performed on the final model to evaluate feature importance.

3.1 Data Collection

Data required for churn prediction were extracted from seven data sets maintained in the Hadoop-based big data platform. The datasets contain details of customers, broadband usage, extra data purchased, payments, faults reported, bill summaries and voice usage. Data about 30,022 voice and broadband subscribers over a period of five months from November 2024 to March 2025 were extracted.

Table 1: Data Set Description

Dataset	Description
Sample	Contains basic customer information.
BB	Contains data about Customer service usage.
ExtraGb	Contains information about extra data purchased.
Payment	Contains payment-related information.
Fault	Contains the churn status of customers.
Bill summary	Contains billing information.
Voice	Contain call count and duration statistics.

3.2 Data Preprocessing

The data preprocessing phase included several vital stages: addressing missing values by substituting nulls with zeroes for the usage-related features as pronounced absence of this activity meant non-usage; resolving duplicate records to preserve the credibility of the data; detecting and managing outliers using Interquartile Range (IQR) methodologies and keeping high-value customers' outliers in the billing data; engineering features using aggregating at the customer level and calculating means for usage, payments and times to resolve faults and encoding categorical variables to prepare for model use by applying one- hot encoding. The pre-processed final dataset consisted of 30 features.

Some of those features:

- 1) Bill_total: Total billed amount for the customer's account.
- 2) Gb_amount: Total extra gigabytes
- 3) I_on_call_count: Average of outgoing call count.
- 4) Days_to_clear_mean: Total days to clear for all faults.

5) Fault_count: Count of fault references

6) Long_call_cnt: Long call count.

7) Balance_out_amt: Average balance out money.

The dependent variable, churn status, exhibited a class imbalance, consisting of 93.4% not churning and 6.6% churning.

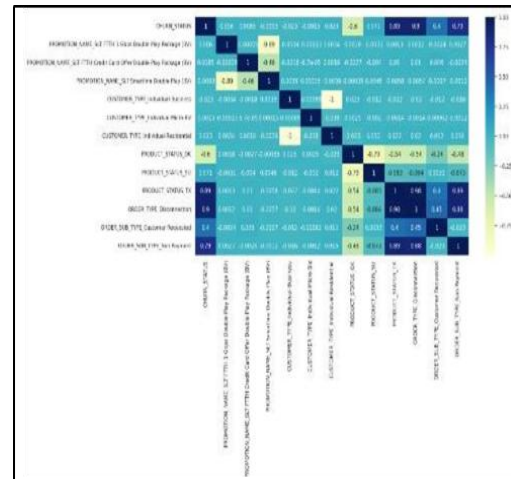


Figure 1: Heat Map of Data Preprocessing

3.3 Model Development

An 80:20 train-test split of the dataset was used to implement three machine learning algorithms: Random Forest, Logistic Regression, and XGBoost. Precision, recall, F1-score, and ROC-AUC metrics were used to assess the model's performance. Since misclassifying real churners as non-churners can result in larger business losses than false positives, recall was given priority in order to reduce false negatives due to the class imbalance in churn data.

The best performing model underwent hyperparameter tuning using the Randomized search CV method (Scikit-learn library in Python) with 50 iterations and 3-fold cross-validation. The Hyperparameters thus optimized were n_estimators, max_depth, learning rate, subsample, colsample_bytree, and scale_pos_weight to offset the class

imbalance.

3.4 Explainable AI

To interpret feature contributions, SHAP values were derived for the optimized model. SHAP is an approach that assigns SHAP values, based on Shapley values from cooperative game theory, in a consistent and locally accurate way.

Summary plots explained which features were important globally, whereas force plots explained the individual predictions for retention interventions.

4. RESULTS AND DISCUSSION

4.1 Model Performance Comparison

Table 2: Comparison of models

Metric	Random Forest	Logistic Regression	XGBoost
Precision	0.76	0.73	0.73
Recall	0.56	0.56	0.71
F1-Score	0.65	0.64	0.72
Accuracy	0.97	0.97	0.97

In the preliminary evaluation of the models' performance, XGBoost was the most effective model with 71% recall, 73% precision and 72% F1 score. In contrast, Random Forest (56% recall) and Logistic Regression (56% recall) both performed substantially worse than the XGBoost model. The XGBoost model classified 199 churners correctly; however, 73 customers were misidentified as churners, therefore producing false positives. The Random Forest model and the Logistic Regression model both classified only 158 churners with 123 false negatives.

4.2 Hyperparameter Optimization

Table 3: Comparison of Tuned XGBoost models

Metric	Random Search	Grid Search
Precision (1)	0.67	0.59
Recall (1)	0.8	0.85
F1-score (1)	0.73	0.7
Accuracy	0.97	0.96
ROC AUC	0.973	0.972

The performance of the XGBoost model demonstrated appreciable improvement for hyperparameter tuning by random search. The tuned model achieved a recall of 80%, a precision of 67%, and an F1-score of 73% along with an ROC-AUC of 0.973. The confusion matrix of the tuned model identified 225 true positives (correctly identified churners) with only 56 false negatives (missed churners); this was a significant 32% reduction in missed churners compared to the original model.

	precision	recall	f1-score	support
0	0.99	0.98	0.98	5557
1	0.67	0.80	0.73	281
accuracy			0.97	5838
macro avg	0.83	0.89	0.86	5838
weighted avg	0.97	0.97	0.97	5838
ROC AUC Score: 0.9729327314400036				
Confusion Matrix				
[[5446 111]				
[56 225]]				

Figure 2: Classification report for the tuned XGBoost model

4.3 Shap Based Feature Importance

The Shap summary plot is given in Figure 3. Each dot represents a customer. Colours ranging from blue to pink indicate feature values (pink is high, blue is low). The position of each feature on the vertical axis (top to bottom) indicates the

overall effect of the feature on churn prediction.

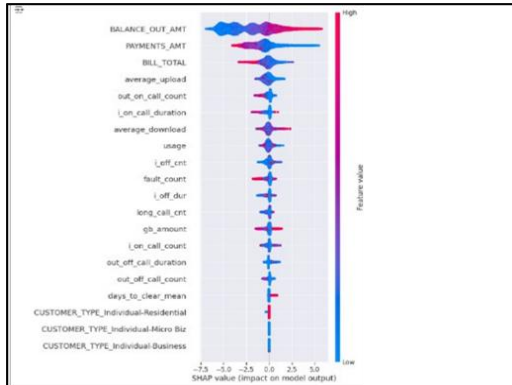


Figure 3: Shap Summary plot

SHAP analysis identified the following Churn identifiers.

- 1) Outstanding balance (balance_out_amt) - highest SHAP value indicating a strong positive correlation with churn.
- 2) Payment amount (payments_amt) – customers who had either erratic
- 3) payments or low payments were more likely to churn.
- 4) Total bill amount (bill_total) - higher bills corresponded to a higher likelihood of churn.
- 5) Service FPS metrics - average upload/download speeds or pattern of total usage significantly influenced predictions, fault count and resolution time - service reliability issues also affect churn probability.

The weakened effect of customer type aligns with results found by Samarasinghe (2016), which indicates that churn patterns are influenced by the use of voice and broadband services, which is evidence of meaningful churn patterns in regions whose customer segmentation is built on traditional service use. It is also certain that

leveraging SHAP with XGBoost assists with issues concerning model understandability previously noted by Hossain (2025). Overall, the integration of XGBoost and the SHAP model results establishes support for the cogency of AI-based retention strategies used in predicting customer churn.

5. CONCLUSION

This research utilized an explainable AI approach using an optimized XGBoost model with SHAP analysis to predict and understand churn among Sri Lanka Telecommunication broadband customers. The model demonstrated strong predictive accuracy, with key financial and service quality factors revealed as primary churn drivers. The study demonstrates how actionable retention strategies, like flexible payment plans and increased network reliability, are made possible by combining explainable AI with machine learning. However, the results are restricted to the broadband segment of a single provider and a particular time period, which may limit generalizability to other telecom operators, product lines, or different market conditions. When predicting results outside of this context, care should be taken because external factors like competitor behaviour and regulatory changes were not taken into consideration. In spite of these drawbacks, the framework provides a repeatable approach to telecom churn analytics that facilitates open, data-driven decision-making and customer retention initiatives.

REFERENCES

- Ahmed, A.L. (2017). A review and analysis of churn prediction methods for

- customer retention in the telecom industry. DOI:10.1109/ICACCS.2017.8014605
- Alotaibi, M.Z. (2024, June). Customer Churn Prediction for Telecommunication Companies using Machine Learning and Ensemble Methods. *Engineering, Technology and Applied Science Research*. doi:<https://doi.org/10.48084/etasr.7480>
 - Asif, D.A. (2025). A data-driven approach with explainable artificial intelligence for customer churn prediction in the telecommunications industry. *Results in Engineering*, 26. doi:<https://doi.org/10.1016/j.rineng.2025.104629>
 - Fauzan, M.A. (2021). Predicting Customer Churn using ensemble learning: Case Study of a Fixed. *International Journal of Technology*, 12(5). doi:<https://doi.org/10.14716/ijtech.V12i55223>
 - Hossain, M.R. (2025). Predicting Customer Churn in telecommunications with machine learning models. doi:<https://doi.org/10.9734/ajrcos/2025/v18i1548>
 - Huang, M.H. (2020). Engaged to a robot? The role of AI in service. *Journal of Service Research*, 155 - 164. <https://doi.org/10.1177/1094670520902266>
 - Jain, H.K. (2020). Churn Prediction in Telecommunication using Logistic Regression and Logit Boost. 167, 101-112. doi:<https://doi.org/10.1016/j.procs.2020.03.187>
 - K, P., & Y, P. (2022). Research on Telecom Customer Churn Prediction. *Journal of Computer and Communications*. doi:10.4236/jcc2022.1011008
 - Lundberg, S., & Lee, S. (2017). A Unified Approach to Interpreting Model Predictions. Retrieved from <https://proceedings.Neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abst.html>
 - ÖZKURT, C. (2024). Comparative Analysis of XAI Techniques on Telecom Churn Prediction Using SHAP. *Research Square*. doi: <https://doi.org/10.21203/rs.3.rs-3917215/v1>
 - Saleh, M.E.A. (2024). Improved Decision Tree, Random Forest, and XGBoost Algorithms for Predicting Client Churn in the Telecommunications Industry. 15(12). Retrieved from https://thesai.Org/Downloads/Volume15No12/Paper_68-Improved_Decision_Tree_Random_Forest_and_XGBoost_AI_algorithms.pdf
 - Samarasinghe, K. (2016). Churn Management in the Sri Lankan Mobile Market. *Engineer: Journal of the Institution of Engineers, Sri Lanka*. doi:<https://doi.org/10.4038/ENGINEER.V41I1.7082>
 - Waweru, T. (2025). *Tridenstechnology.com*. Retrieved from Why Telecom Customers Churn and How to Measure it?: <https://tridenstechnology.com/telecom-churn/>



High Labor Turnover in the Garment Manufacturing Industry: Evidence from a Leading Apparel Factory in North Western Province of Sri Lanka

Sewwandi MP¹, Jayasena DM²

^{1,2}Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka
1999pershanisewwandi@gmail.com¹

ABSTRACT

This study identifies reasons behind high labor turnover in the garment industry through a case study of a leading factory in North Western Province, Sri Lanka. The continuous resignation of workers disrupts production and increases costs. The research employed a qualitative case study methodology with semi-structured interviews of 18 former employees using purposive sampling. Guided by Herzberg's Two-Factor Theory, Maslow's Hierarchy of Needs, and the Push-Pull Theory, the study examined both financial and non-financial factors affecting turnover decisions. Thematic analysis revealed that wage dissatisfaction was reported by 67% of respondents, poor supervisory practices by 50%, lack of career development by 44%, poor workplace facilities by 39%, and excessive overtime work by 56%. Analysis reveals that the turnover decisions stem from a combination of financial dissatisfaction, demanding work conditions, unsupportive supervisory practices, limited career progression, and inadequate workplace facilities. The study recommends improving wage structures, strengthening supervisor training, reduce excessive overtime, and enhancing career development and welfare provisions to improve employee retention in the garment sector

Keywords: Employee Retention, Garment Industry, Labor Turnover, Sri Lanka

1. INTRODUCTION

Labor turnover remains a persistent challenge in the garment industry, with direct implications for productivity, quality, and continuity of operations. Most companies in the apparel sector face high labor turnover as their main operational challenge. The factory examined in this study has experienced an unusually high rate of employee exits in recent months, prompting management to reassess its workforce-related challenges. Although prior research has examined labor turnover in Sri Lanka's garment industry, most studies provide sector-level analysis with limited attention to factory-specific drivers and employee-centered insights. As a result, there is insufficient empirical understanding of the contextual and organizational factors influencing turnover at the individual

factory level. This gap restricts management's ability to design evidence-based strategies that respond to the unique challenges faced by their workforce. Therefore, this study addresses this gap by identifying the key factors contributing to high labor turnover in a leading garment factory in the North Western Province and by proposing practical measures to reducing the turnover rate. This research is significant in a labor-intensive industry where workforce stability is essential; high labor turnover leads to increased training costs, lower morale among remaining staff, and disrupts production flow. Understanding the reasons for employee departures is therefore critical for ensuring operational efficiency and sustaining competitiveness in Sri Lanka's apparel manufacturing sector.

2. LITERATURE REVIEW

Labor turnover refers to the frequency at which employees voluntarily or involuntarily leave an organization and are replaced by new hires. For labor-intensive industries such as garment manufacturing, labor turnover is not merely a human resource concern but a major operational challenge that directly impacts productivity, product quality, and overall competitiveness (Jayasinghe, 2021). Labor turnover can be classified as voluntary or involuntary. Voluntary turnover occurs when an employee chooses to leave the organization, often for better pay, improved working conditions, or personal reasons. Involuntary turnover is initiated by the employer due to layoffs, misconduct, or redundancy (Smith, 2018). The global garment sector is characterized by high pressure, low margins, and fast-paced production cycles. These conditions contribute to persistently high turnover rates, especially in low-income countries. The International Labor Organization (ILO, 2021) reports that inadequate social protection, minimal job security, and high occupational stress are common issues. In Bangladesh and Vietnam, turnover rates in garment factories can exceed 50% annually (Ferdous & Rafiq, 2017). Sri Lanka's garment sector employs hundreds of thousands of workers, many of whom are women from rural areas. Despite having labor laws and employee protections in place, high turnover remains a persistent issue. Jayasinghe (2021) highlights how cultural expectations, family responsibilities, low wages, and lack of recognition contribute to this phenomenon. Numerous studies identify a combination of personal,

organizational, and environmental factors contributing to labor turnover. Compensation remains a primary concern. Khan and Rahman (2019) found that low pay and inconsistent bonuses were key motivators for quitting in the Bangladesh garment sector. Long hours, repetitive tasks, lack of breaks, and inadequate health and safety measures are common. ILO (2021) indicates that such conditions lead to both physical and mental burnout. Wickramasinghe (2020) noted that limited opportunities for training and promotion were major turnover drivers in Sri Lankan garment firms. Poor grievance redressal systems, lack of feedback mechanisms, and authoritarian supervisory styles contribute to dissatisfaction (Gunasekara, 2022). This study is grounded in Herzberg's Two-Factor Theory (1959), which distinguishes between hygiene factors such as pay and work conditions, and motivators such as recognition and responsibility. Maslow's Hierarchy of Needs (1943) suggests that workers need to fulfill basic physiological and safety needs before aspiring to higher levels. The Push-Pull Theory explains that workers are pushed by dissatisfaction and pulled by better opportunities elsewhere.

3. METHODOLOGY

This study employs a qualitative case study methodology to explore the reasons for high labor turnover in the Sri Lankan garment industry. A qualitative case study was selected as the most appropriate research design because this study aims to understand employees' lived experiences and personal narratives behind turnover decisions. The population for this study includes all workers who have resigned from the selected garment

manufacturing company. A purposive sampling technique was used to select participants who were willing to share their reasons for leaving the company. Data were gathered using semi-structured interviews conducted in person and via telephone depending on participant availability. Open-ended questions explored the participant's experience, job conditions, management interaction, and personal reasons for leaving. Sample interview questions included: When did you join and leave the job? What were your daily experiences like at work? What were the main reasons that led you to leave the job? Were there any specific events or conditions that influenced your decision? Thematic analysis was used to interpret the data. Interview transcripts were coded manually to identify recurring themes and patterns that represent the underlying causes of labor turnover. Ethical considerations were given high priority throughout the study. Informed consent was obtained from all individuals prior to their participation. Confidentiality and anonymity were strictly maintained by using pseudonyms and securely storing collected data. Participants were informed of their right to withdraw from the study at any stage without facing consequences.

4. RESULTS AND DISCUSSION

Data were collected from eighteen former employees across different roles. The responses were analyzed through Person Environment Fit Theory (PEFT) dimensions: person-job fit, person-organization fit, person-supervisor fit, and person-team fit.

Person-Job Fit: Several operators reported skill mismatches and unrealistic performance expectations. One female

operator with 1 year experience stated: "I was trained only for overlock machines, but they assigned me to operations I had never done before and still expected me to meet targets. It's impossible to reach 70-80 pieces per hour without proper practice." Additionally, 56% reported excessive overtime. A male helper with 2 years experience explained: "We often work late into the night even after 8 p.m., and transport is only provided halfway. Even Saturdays are not holidays anymore." This aligns with Siva and Krishnan (2015) who found long working hours influence turnover.

Person-Organization Fit: Wage dissatisfaction emerged as the primary factor with 67% reporting concerns. A female machine operator with 8 months experience noted: "The salary is too low. They promised to increase it, but it never happened. Other factories pay at least Rs. 5,000 more than here." This supports Mbah and Ikemefuna (2012), who linked low satisfaction with wages to turnover. Furthermore, 44% cited a lack of career development. A finishing section worker with 2 years of experience stated: "There are no promotions now. Only production matters. My salary hasn't been increased even once." Yang et al. (2008) confirmed jobs facilitating promotion have reduced turnover tendency.

Person-Supervisor Fit: Poor supervisory practices affected 50% of respondents. A female ironing section worker with 3 years experience revealed: "They shout for no reason. The younger workers don't even want to come anymore. When we already work under stress, they still keep yelling." Ugoani (2016) illustrated unfair managers increase stress and turnover intention.

Person-Team Fit: Mixed experiences were observed regarding teamwork. Some respondents reported supportive colleagues, while others noted selfish and irresponsible team members. However, this factor did not emerge as a primary turnover driver.

General Factors: Workplace facilities influenced 39% of decisions. A male line supervisor with 5 years of experience described: "They only care about output. The machines are old, difficult to run, and the chairs are uncomfortable. I have continuous back pain now". Al-Omari and Okasheh (2017) have confirmed that poor physical conditions increase turnover rates.

The findings strongly support Herzberg's Two-Factor Theory, as hygiene factors including pay, working conditions, and supervision were inadequately met, contributing to dissatisfaction. Maslow's Hierarchy of Needs is evident as basic physiological and safety needs remain unmet through low wages and unsafe conditions. The Push-Pull Theory explains employees are pushed by dissatisfaction and pulled by better opportunities elsewhere, particularly regarding wages where competing factories offer significantly more.

5. CONCLUSION

This study successfully identified reasons for high labor turnover at the selected factory. Results revealed that turnover is driven by a combination of interrelated factors, including low wages, poor working conditions, lack of career development opportunities, unsupportive supervisory practices, and weak workplace welfare. Low wages and poor

benefits remain the primary reason for employees leaving their jobs, with 67% reporting wage dissatisfaction. Many employees expressed frustration at being stuck in the same positions for years without promotion or recognition, with 44% citing lack of career development. Workplace stress driven by excessive overtime (56%), unsafe working conditions, and unrealistic production targets placed immense pressure on employees. Supervisory practices were identified as a major concern with 50% reporting poor supervisory practices that reduced morale and trust. The general workplace environment played a key role with 39% reporting poor workplace facilities that added to physical and emotional strain. Based on findings, several strategies are recommended. The factory should revise wage structures to align with industry standards, reintroduce performance-based incentive schemes, and provide additional allowances for attendance, transport, and meals. Establishing clear promotion pathways, implementing continuous training programs on machine handling and technical skills, and providing leadership training for aspiring supervisors would enhance career development opportunities. Management should limit overtime hours, ensure employees receive rest days as per labor regulations, rotate workers fairly across different tasks, and maintain machinery regularly to minimize breakdowns. Supervisors should be trained in effective leadership, communication, and employee relations. Establishing proper feedback mechanisms would allow employees to voice concerns safely. Creating positive team environments through collaborative work practices, recognition of cooperative

teams, and team-building programs is vital. Management must ensure adequate fans, lighting, and ergonomic seating are provided, maintain clean washrooms and break areas, provide appropriate personal protective equipment, and introduce wellness programs including periodic health checks and counseling support. Implementing retention programs such as regular employee satisfaction surveys, recognition of long-serving employees, and recreational welfare activities would ensure long-term workforce stability.

REFERENCES

- Ferdous, F., & Rafiq, R. (2017). Labor turnover in the garment industry: Causes and consequences. *Journal of Human Resource Management*, 5(2), 45-56.
- Gunasekara, T. (2022). Leadership styles and employee turnover: Evidence from Sri Lanka. *Journal of Organizational Behavior Studies*, 14(1), 56-70.
- International Labor Organization. (2021). *Decent work in global supply chains: Lessons from the garment sector*. ILO Publications.
- Jayasinghe, S. (2021). Employee attrition in Sri Lanka's Free Trade Zones. *South Asian Journal of Management Studies*, 10(3), 87 - 102.
- Khan, A., & Rahman, M. (2019). Factors influencing employee turnover in Bangladesh's ready-made garment industry. *International Journal of Business Research*, 12(4), 78-92.
- Wickramasinghe, N. (2020). HRM practices and labor retention in the apparel industry. *International Review of Management and Business Research*, 9(1), 112-125.



Impact of Packaging Designs on Consumer Purchasing Behavior in the Packaged Food Industry

Sewwandi MP¹, Jothirathne PAAU²

^{1,2} Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka
*1999pershanisewwandi@gmail.com*¹

ABSTRACT

Packaging extremely influences item purchase according to customer preference particularly in the food industry where competing products often have similar quality and price. This study examines how visual and functional packaging designs influence customer preference in purchases in food industry. A structured questionnaire was administered to 385 regular consumers selected through stratified random sampling. Data were analyzed using SPSS Version 26, including descriptive statistics, reliability analysis (Cronbach's alpha), validity testing (KMO and Bartlett's Test), correlation analysis, and multiple regression analysis to assess the predictive influence of packaging attributes. Descriptive analysis was conducted to summarize demographic characteristics and responses to each construct. In addition, Pearson correlation analysis was used to examine the strength and direction of the relationships between visual packaging elements ($\beta = 0.422$, $p < 0.001$), functional packaging elements ($\beta = 0.492$, $p < 0.001$), and consumer purchasing behaviour before conducting regression analysis. Together, these predictors explained 59.6% of the variance in consumer behaviour ($R^2 = 0.596$). The findings highlight that while visual appeal significantly shapes perceptions of attractiveness, quality, and brand image, functional aspects such as reliability, durability, and eco-friendliness have a slightly greater impact on consumer satisfaction and repeat purchases. This study contributes to existing literature by simultaneously assessing visual and functional packaging dimensions, offering a comprehensive perspective. It recommends that organizations prioritize enhancing functional features while also investing in attractive visual design to strengthen consumer engagement, foster brand loyalty, and maximize competitiveness in the food market.

Keywords: Consumer Behaviour, Functional Elements, Packaging Design, Visual Elements

1. INTRODUCTION

Packaging has evolved beyond its traditional role of protecting products to become a strategic marketing instrument within the modern packaged food industry. Often described to as the "silent salesman," packaging communicates product value, quality, and brand identity without direct human interaction. In increasingly competitive markets where food products frequently share comparable quality and price attributes,

packaging design becomes a decisive factor in brand differentiation and consumer choice.

The packaging of food products serves as the first point of contact between consumer and the brand significantly shaping initial perceptions and influencing purchasing behaviour. Researchers indicated that consumers often make purchase decisions within a few seconds of encountering a product on the shelf, highlighting the critical

importance of visual designs (colour schemes, typography, imagery) and functional attributes (resealability, portability, eco-friendliness). This study investigated how visual packaging elements and functional features collectively influence consumer purchasing behaviour in the packaged food industry. This aimed to assess the relative impact of these design dimensions on consumer decision-making and to propose practical strategies for optimizing packaging designs that enhance product appeal, strengthen brand competitiveness, and promote consumer loyalty in the packaged food sector.

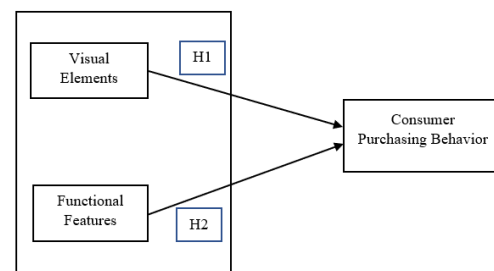
2. LITERATURE REVIEW

Packaging serves as more than a container; it functions as a powerful marketing and communication tool that confirms psychological and emotional connections with consumers. According to Rajkumar and Jain (2021) visual elements such as colour, typography, and imagery act as stimuli that evoke trust, excitement, or nostalgia, thereby shaping purchase intentions. Similarly, Ahmad et al. (2012) identified colour as the most influential packaging attribute, followed by imagery and labelling, with colours creating strong psychological associations that enhance product recall. Mazhar et al. (2015) further observed that packaging which effectively integrates appealing visuals with durability and practical design builds consumer confidence and strengthen brand credibility. Santhoshi et al. (2023) found that younger consumers prioritize both functionality and sustainability, with eco-friendly packaging gaining increased importance among eco-conscious consumers. Al-Tunisi and Bagwell (2022) highlighted

that effective packaging design enhances product differentiation and communicates brand values, particularly critical in the packaged food industry where authenticity and freshness perceptions matter. The literature highlighted that packaging integrates visual aesthetics, functional convenience, and sustainability to influence consumer perceptions and influencing purchasing behavior.

3. METHODOLOGY

This study adopted a quantitative research design to investigate the



influence of packaging design elements on consumer purchasing behavior.

Primary data have collected through a structured questionnaire. The conceptual framework identified visual elements and functional features as independent variables and consumer purchasing behaviour as the dependent variable.

Two hypotheses have formulated:

H1: Visual elements of packaging positively influence consumer purchasing behaviour.

H2: Functional features of packaging positively influence consumer purchasing behaviour.

The total of 385 (according to The Morgan table) respondents—had been selected using stratified random sampling to ensure demographic diversity across age, gender, and income groups. The

survey instrument employed five-point likert scale to measure respondents' attitudes. The questionnaire distributed both online and physically. Analysis was performed using SPSS, including reliability testing through Cronbach's alpha, validity assessment via Kaiser-Meyer-Olkin measure and Bartlett's test, and multiple regression analysis to examine relationships between packaging attributes and consumer purchase behaviour.

4. RESULTS AND DISCUSSION

A pilot test with 70 responses confirmed the internal reliability of the measurement scale, with Cronbach's alpha values exceeding 0.7 for all variables after minor adjustments. The final data set has included 385 valid responses. Age distribution of the respondents is 35% aged between 18-25 years, 44% aged 26-35, 16% aged 36-45, and 5% above 45 years, The sample included 64% female and 36% male participants. Data cleaning procedures identified outliers, which have retained as they represented valid consumer perspectives rather than measurement errors. Validity testing confirmed appropriateness of factor analysis with KMO values and significance levels below 0.001 for all variables. Cumulative variance percentages exceeded 50%, validating the adequacy of the measurement constructs.

Table 1: Regression Model Summary

R	R ²	Adjusted R ²	Std. Error	F	Sig.
.772	.596	.593	.349	281.2	<.001

The overall regression model has confirmed predictive power, with $R^2 = 0.596$ which denotes that approximately

59.6% of the variance in consumer purchasing behavior has explained by the two predictors visual and functional packaging elements.

Table 2: Regression Coefficients

Variable	B	β	t	Sig.	VIF
Constant	.029	-	.178	.859	-
Visual Elements	.400	.422	10.225	<.001	1.216
Functional Features	.564	.492	11.931	<.001	1.216

Correlation analysis indicated strong positive relationships between: visual elements and consumer behavior ($r = 0.630$, $p < 0.01$) and functional features and consumer behavior ($r = 0.670$, $p < 0.01$). The regression analysis further illustrated that; Functional features ($\beta = 0.492$) have a stronger impact than visual elements ($\beta = 0.422$). Together, both predictors explained 59.6% of the variance in consumer purchase behavior, confirming a considerable and statistically significant relationship.

These findings confirmed that visual elements significantly influenced consumer perceptions of quality, attractiveness, and brand image, consistent with literature demonstrating packaging aesthetics as persuasive communication tools. Functional attributes including resealability, eco-friendliness, and material durability showed strong associations with consumer satisfaction and loyalty, extending influence beyond initial purchase to repeat buying behaviour. The results also highlighted growing consumer awareness regarding sustainable packaging practices, reflecting broader industry trends emphasizing environmental responsibility. Both visual appeal and

functional utility function as interdependent dimensions collectively shaping consumer behaviour from initial evaluation through brand loyalty development.

5. CONCLUSION

The empirical results supported by reliability, validity, correlation, and regression analyses confirm that both visual and functional packaging elements significantly influence consumer purchasing behaviour. The findings confirmed that visual elements including colour, typography, and imagery play a critical role in shaping consumer perceptions of product quality, attractiveness, and brand image. These visual cues act as a powerful differentiator in competitive retail environments where consumers often make rapid purchase decisions. Functional attributes such as resealability, eco-friendliness, and durability have found to enhance consumer satisfaction and convenience, and trust. These features contribute to the initial purchase decision and to repeat purchase behavior and long-term brand loyalty. The results thus supported both hypotheses, confirming that packaging design functions as a strategic determinant of consumer decision making, brand differentiation, and market competitiveness in the packaged food sector.

From a managerial perspective, organizations have encouraged to invest in innovative packaging designs that integrate aesthetic appeal with functional usability. Designing packaging that harmonizes visual attraction, and practical utility can strengthen brand image and enhance consumer engagement. Furthermore, companies should position

sustainability and eco-friendly materials as strategic advantages, responding to growing consumer awareness of environmental responsibility and aligning with global trends toward sustainable consumption.

Future research could expand by exploring the role of labelling and nutritional information in influencing purchase behavior, exploring cross-cultural variations in packaging preferences, and examining the impact of digital packaging innovations such as QR codes or smart labels on consumer engagement.

REFERENCES

- Ahmad, N., Billoo, M., & Lakhan, A. (2012). Effect of product packaging in consumer buying decision. *Journal of Business Strategies*, 6(2), 1 - 10.
- Al-Tunisi, I. A., & Bagwell, T. J. (2022). Influence of product packaging on consumer purchase behavior among university undergraduates in Pakistan. *International Journal of Marketing Management Sciences*, 1(2).
- Magnier, L., & Cri  , D. (2015). Communicating packaging eco-friendliness: An exploration of consumers' perceptions of eco-designed packaging. *International Journal of Retail & Distribution Management*, 43(4/5), 350 - 366.
- Mazhar, M., Daud, S., Bhutto, S.A., & Mubin, M. (2015). Impact of product packaging on consumers' buying behavior: Evidence from Karachi. *Journal of Marketing and Consumer Research*, 16, 35-42.
- Rajkumar, A., & Jain, V. (2021). A literature study on the product packaging influences on the customers' behavior. *Journal of Contemporary Issues in Business and Government*, 27(3), 779 - 785.
- Rundh, B. (2016). The role of packaging within marketing and value creation. *British Food Journal*, 118(10), 2491-2511.
- Santhoshi, E., Bindu, E.S.H., Pavani, N., & Kodali, S. (2023). Influence of

IMPACT OF PACKAGING DESIGNS ON CONSUMER PURCHASING BEHAVIOR IN THE
PACKAGED FOOD INDUSTRY

packaging on the buying decisions of college students. *Biological Forum – An International Journal*, 15(10), 182 - 186.



Factors Affecting to the Sola Panel Adoption in Sri Lanka

Bimsara MU¹, Jayasena DM²

^{1,2} Department of Industrial Management, Faculty of Applied Sciences, Wayamba University of Sri Lanka
Bimsaramunasinghe1@gmail.com

ABSTRACT

With the increase of global fossil energy prices after the COVID-19 pandemic, the world is moving to the sustainability of renewable energy sources, and it's becoming one of the key global technologies with an affordable cost. After the economic crisis in Sri Lanka, the country has more potential for solar adoption. Because people are financially stable. But the problem is that still solar adoption rate is still very low (10.16%). There are some reasons behind this lower adoption rate. The main objective is to find the factors affecting the lower adoption of solar panels. This study mainly focuses on the factors like economics, knowledge and attitudes, and government impact for the adoption of solar energy by individual responders. This research was conducted across Sri Lanka with a sample size of 381 respondents using a stratified sampling technique, and the structural equation model and correlation analysis to interpret the results. The results showed that people's perceptions and expectations regarding solar systems are very good, and economic concerns are significantly high. Respondents are not very positive about government intervention in solar adoption, and they are not satisfied with government support. This research is supposed to identify solutions to these problems, like awareness of the rural people, easy access to loans, rules, and regulations, with the support of policymakers. To sustain positive attitudes towards solar systems in society, the government and policymakers should introduce a convenient and affordable program for homeowners.

KEYWORDS: Government Impact, Knowledge and Attitudes, Renewable Energy, Solar Adoption

1. INTRODUCTION

With the development of modern technology, the world is quickly moving towards sustainability adoptions. As a result of that world interest in adopting some renewable energy sources based on sustainability aspects and financial benefits. (Bandara, 2018). Most of countries move to renewable energy sources like solar panels. But in the Sri Lankan context, it is still struggling with the past economic crisis of Sri Lanka. But now it is moving in the right way. The purpose of this research is to find what are the issues that are affecting the solar panel adoption in Sri Lanka and propose the appropriate solutions based on the results. Solar energy is abundant and sustainable, especially in tropical countries like Sri

Lanka, which receives consistent solar radiation throughout the year.

Despite the enormous potential for utilising solar energy in Sri Lanka, the current adoption at the household level is far from optimal. Lack of knowledge and attitudes towards the solar panel, struggling with some financial and economic problems and government lower impact and tough rules and regulations are some of the major factors contributing to the very low diffusion rate of solar panels.

This study provides complete solutions based on these three main factors because they cover all the problems that can arise from solar panel adoption. In the Sri Lankan context, people mainly use solar panels to generate electricity connected to the grid. The main

solar panel systems are monocrystalline and polycrystalline. This research investigates the knowledge of these two types of solar panels among responders and their financial possibilities, and the opinion of the government's impact on

2. LITERATURE REVIEW

The adoption of solar energy technology in households has been widely studied across various global contexts, revealing numerous barriers and drivers influencing its acceptance. Solar energy, being abundant and sustainable, is particularly suited for tropical nations like Sri Lanka (Bandara, 2018). However, despite its potential, household-level adoption remains suboptimal, primarily due to financial, economic, technological, and socio-cultural challenges.

It employs a dual-method approach (SEM and ANN) to analyse financial, technical, and awareness-related factors affecting adoption. This study explores solar panel adoption in Sri Lanka, emphasizing variables like economy, knowledge and attitudes, and government impact. This study differentiates by addressing Sri Lanka's unique socio-economic context, policy environment, and enriching literature on regional disparities in renewable energy adoption. (Ravindu Yahanpath, W.M.C.L.K. Wijekoon, 2024)

Government policies and incentives significantly influence adoption rates. The Renewable Energy Efficiency Fund in Jordan, which covers a substantial portion of solar PV installation costs, is a successful example of policy-driven adoption. (Amin Al-Habaibeh a, Bashar Al-haj Moh'd b, Hiba Massoud c, 2023) Sri Lanka's National Energy Policy

recognizes the importance of diversifying energy sources but requires more robust frameworks to overcome socio-economic barriers and incentivize solar energy adoption

This study explores strategic approaches to optimize consumer adoption and increase the utilization of energy generated by owned solar panels. This study focuses on the key factors influencing consumer behavior, including the financial benefits and government support and based on their Knowledge

3. METHODOLOGY

The current research utilized a quantitative method using primary data that was obtained from Households in Sri Lanka using the online questionnaire within one month period. The quantitative data in this study were analyzed using SPSS software and Smart PLS, and its results are reported statistically. A Spearman rank one-tailed correlation is conducted, which identifies the strength and corresponding direction between the variables.

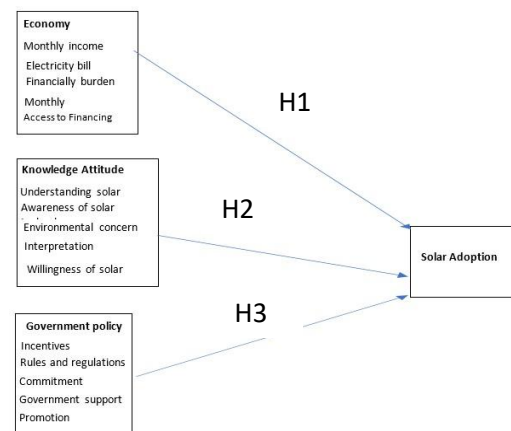


Figure 1: Conceptual Framework

Descriptive statistics such as means, standard deviations, and percentages were calculated based on demographic characteristics: gender, education, age

group, occupation, province, monthly electricity bill, and household income. The rest of the questionnaire measured five variables representing the Economic factors, Knowledge and attitudes, government impact in relation to the adoption of solar energy. Before all the calculations, test the validity, reliability and outlier tests to get the accurate results.

Structural equation model, correlation analysis and descriptive statistics were calculated to identify the relationship of those variables with solar panel adoption.

The structural equation model was calculated by using SmartPLS. Descriptive analysis and correlation analysis were calculated by using SPSS software.

4. RESULTS AND DISCUSSION

4.1 Responders Demographic Analysis

The demographic analysis of all the responders is helped to identify their background how important to this research.

Sex					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	144	37.8	37.8	37.8
	Male	237	62.2	62.2	100.0
	Total	381	100.0	100.0	

Figure 2: Gender

Age_Range					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	20-29	93	24.4	24.4	24.4
	30-39	91	23.9	23.9	48.3
	40 - 49	119	31.2	31.2	79.5
	50-59	43	11.3	11.3	90.8
	60 or above	35	9.2	9.2	100.0
	Total	381	100.0	100.0	

Figure 3: Age Range

Highest_Educational_Qualification					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Advance Level	132	34.6	34.6	34.6
	BA/BSc	98	25.7	25.7	60.4
	Diploma	49	12.9	12.9	73.2
	MA/MSc	21	5.5	5.5	78.7
	Ordinary Level	43	11.3	11.3	90.0
	Other	32	8.4	8.4	98.4
	PhD	6	1.6	1.6	100.0
	Total	381	100.0	100.0	

Figure 4: Educational Qualification

Province					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Central Province	47	12.3	12.3	12.3
	Eastern Province	21	5.5	5.5	17.8
	North Central Province	39	10.2	10.2	28.1
	North Western Province	35	9.2	9.2	37.3
	Northern Province	32	8.4	8.4	45.7
	Sabaragamuwa Province	42	11.0	11.0	56.7
	Southern Province	95	24.9	24.9	81.6
	Uva Province	24	6.3	6.3	87.9
	Western Province	46	12.1	12.1	100.0
	Total	381	100.0	100.0	

Figure 5: Province

Occupation					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Businessman	67	17.6	17.6	17.6
	government worker	96	25.2	25.2	42.8
	Non-Government Worker	95	24.9	24.9	67.7
	Other	71	18.6	18.6	86.4
	Retired	52	13.6	13.6	100.0
	Total	381	100.0	100.0	

Figure 6: Occupation

Household_Size					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	8	2.1	2.1	2.1
	2	48	12.6	12.6	14.7
	3	78	20.5	20.5	35.2
	4	137	36.0	36.0	71.1
	5	76	19.9	19.9	91.1
	More than 5	34	8.9	8.9	100.0
	Total	381	100.0	100.0	

Figure 7: Household Size

House_Type					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Other	3	.8	.8	.8
	Private flat	1	.3	.3	1.0
	Private house	348	91.3	91.3	92.4
	Rented flat	8	2.1	2.1	94.5
	Rented house	21	5.5	5.5	100.0
	Total	381	100.0	100.0	

Figure 8: House Type

According to the above 6 figures, most of the responders' backgrounds are suitable for solar panel adoption. That

means they have the possibility for solar panel adoption. As an example, most of the responders' house type is private house. That means they can go for solar adoption. There are some other reasons behind that. But they have some possibility. This demographic analysis helps to get a basic idea of the responders before moving to the analysis part. So, as the summary responder's educational level, province, age range, house type, and household size are in good positions for the solar panel adoption in Sri Lanka.

4.2 Descriptive Statistics

Descriptive statistics provide a summary of all the independent and dependent variables. According to Table 01, the results show all the variables' mean values are close to 4 on 5 5-point scale. This suggests a positive perception. Economic Factors ($M = 3.902$, $SD = 0.739$) show that responders, on average, moderately consider that economic considerations are for solar adoption. Knowledge and Attitudes ($M = 3.897$, $SD = 0.728$) reflect a high level of awareness and favourable attitudes on solar energy. That means the knowledge and attitudes of the solar adoption of the people are quite good. Government Impact ($M = 3.838$, $SD = 0.730$) shows that responders perceive government policies and influence as moderately supportive in shaping solar adoption decisions. The dependent variable, Solar Adoption ($M = 4.001$, $SD = 0.774$), has the highest mean score, indicating that respondents have a good perception of solar.

	N	Descriptive Statistics			
		Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic
Economic_Factors	381	2.2	5.0	3.902	.7388
Knowledge_Attitudes	381	2.2	5.0	3.897	.7282
Government_Impact	381	2.4	5.0	3.838	.7304
Solar_Adoption	381	2.00	5.00	4.0007	.77428
Valid N (listwise)	381				

Table 1: Descriptive Statistics

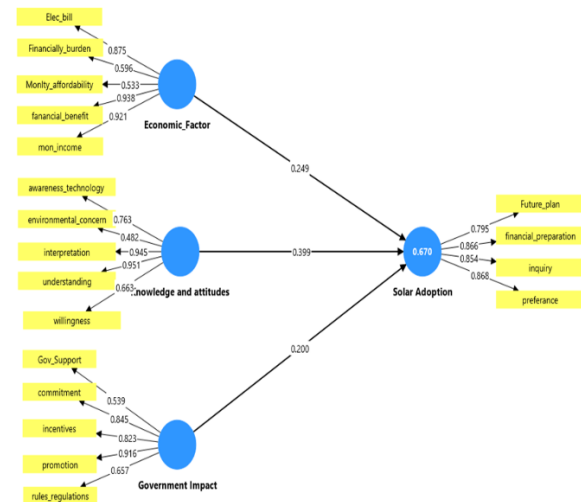


Figure 9: Structural Equation Model

This structural equation model indicates how these three independent variables influence solar adoption, with the overall explanatory power of $R^2 = 0.670$. This shows that 67% of the variation in solar adoption can be explained by these three variables.

According to this model, the highest influence factor is knowledge and attitudes. It is nearly 0.399. means responders' knowledge and attitudes are quite good for the solar adoption, including their awareness, environmental concern, understanding and willingness.

Economic factors moderately influence solar adoption. Nearly 0.249. That means responders have a moderate economic background for solar, including the responder's monthly affordability, financial burden of electricity bill, monthly electricity bill, and considering

financial benefits from solar panels and their monthly income.

Government impact is the lower influence factor for solar panel adoption. (0.200). that means responders are dissatisfied with the government's impact on solar adoption, including their support, commitment, incentives, promotion, rules and regulations.

Correlations					
		Economic_Factors	Knowledge_Attitudes	Government_Impact	Solar_Adoption
Economic_Factors	Pearson Correlation	1	.845**	.782**	.740**
	Sig. (2-tailed)		.000	.000	.000
	N	381	381	381	381
Knowledge_Attitudes	Pearson Correlation	.845**	1	.780**	.740**
	Sig. (2-tailed)	.000		.000	.000
	N	381	381	381	381
Government_Impact	Pearson Correlation	.782**	.780**	1	.731**
	Sig. (2-tailed)	.000	.000		.000
	N	381	381	381	381
Solar_Adoption	Pearson Correlation	.740**	.740**	.731**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	381	381	381	381

** Correlation is significant at the 0.01 level (2-tailed).

Figure 10: Correlation Analysis

Correlation analysis shows how each independent variable correlates with variables. According to Figure 9, all the variables have a strong positive relationship with the dependent variable, solar adoption. All the correlation values are positive and strong. As well as all the p-values are less than 0.05. Therefore, we can reject the null hypothesis. That means all the variables have a significant relationship with the dependent variable.

5. CONCLUSION

This study provides a comprehensive examination of Solar panel adoption in Sri Lanka based on the three main factors that can affect solar panel adoption. According to the results and interpretation, the variables such as economics, knowledge and attitudes, and government impact are positively affected for the solar panel adoption in Sri Lanka. increasing the impact of these three variables can improve the adoption rate.

Based on these results most impactful factor is knowledge and attitudes. We should try to inform the advantages of solar adoption for the rural people as well. The government should focus more on their impact on solar adoption.

Because people don't have a good perception of the government's impact on solar adoption. As well as if people have any financial facilities like bank loans, adoption can go high. People have good perceptions about solar panels. But they are struggling with some issues, like financial issues, long waiting time periods, and high initial costs. If we can mitigate these problems, household owners are motivated for the solar adaptation

For the high solar panel adoption, Sri Lanka should prioritize strategies such as:

- Introduce low-interest loan schemes and flexible instalment payment plans.
- Provide tax reductions or subsidies for purchasing solar systems
- Provide training programs for technicians and installers to improve service quality
- Implement clear and stable renewable energy policies with long-term guarantees.
- Streamline approval and installation processes to reduce delays
- Partner with private companies through public-private partnerships to expand access to solar technologies
- Conduct awareness workshops in both urban and rural communities

The above suggestions are possible and reliable solutions to increase solar panel adoption in Sri Lanka.

REFERENCES

- Amin Al-Habaibeh a, Bashar Al-haj Moh'd, Hiba Massoud, Onyebuchi Benjamin Nweke a, (2023). Solar energy in Jordan: Investigation Challenges and opportunities of using a domestic solar energy system. *World Development Sustainability*.
- Bandara, U. (2018). Factors Influencing Solar Energy Technology Adoption by Households in Western Province in Sri Lanka. 15th International Conference on Business Management (ICBM 2018).
- Dheerasingha D.A.H.N., Jayasekara K. D. De. S., Gunawardena U.A.D.P. (2024). Factors Influencing Photovoltaic Solar Technology Adoption Among Households in Hambantota District, Sri Lanka. Proceedings of the 10th Undergraduate Research Symposium.
- Nadarajah Kannann, DivagarVakeesan. (2016). solar energy for future world: -A review. *Renewable and Sustainable Energy Reviews*.
- Ravindu Yahanpath, W.M.C.L.K. Wijekoon, (2024). Exploratory Study on Adaptability.
- Smita Nanasaheb Warhade, Dr. Amit Aggrawal. (2024). Optimizing Consumer Adoption: Strategic Approaches to Increase Utilization of Solar Energy from Owned Solar Panels. *International Research Journal on Advanced Engineering Hub (IRJAEH)*.



Workplace Ostracism and Deviant Workplace Behaviour: Systematic Literature Review Based on PRISMA Method

Weersinghe RD¹, Janadari MPN²

^{1,2} *Department of HRM, Faculty of Commerce & Management Studies, University of Kelaniya, Sri Lanka*
wprdi231@kln.ac.lk¹, njanadari@kln.ac.lk²

ABSTRACT

Employees are the active resources of every organization, and the way they are treated brings sustainable outcomes to the organization in the future. However, any form of mistreatment could lead them to deviate from their work, creating threats to overall organizational well-being. This conceptual review aims to explore the prevailing literature on workplace ostracism and deviant workplace behaviour and to investigate future research opportunities amidst Industry 4.0. A Systematic Literature Review (SLR) using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method was utilized to review the articles in the literature. To meet the research aims, 265 published documents from 134 sources were retrieved from the Scopus database within the period from year 2015 to 2025 to ensure the quality of the input, and bibliometric analysis was used to perform keyword co-occurrence to reflect areas that require further research attention. The review highlights the significant effect of workplace ostracism on an employee's psychological and physical behaviour, distancing him or her from workplace tasks. Further, the results revealed that scholarly attention should be directed towards leadership, culture, and individual traits to strengthen citizenship behaviours at work. In conclusion, this review provides valuable insights for policy formulation, future research, and practical applications to promote sustainable and harmonious workplace behaviours in the era of Industry 4.0.

KEYWORDS: Deviant Behaviour, PRISMA, Workplace Behaviour, Workplace Ostracism

1. INTRODUCTION

Every organization achieves its highest performance with the support of its workforce, which is fueled by psychological and physical well-being at work (BELLA, 2023). Employees are the key drivers of organizations, guiding other resources, including human resources, toward the organization's desired future. Therefore, the way they are treated has drawn the attention of scholars over the past decades, as it shapes employees' behaviours and influences their engagement in productive tasks.

Workplace Ostracism (WO), also known as workplace exclusion, is the perception employees have of being ignored or excluded by others in the organization (Robinson & Schabram, 2019). It is a passive form of mistreatment

toward employees, causing them to become isolated from the social environment in the workplace.

This workplace mistreatment underscores several possible reasons, categorized into three main areas: intrapersonal (individual level, such as personality traits), interpersonal (relational level, such as conflict, jealousy, or rivalry within the team), and social and organizational (context level, such as workplace conditions, culture, structures, or leadership) characteristics. Further, there are internal factors (personal personality) and external factors (organizational climate) that play a crucial role in WO (Zhu, Shu, Feng, & Tan, 2019).

As Zheng and Yang (2016) showed in their study, WO is strongly correlated

with Deviant Workplace Behaviours (DWB) among employees (Jahanzeb, Giles, & Mushtaq, 2023), causing adverse deterioration in their performance, well-being, and organizational perceptions.

2. LITERATURE REVIEW

According to the prevailing literature, WO is considered a “social death,” reflecting its significant impacts on individuals’ attitudes, well-being, and behaviours (Dash, Ranjan, Bhardwaj, & Rastogi, 2024). Moreover, it has been reported that there is a strong positive relationship between WO and employees’ propensity to engage in DWB as well (Attia, Saleh, & Mohamed, 2025). However, the effects of WO on belongingness, job satisfaction, emotional exhaustion, interpersonal deviance, and organizational deviance have shown a strong association in individualistic contexts, while its relationship with organizational identification and organizational citizenship behaviour becomes stronger in collectivist contexts (Li et al., 2021).

Meanwhile, the study conducted by Ruslan and Wahat (2024) stated that addressing the primary drivers of deviant behaviour—such as leadership styles, organizational justice, and the workplace environment—through targeted interventions could reduce workplace deviance, ultimately enhancing organizational performance and employee satisfaction. Ramsey and Jones (2015) indicated that WO could be prevented through Mindfulness-Based Interventions (MBIs), which help individuals stay connected to present-moment attention and awareness. Further, through MBIs, employees may experience non-

judgmental and non-reactive responses in their workplace behaviour as well.

WO is regarded as a universal phenomenon that creates social pain for individuals at work. An individual employee may experience disturbances in psychological well-being when he or she encounters resource loss or the threat of potential loss, as they perceive resources as valuable, according to the Conservation of Resources Theory (Oberai, 2021). In addition, Zhou and Rainone (2021) reported that Social Identity Theory and Sociometer Theory are used to explain prosocial organizational behaviours, while antisocial organizational behaviours are identified using the Need to Belong Theory, the Transactional Stress Model, Social Exchange Theory, and Self-Verification Theory. Moreover, these theories also suggest several moderators and mediators.

2.1 *Prosocial organizational behaviors*

Organizational factors (person-organization fit), social factors (social support), and psychological well-being variables (self-esteem) are examples.

2.2 *Antisocial organizational behaviors*

The examples are organizational factors (360 feedback), social factors (high team identification), and individual differences (personality and self-esteem). Especially during the COVID-19 period, WO became more severe as employees were restricted to their homes or confined to virtual spaces. However, moral leadership was able to reduce WO by 79.3% and counterproductive behaviour by 36.7% (Ali & Sayed, 2023).

As prevailing studies show, both WO and DWB have become critical areas of investigation in organizational behaviour research amidst Industry 4.0,

emphasizing the continued need for systematic literature reviews to identify trends, fill gaps, and suggest possible areas for future research (Adekanmbi & Ukpere, 2022). This Systematic Literature Review (SLR) has been conducted with the following objectives: to investigate how these concepts best fit within Industry 4.0, including technologies such as Artificial Intelligence (AI) and Virtual Reality (VR), and to examine their contributions to SDG Goal No. 3: Good Health & Well-Being, Goal No. 10: Reduced Inequalities, and Goal No. 16: Peace, Justice, and Strong Institutions.

- To identify the current knowledge on WO and DWB.
- To identify the keywords and themes where empirical research is stronger in centrality and density.
- To identify in which context research has primarily been conducted on WO and DWB.

3. METHODOLOGY

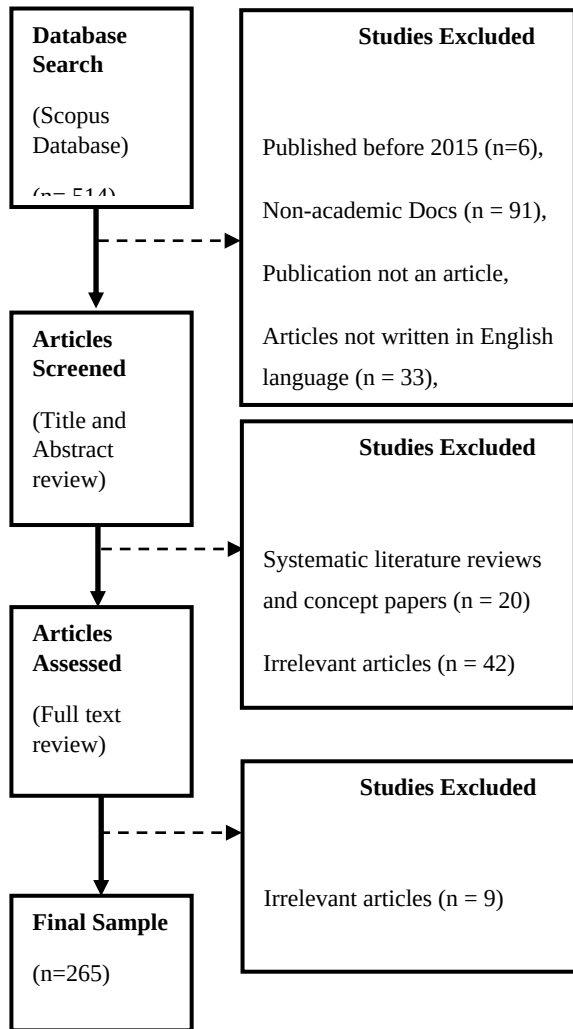
To achieve the research objectives mentioned above, the SLR was conducted using 265 high-quality research papers published in English from 2015 to 2025. This time range was selected to provide a recent perspective on how these concepts have been challenged prior to the upcoming technological revolution. According to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) article selection guidelines, the documents were retrieved from the well-established Scopus database, which is considered to have the highest number of publications with significant citations in scientific fields (Moed, Bar-Ilan, & Halevi, 2016). This paper undertakes a systematic

investigation of the literature concerning WO and DWB, with a specific focus on their influence and outcomes. The study spans diverse aspects, including HRM, management, corporate governance, leadership, workplace behaviour, and employee psychological and overall well-being, to examine their connectivity with employee satisfaction and organizational performance.

In this paper, the SLR process was adopted following the PRISMA guidelines, where the objective selection and analysis of articles were conducted according to the stages illustrated in Figure 01 (Liberati et al., 2009). This approach helps to identify the most relevant documents for review.

3.1 Identification

During the identification stage, the databases, search phrases, and search criteria were determined. High-impact publications were retrieved from Scopus. By using the "AND" and "OR" operators with the most popular search terms; "workplace ostracism," "ostracism," "incivility," "counterproductive work behaviour," "workplace deviant behaviour," "deviant workplace behaviour," and "ethical leadership"—a total of 514 sources were initially identified before being checked for duplicates.



(Source: Author's compilation)

Figure 1: PRISMA Framework

3.2 Screening

At the screening stage, the inclusion and exclusion criteria were applied to filter the articles, as depicted in Table 01. These criteria ensure sufficient consistency in methodological quality to produce relevant results that meet internal validity requirements (Petticrew & Roberts, 2006).

Table 1: Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Published between 2015 and 2025	Publication before the year 2015

Academic journals (inherent academic rigor, quality control via the peer-review process)	Non-academic databases such as books, book chapters, Magazines, conference papers, and editorials
Publication as an article	The publication is not an article.
Written in English	written in any other language
Availability of access to full docs	Unavailability of access to full papers

(Source: Author's compilation)

3.3 Ensuring Eligibility

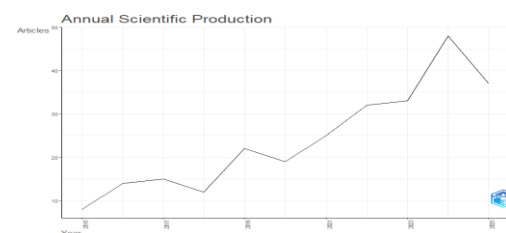
The evaluation was based on research discipline, population, sample, methodology, methods, design, techniques, and related factors. Documents were excluded if they employed unreasonable methodologies or contained confusing and ambiguous approaches that required additional information from the author(s). All exclusions were made through the collective agreement.

3.4 Deciding on Inclusion

Finally, at the inclusion stage, the final sample was determined through consensus among the authors regarding the general relevance and quality of the eligible articles.

4. RESULTS

4.1 Annual Production of Articles



(Source: Author's compilation)

Figure 2: Annual Production of Articles

The yearly release of articles illustrates a 16.5% annual growth rate of these concepts and demonstrates how they have evolved over the years. The publication trend for WO and its impact on DWB in the Scopus database from 2015 to 2025 is depicted in Figure 2. Most articles were published after 2015, with 2023 and 2024 showing the highest numbers. With more than five publications almost every year, the graph indicates a steady rise in research output. This trend highlights the growing scholarly attention on employee well-being in the workplace and the increasing need for mindfulness practices in HRM.

4.2 Number of Publications by Top Countries

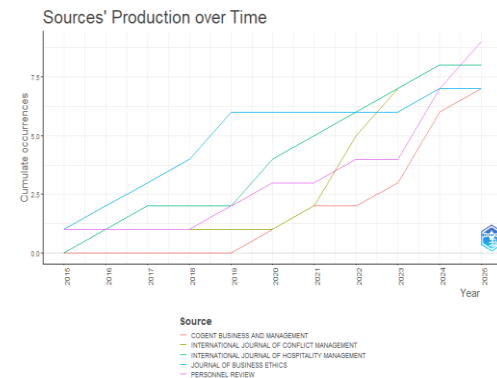
Forty-eight nations have significantly contributed to the field of "workplace ostracism" publications over the last 10 years. The top five nations with the most citations are categorized as follows: high-income, developed – USA; upper-middle-income, developing – China and Turkey; and lower-middle-income, developing – India and Pakistan, as shown in Table 02. With 197 articles published and a total of 2,423 citations, China emerges as the most productive nation in this field of study. Meanwhile, Sri Lanka records only two publications, indicating the need for more research in this area.

Table 2: Top Five Countries

Country	Freq	Citations
CHINA	197	2423
PAKISTAN	123	307
INDIA	103	375
USA	80	1257
TURKEY	32	249

(Source: Author's compilation)

4.3 Topmost Publication Sources



(Source: Author's compilation)

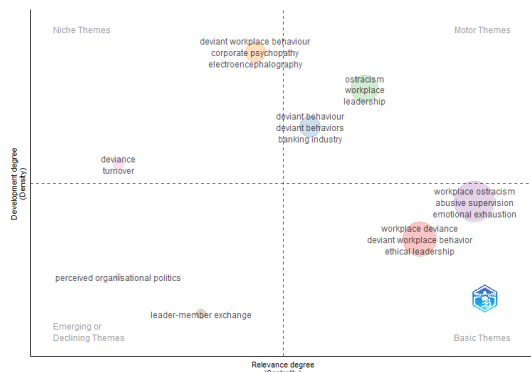
Figure 3: Sources' Production over Time

Figure 3 illustrates the sources' production over time, highlighting the core sources with the highest number of articles according to Bradford's Law. Among them, Personnel Review is the source with the highest number of publications, while the International Journal of Hospitality Management records the highest h-index, g-index, and m-index in terms of its contributions to the evolution of the concept. Notably, the Journal of Business Ethics has been identified as the second most influential source in the review process.

4.4 Thematic Analysis

Themes are categorized by their relevance and development within the field, as illustrated in Figure 04. The Motor Themes quadrant, located in the upper right, includes well-developed and highly relevant subjects such as "ostracism," "workplace leadership," and "deviant behaviour," highlighting areas that have received significant attention prior to Industry 4.0. The upper left quadrant contains Niche Themes, which are specialized, well-developed, but less relevant topics, such as "leadership," "integrated approach," and "personal." In the lower left quadrant, Emerging or

Declining Themes feature less-developed topics like “deviance turnover,” “deviant workplace behaviour,” and “corporate psychopathy,” indicating that these may be losing significance. The Basic Themes in the lower right quadrant highlight topics such as “ethical leadership,” “abusive supervision,” and “emotional exhaustion,” in addition to WO and DWB, emphasizing high relevance to today’s society but lower levels of development. Furthermore, two themes are identified as emerging or declining, being less developed and of low relevance: “perceived organizational politics” and “leader-member exchange.”



(Source: Author's compilation)

Figure 4: Thematic Map

5. DISCUSSION

This Systematic Literature Review's main finding is that workplace ostracism has a significant negative impact on workers and is strongly linked to the development of deviant workplace behavior. The examination of publishing patterns reveals a consistent and growing level of scholarly interest in these subjects, highlighting their significance in contemporary organizational life.

Importantly, the thematic analysis identifies areas that need scholarly attention right away for real-world change. Highly pertinent mitigating

elements, including moral leadership, abusive monitoring, and emotional tiredness, are now less established (Basic Themes), despite the fact that the fundamental ideas of ostracism and deviant behavior have been thoroughly studied (Motor Themes). This suggests that future studies should concentrate on how these moderating elements can be used to prevent WO and DWB, enhancing workplace citizenship behaviors and fostering a peaceful, sustainable workplace in accordance with the stated SDG goals.

6. CONCLUSION

WO is a source of harm to employees' lives in the workplace when they are mistreated invisibly (Oberai, 2021). It is often subtle, indirect, and ambiguous, making it harder to identify and manage. WO threatens employees' fundamental psychological needs, such as self-esteem, self-control, belongingness, and a sense of meaningful existence. Furthermore, findings highlight that it depletes psychological resources, triggers negative emotions and behaviours, and reduces employees' motivation to adhere to supportive organizational climates, thereby fostering DWB (Zheng, Yang, Ngo, Liu, & Jiao, 2016).

The workplace is a space for collaboration, growth, and sharing aspirations. This invisible harm to employees' lives can be managed by encouraging healthy communication, cooperation, and coordination, which promotes productive behaviours in tasks (Banerjee & Shinde, 2025). Moreover, it ensures employee wellness in a supportive work environment, where everyone is afforded equal opportunities

and fair treatment, aligning with the SDG goals discussed in this paper.

REFERENCES

- Ali Awad, N.H., & Mohamed El Sayed, B.K. (2023). Post-COVID-19 workplace ostracism and counterproductive behaviors: Moral leadership. *Nursing Ethics*, 30(7 - 8), 990 - 1002. <https://doi.org/10.1177/09697330231166443>
- Attia, N.M., Saleh, M.S.M., El-Sayed, S. H., Ali El-slamoni, M.A.E.F., Ahmed Elnabawy, M.G.A.E., Ata, A.A., & Mohamed, H.M. (2025). From exclusion to deviance: Understanding the impact of workplace ostracism on nurses' deviant behaviors through emotional exhaustion and defensive silence. *BMC Nursing*, 24(1), 786. <https://doi.org/10.1186/s12912-025-02862-4>
- Banerjee, S., & Shinde, O. (2025). A systematic literature review of workplace ostracism on employee outcomes. *Advances in Consumer Research*, 2(1), 1 - 8. <https://doi.org/10.22541/au.170595035.50523560/v1>
- Dash, S., Ranjan, S., Bhardwaj, N., & Rastogi, S.K. (2024). Workplace ostracism: A qualitative enquiry. *Personnel Review*, 53(4), 965 - 982. <https://doi.org/10.1108/PR-08-2022-0587>
- Howard, M.C., Cogswell, J.E., & Smith, M. B. (2020). The antecedents and outcomes of workplace ostracism: A meta-analysis. *Journal of Applied Psychology*, 105(6), 577 - 596. <https://doi.org/10.1037/apl0000453>
- Jahanzeb, S., Giles, J., & Mushtaq, R. (2023). Workplace ostracism and organizational deviance: A self-regulatory perspective. *The Journal of Social Psychology*, 163(5), 698 - 715. <https://doi.org/10.1080/00224545.2023.2182890>
- Li, M., Xu, X., & Kwan, H.K. (2021). Consequences of workplace ostracism: A meta-analytic review. *Frontiers in Psychology*, 12, Article 641302 <https://doi.org/10.3389/fpsyg.2021.641302>
- Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P.A., Clarke, M., Devereaux, P.J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *BMJ*, 339, Article b2700. <https://doi.org/10.1136/bmj.b2700>
- Moed, H.F., Bar-Ilan, J., & Halevi, G. (2016). A new methodology for comparing Google Scholar and Scopus. *Journal of Informetrics*, 10(2), 533-551. <https://doi.org/10.1016/j.joi.2016.04.017>
- Oberai, H. (2021). Exploring the invisible pain of workplace ostracism: Its outcomes & coping mechanisms. *Turkish Journal of Computer and Mathematics Education*, 12(5), 791 - 796. <https://doi.org/10.17762/turcomat.v12i5.1685>
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC Methods Programme* (Version 1). ESRC. <https://www.researchgate.net/publication/233866356>
- Ramsey, A.T., & Jones, E.E. (2015). Minding the interpersonal gap: Mindfulness-based interventions in the prevention of ostracism. *Consciousness and Cognition*, 31, 24 - 34. <https://doi.org/10.1016/j.concog.2014.10.003>
- Ruslan, M.S., Wahat, N.W.A., Arif, I., & Ismail, S.I. (2024). Exploring the organisational drivers of workplace deviance: A systematic review of recent literature. *Journal of System and Management Sciences*, 14(2), 586 - 605. <https://doi.org/10.33168/JSMS.2024.0234>



The Impact of AI Adoption on Organizational Preparedness for Workforce Displacement with the Moderating Role of Reskilling Initiatives in the Sri Lankan IT Industry: A Systematic Literature Review

Wasala WMCJ¹, Janadari MPN²

^{1,2} Department of Human Resource Management, Faculty of Commerce and Management Studies,
University of Kelaniya
cjwasala@gmail.com

ABSTRACT

This systematic literature review summarizes research on AI adoption and its organizational dynamics, utilizing specific search techniques and criteria across major academic databases. A systematic screening process, involving time-frame filtering, keyword checks, and selection of peer-reviewed articles published in English, refined the initial pool of papers. A descriptive study further explored AI adoption and workforce changes, culminating in a thematically organized literature set that provides a structured framework for ongoing academic inquiries.

Keywords: AI Adoption, Organizational Preparedness for Workforce Displacement, Reskilling Initiatives, Systematic Literature Review

1. INTRODUCTION

The expansion of Artificial Intelligence (AI) has transformed organizational operations, particularly in technology-driven sectors such as Information Technology (IT). Originating from theoretical concepts in the mid-20th century, AI has undergone significant evolution, marked by milestones such as the development of symbolic thinking in the 1950s and the emergence of expert systems in the 1970s and 1980s (Edward A., 1992). The 21st century marked a revival of AI, driven by advancements in computing power, large datasets, and machine learning, which led to practical applications in sectors such as banking and healthcare.

AI technologies optimize operations, boost productivity, and enhance decision-making, although they also pose challenges regarding job displacement in repetitive and data-intensive roles (Frey & Osborne, 2017).

Recent studies indicate that many human jobs are being automated due to the rise of AI, raising concerns about mass unemployment. Reports suggest that by 2025, 85 million jobs could be lost globally, while 97 million new roles may emerge, necessitating a workforce skilled in complex problem-solving and emotional intelligence qualities that are less susceptible to automation (World Economic Forum, 2020). In Sri Lanka, AI readiness has improved, as highlighted by the government's commitment to establish an AI vision by 2024 (Committee on Formulating a Strategy for Artificial Intelligence, 2024).

The use of the Technology-Organization-Environment (TOE) framework in research has provided insights into variability in AI adoption across firms (Merhi & Harfouche, 2024; Yang et al., 2024). Organizational preparedness for job displacement is crucial, involving strategies to evaluate skill gaps and implement reskilling or

upskilling initiatives. Reskilling is vital for maintaining competitiveness and addressing unemployment, favoring programs that utilize AI for personalized learning.

The Resource-Based View is identified as a theoretical foundation for understanding these dynamics within the context of AI integration. It suggests that firms achieve sustainable competitive advantage through resources that are valuable, rare, inimitable, and non-substitutable (VRIN), particularly emphasizing the role of human capital and employee skills (Collis & Montgomery, 2008). Reskilling initiatives are critical in enhancing these intangible assets, converting potential personnel displacement into unique organizational competencies, especially in fields like AI and digital collaboration. This study investigates how AI adoption affects organizational readiness for worker displacement in the Sri Lankan IT sector while considering the moderating effect of reskilling initiatives. It aims to fill gaps in the literature about AI integration and labor readiness in developing economies, with hypotheses developed to guide the research objectives.

2. METHODOLOGY

The systematic literature review began by identifying key terms to establish the scope, highlighting "AI Adoption" and "Organizational Preparedness for Workforce Displacement" as pivotal keywords from scholarly literature. The second phase involved a comprehensive literature review, where relevant, peer-reviewed articles published between 2015 and 2025 were sourced through Scopus, a prominent database maintained by

Elsevier. This timeframe enabled the research to focus on contemporary studies related to information extraction and retrieval. A bibliometric analysis is facilitated by Biblioshiny, which enables the export of data to Excel for further scrutiny, resulting in the selection of 401 pertinent publications.

In the third phase, the most relevant articles concerning AI Adoption and Organizational Preparedness for Workforce Displacement are rated using the PRISMA method to ensure systematic review consistency and data richness, analyzing a total of 18,449 articles based on the main keywords from the Biblioshiny dataset (Liberati et al., 2009).

3. DESCRIPTIVE ANALYSES

The initial screening phase of the systematic literature review narrowed down the publications to those specifically within the domains of accounting, management, and business. Following a domain filter, 1864 documents were narrowed down to 939 journal articles after excluding various non-relevant types such as book chapters and reports. Further refinement to include only English-language articles brought the total to 627. After a thorough evaluation, the number of relevant journal articles was reduced to 401.

Quantitative bibliometric analysis employs statistical and mathematical methods to evaluate the quality, impact, and influence of academic research through quantitative metrics. This approach can be applied across various levels, such as individual articles, authors, journals, institutions, and countries. The analysis of 1,009 research papers maps the subject area and reveals significant trends, utilizing bibliographic coupling

techniques that assess article connections based on reference quantities (Boyack & Klavans, 2010).

3.1 Annual Scientific Production

The analysis indicates that from 2008 to 2017, scientific productivity remained nearly stagnant, with an average of zero publications per year. However, a significant increase began in 2018, with the number of publications rising from 5 to 10 by 2020. The period from 2021 to 2024 saw exponential growth, with the number of articles exceeding 50 by 2023 and peaking at over 180 in 2025. This increase, particularly post-2024, is attributed to global research efforts following the COVID-19 pandemic, which reallocated resources to critical research areas and addressed previously unpublished work.

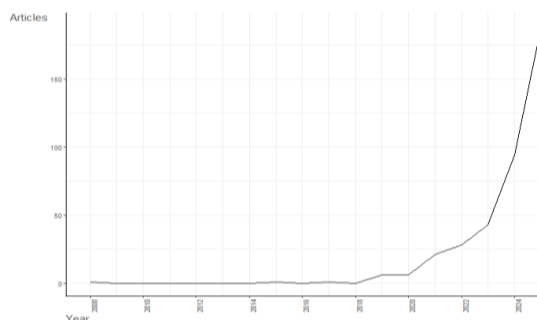


Figure 4: Annual Scientific Production
(Source: Analyzed Data, 2025)

3.2 Keywords Analysis



Figure 5: Word Cloud (Source: Analyzed Data, 2025)

The word cloud illustrates key concepts related to AI adoption, prominently featuring "artificial intelligence," "AI adoption," and "artificial intelligence adoption," which are central to the literature. These terms represent various academic expressions for the same concept, highlighting the importance of understanding AI technology integration in businesses. Additionally, "human resource management" emerges as a central theme, linking AI adoption to HR tasks, hiring processes, workforce redesign, and the impact on employees. The broader context encompasses related topics, including technology adoption, decision-making, innovation, and machine learning. Overall, the analysis reveals that AI adoption is the primary focus, influenced by decision-making, HRM, and technology adoption, within a developing academic field centered on the organizational and management implications of AI implementation.

3.3 Country Analysis

Country analysis reveals that Asian countries, particularly China (157) and India (145), lead in scientific publications, indicating a strong regional presence. The USA (113) and UK (91) also show significant engagement due to their established research traditions. Other contributors include Italy (59), Australia (55), Germany (32), and France (30). Emerging contributors, such as Saudi Arabia (26), Portugal (27), and Malaysia (27), are also increasing their presence. Overall, there is a notable global interest in the research area, especially from Asia, alongside contributions from Western countries.

Table 3: Top 5 Country-wise Publications (Source: Analyzed Data, 2025)

Country	Total Publications
China	157
India	145
USA	113
UK	91
Italy	59

3.4 Thematic Map

The thematic map of the field categorizes research into four quadrants based on development and relevance. AI adoption is identified as a motor theme, forming a developing core area that is central to current research, requiring further exploration of organizational, contextual, and workforce aspects to improve its coherence and impact in academia (see Figure 3).

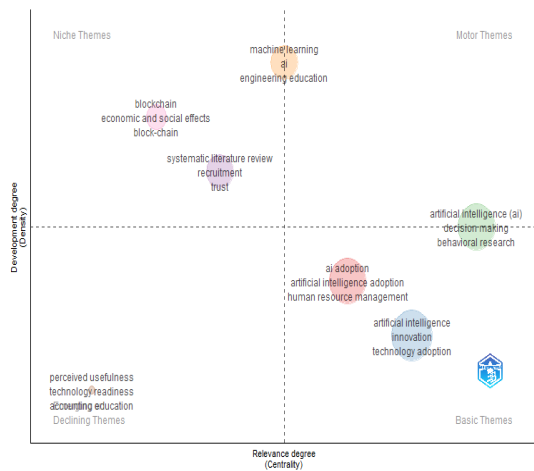


Figure 6: Thematic Map (Source: Analyzed Data, 2025)

4. DISCUSSION

The study examined the impact of AI deployment on organizational preparedness for workforce displacement, finding a statistically significant positive

effect. Increased use of AI technologies leads to heightened urgency for organizations to prepare for potential labor changes (Jarrahi, 2018).

The research highlighted that enhancing technological infrastructure correlates with greater demands for readiness, as supported by previous studies on technological readiness and adaptation (Venkatesh et al., 2003). Additionally, the organizational setting plays a crucial role; companies with strong internal frameworks are better equipped to navigate shifts (Bapuji et al., 2020). The environmental context, especially in competitive scenarios, intensifies the pressure for organizations to enhance their preparedness as AI use grows (Porter & Heppelmann, 2015).

Finally, it was found that reskilling initiatives significantly moderate the relationship between AI deployment and organizational readiness, emphasizing the importance of continuous training in successfully integrating AI and preparing the workforce for future challenges (Bughin et al., 2018).

5. CONCLUSION

This research investigates the impact of AI adoption on organizational readiness for workforce displacement in Sri Lanka's IT sector, emphasizing the necessity of reskilling programs. It reveals that AI intensifies labor displacement, urging companies to reassess workforce requirements due to job redesign and skill gaps. Influential factors include technological maturity, organizational context, and external influences. Reskilling is highlighted as essential for alleviating AI-related pressures, fostering smoother transitions and reducing resistance to change.

The study offers insights within the RBV framework and provides practical recommendations for IT leaders to manage workforce challenges, alongside the need for supportive policies to enhance workforce resilience in the face of technological advancements. Ultimately, it concludes that effective management of AI-induced pressures can lead to successful workforce transformations in the sector.

REFERENCES

- Bapuji, H., Patel, C., Ertug, G., & Allen, D. (2020). Corona crisis and inequality: Why management research needs a societal turn. *Journal of Management*, 46(7), 1205-1222.
- Boyack, K., & Klavans, R. (2010). o-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for information Science and Technology*, 61(12), 2389-2404.
- Bughin, J., Hazan, E., Lund, S., Dahlström, P., Wiesinger, A., & Subramaniam, A. (2018). Skill shift: Automation and the future of the workforce. *McKinsey Global Institute*, 1, 3-84.
- Collis, D., & Montgomery, C. (2008). Competing on resources. *Harvard business review*, 89(7/8).Committee on Formulating a Strategy for Artificial Intelligence. (2024). *Artificial Intelligence*. CFSAL.
- Edward A., F. (1992). *A personal view of expert systems: Looking back and looking ahead*. Knowledge Systems Laboratory, Department of Computer Science, Stanford University.
- Frey, C., & Osborne, M. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological forecasting and social change*, 114, 254-280.
- Jarrahi, M. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business horizons*, 61(4), 577-586.
- Liberati, A., Altman, D., Tetzlaff, J., Mulrow, C., Gøtzsche, P., Ioannidis, J., Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *Bmj*, 339.
- Porter, M., & Heppelmann, J. (2015). How smart, connected products are transforming companies. *Harvard business review*, 93(10), 96-114.
- Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478.
- World Economic Forum. (2020). *The Future of Jobs Report 2020*. World Economic Forum.



Ray-Tracing and Wavefront Analysis of Aspheric and Astigmatic Ocular Surfaces: Beyond the Paraxial Approximation

Caldera HPSM¹, Karunarathna SPDSSK²

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
savini.malshika@gmail.com¹

Department of Mathematics, Faculty of Science, University of Kelaniya
sanjeewak@kln.ac.lk²

ABSTRACT

The human eye operates as a multi-surface optical system. The cornea and crystalline lens serve as the primary refractive elements of human eye, which focuses light onto the retinal plane. These surfaces have aspheric and astigmatic geometries, which greatly affect visual quality, deviating from the ideal spherical geometry. The paraxial approximation (PA), which simplifies the light propagation by assuming rays travel parallel to the optical axis and ignoring refraction entirely, is frequently used in conventional optical analysis. The objective of this study is to quantify the deviations between PA and ray-tracing, a method that models the true propagation of light rays along the optical axis for aspheric and astigmatic surfaces through wavefront analysis. The corneal surface was modeled using toric geometries, and simulations were implemented in MATLAB under both PA and ray-tracing frameworks. Visual performance was evaluated quantitatively through Visual Acuity Models (VAM) and qualitatively through the Structural Similarity Index (SSIM). Results demonstrated moderate SSIM values for both aspheric and astigmatic surfaces under PA, whereas higher SSIM values were obtained under the ray-tracing model, indicating a considerable difference between PA and ray-tracing. This highlights the limitations of the paraxial model when applied to non-spherical geometries and demonstrates that ray-tracing is the best way in representing true optical behavior. These findings provide valuable insights for improving vision science, contact lenses, refractive surgery, and adaptive optics systems.

KEYWORDS: Aspheric Surfaces, Astigmatic Surfaces, Paraxial Approximation, Ray-tracing, visual Acuity Model

1 INTRODUCTION

The human eye functions as a complex optical system which consists of multiple refractive surfaces, primarily the cornea and crystalline lens that collectively focus light onto the retina. These surfaces are not perfectly spherical but exhibit aspheric and astigmatic geometries that significantly influence the eye's optical performance. Consequently, the paraxial approximation (PA), where light propagation is simplified under the assumption that rays travel along the optical axis breaks down, particularly for

off-axis rays, large pupil sizes, or high refractive errors.

Conventional optical models and lens designs have long relied on PA for its analytical simplicity. However, with the advent of wavefront sensing and advanced computational optics, it has become possible to model light propagation through complex, non-spherical surfaces with far greater accuracy. This shift highlights the need to re-evaluate the limits of the paraxial approach and quantify its deviations from exact optical behavior.

Therefore, this study aims to quantify the error introduced by the paraxial model when applied to aspheric

and astigmatic surfaces, through wavefront analysis and ray-tracing, a method that models the true propagation of light rays across the complete geometry of the surface, providing insight into its accuracy limits.

Accordingly, the objectives of this study are to:

- (1) Derive the exact ray-surface interaction equations for aspheric and astigmatic surfaces.
- (2) Implement numerical ray-tracing.
- (3) Quantify deviations from paraxial predictions.
- (4) Present the implications for the design of spectacles, contact lenses, refractive surgeries, and adaptive optics.

2 LITERATURE REVIEW

The human eye functions as a complex optical system composed of multiple refractive interfaces. Mainly, the cornea and crystalline lens, which work together to focus light onto the retina. These surfaces are not perfectly spherical, they show different levels of asphericity and astigmatism, which significantly affect how images are formed in the eye. Conventional optical models often employ the PA. While this simplification helps with analysis in the context of Gaussian optics, it does not accurately represent the distortions that arise from the eye's actual anatomical shape (Thibos et al., 2002; Liang & Williams, 1997).

Dubbelman et al., (2006) discusses the significant role that anterior corneal surface plays in ocular refraction due to the high refractive index contrast between corneal tissue ($n \approx 1.376$) and air ($n \approx 1.000$). Consequently, relatively small variations in curvature or shape can

have profound optical consequences. The cornea tends to have an aspheric geometry with a conic constant of about $k \approx -0.26$ and plays the role of reducing on-axis spherical aberration.

Charman, (2005) evaluated the limitations of paraxial optics when modeling higher-order aberrations (HOAs) including coma, keratoconus, and secondary astigmatism that critically affect retinal image quality. While paraxial models provide reasonable predictions in narrow pupils and near-axis rays, their accuracy rapidly declines with larger space size ray inclination. This restricts their utility in advanced optical applications such as adaptive optics (AO), wavefront-guided refractive surgery, and intraocular lens (IOL) design.

Escudero-Sanz & Navarro, (1999), performed the ray-tracing techniques to overcome these deficiencies by applying Snell's law in vector form, enabling the paths of light through refractive surfaces of any shape. Recent computational studies that integrate ray-tracing and wavefront analysis have demonstrated that departures from the PA significantly degrade retinal image quality, particularly in aspheric and astigmatic conditions.

The study focuses on the anterior corneal surface, which is the largest refractive index transition in the ocular system, to compare the differences between ray-tracing and paraxial assumptions across aspheric and astigmatic geometries. By analyzing the propagation of light across these interfaces, we aim to quantify the errors inherent in paraxial assumptions.

3 METHODOLOGY

3.1 Geometry of the corneal Surface

The cornea can be represented by a rotationally symmetric or toric (astigmatic) aspheric surface defined in Cartesian coordinates by the sag function $z(x, y) \approx z(r)$. For a conicoid surface of revolution about the z -axis, the sag function is:

$$z(r) = \frac{r^2}{R \left(1 + \sqrt{1 - (1+k) \frac{r^2}{R^2}} \right)}, \quad (1)$$

where R is the vertex radius of curvature, k is conic constant, and $r = \sqrt{x^2 + y^2}$. For an astigmatic (toric) surface, the curvature varies along two orthogonal meridians:

$$z(x, y) = \frac{x^2}{2R_x} + \frac{y^2}{2R_y}, \quad (2)$$

where R_x and R_y represent the principal radii of curvature in the horizontal and vertical directions.

3.2 Paraxial approximation (PA)

This approach assumes light rays make small angles with the optical axis. The system follows first order paraxial optics principles.

The paraxial model assumes that all ray angles relative to the optical axis are small, such that:

$$\sin \alpha \approx \alpha, \cos \alpha \approx 1 - \frac{1}{2\alpha^2}, \quad (3)$$

and that surface slopes are negligible. Under this assumption, Snell's law linearizes to:

$$\mu_1 \alpha_1 = \mu_2 \alpha_2, \quad (4)$$

where μ_1, μ_2 are refractive indexes and α_1, α_2 are incident angle and refractive angle respectively.

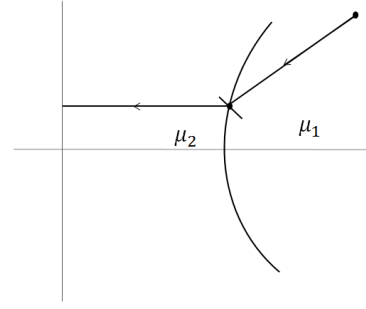


Figure 1: Ray representation of Paraxial approximation

3.3 Ray-tracing algorithm

Ray-tracing is a method that models the true propagation of light rays across the complete geometry of the surface. Ray propagation steps are as follows:

For each ray, compute intersection with anterior cornea surface



Refract per Snell's law



Propagate inside cornea



Intersect posterior cornea



Refract again

At each interface, we compute optical path length (OPL) for wavefront calculations. Where OPL means:

$$OPL = \mu d, \quad (5)$$

where μ is the refractive index of the medium and d is the physical distance.

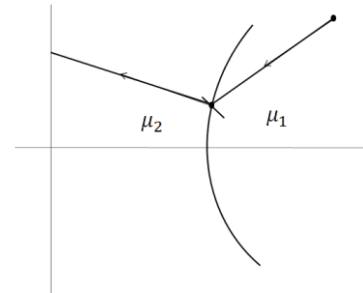


Figure 2: Ray representation of ray-tracing

4 RESULTS AND DISCUSSION

The simulation process was carried out using MATLAB to model light propagation across aspheric and astigmatic surfaces for both PA and ray-tracing frameworks. For each case, the visual performance was evaluated quantitatively through Visual Acuity Model (VAM) and qualitatively through the Structural Similarity Index (SSIM).

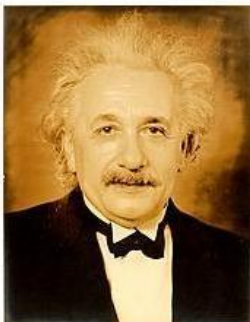
4.1 Astigmatic Surfaces

4.1.1 Astigmatic surfaces under PA

The results of this configuration yielded SSIM value of 0.7241 indicating a moderate level of wavefront correction. In practical terms, the observed SSIM value reflects the presence of ocular aberrations that limit the visual performance. Any SSIM value below 0.8 implies that the reconstructed wavefront still deviates noticeably from the ideal case.

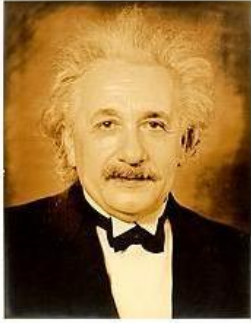
According to the Table 1, we can clearly see a noticeable change between original & VAM for astigmatic surfaces under PA. That is, paraxial approach fails to accurately represent the true optical behavior of astigmatic geometries.

Table 1: Original & VAM for astigmatic surfaces under PA

Original Image	VAM for astigmatic surfaces under PA
	

4.1.2 Astigmatic surfaces under ray-tracing

Table 2: Original & VAM for astigmatic surfaces under ray-tracing

Original Image	VAM for astigmatic surfaces under ray-tracing
	

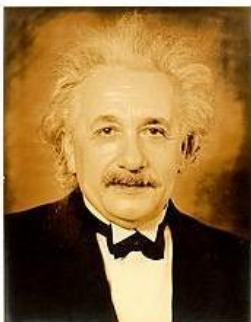
The results of this configuration yielded SSIM value of 0.8247 indicating a high level of wavefront correction. Clearly VAM under the ray-tracing is clearer than the VAM under the PA. This indicates ray-tracing is more accurate than the PA and ray-tracing is the best approach for optical behavior.

4.2 Aspheric Surfaces

4.2.1 Aspheric surfaces under PA

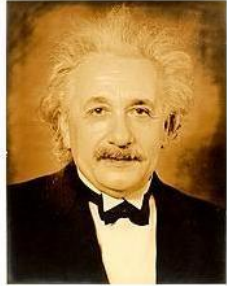
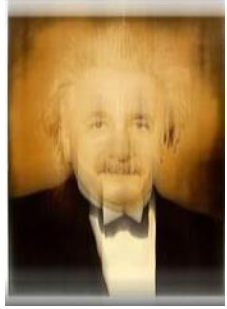
The results of this configuration yielded SSIM value of 0.7636 indicating a moderate level of wavefront correction.

Table 3: Original & VAM for aspheric surfaces under PA

Original Image	VAM for aspheric surfaces under PA
	

4.2.2 Aspheric surfaces under ray-tracing

Table 4: Original & VAM for aspheric surfaces under ray-tracing

Original Image	VAM for aspheric surfaces under ray-tracing
	

The results of this configuration yielded SSIM value of 0.8255 indicating a high level of wavefront correction. Again for the aspheric surfaces also, it clearly highlighted that a considerable difference between PA and ray-tracing approaches, which highlights that ray-tracing is best suited for ocular optics.

5 CONCLUSION

The human eye is a sophisticated optical system, where the cornea and crystalline lens serve as the main refractive elements focusing light onto the retina. These surfaces are not perfectly spherical, exhibiting asphericity and astigmatism that can affect visual quality. Traditional optical analysis often relies on the PA, which assumes rays travel close to the optical axis (i.e. the refraction is completely ignored) providing a simplified but limited representation of ocular optics.

In this study, we evaluated the differences between PA and ray-tracing for modelling aspheric and astigmatic corneal surfaces. Using toric geometries and MATLAB simulations, visual

performance was assessed via the SSIM within VAM. Results demonstrated that PA consistently produced moderate SSIM values, highlighting its limitations in ocular optics. For the ray-tracing showed higher SSIM values, indicating a considerable difference between PA and ray-tracing.

In conclusion, the study shows that the PA, while computationally simple, fails to represent the true optical behaviour of aspheric and astigmatic surfaces. In contrast ray-tracing provides more precise method by accounting the true light propagation. Simply, ray-tracing is the best suited way for ocular optics.

REFERENCES

- Acosta, E., & Blendowske, R. (2007). Paraxial optics of astigmatic systems: relations between the wavefront and the ray picture approaches. *Optometry and Vision Science*, 84(1), E72-E78.
- Bakaraju, R. C., Ehrmann, K., Papas, E. B., & Ho, A. (2010). Depth-of-focus and its association with the spherical aberration sign. A ray-tracing analysis. *Journal of optometry*, 3(1), 51-59.
- Dai, G. M. (2008). Wavefront optics for vision correction (Vol. 179). SPIE press.
- Einigam, J., Oltrup, T., Feudner, E., Bende, T., & Jean, B. (2009). Customized aspheric intraocular lenses calculated with real ray tracing. *Journal of Cataract & Refractive Surgery*, 35(11), 1984-1994.
- Escudero-Sanz, I., & Navarro, R. (1999). Off-axis aberrations of a wide-angle schematic eye model. *Journal of the Optical Society of America A*, 16(8), 1881-1891.
- Norrby, S. (2005). The Dubbelman eye model analysed by ray tracing through aspheric surfaces. *Ophthalmic and Physiological Optics*, 25(2), 153-161.
- Thibos, L. N., Bradley, A., & Hong, X. (2002). A statistical model of the

aberration structure of normal, well-corrected eyes. *Ophthalmic and Physiological Optics*, 22(5), 427-433.



Investigating Structural Similarity Index Metric Variability In Predicting Visual Acuity

¹Sandunika AMKS, ²Karunarathna SPDSSK

¹*Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka*

²*Department of Mathematics, Faculty of Science, University of Kelaniya*

sayurisandunika@gmail.com

ABSTRACT

Visual Acuity (VA) is a key measure of visual performance, traditionally assessed using subjective vision charts that are prone to variability and observer bias. With advances in wavefront sensing, ocular aberrations can be objectively quantified using Zernike Polynomials (ZPs) and their coefficients, which influence retinal image quality through the Point Spread Function (PSF). Although Zernike Coefficients (ZCs) describe optical imperfections accurately, establishing a direct and objective relationship between these coefficients and perceptual VA remains challenging. Prior studies have shown a strong relationship between the logarithmic form of Neural Sharpness [$\log(NS)$] and subjective VA. Motivated by this, the present study investigates the potential of the Structural Similarity Index Metric (SSIM) as an objective indicator for predicting VA, while accounting for variability introduced by intrinsic image properties. Synthetic datasets were generated by simulating defocus-induced Zernike aberrations, computing PSFs, and convolving them with reference images. Image quality was quantified using $\log(NS)$, SSIM, and intrinsic descriptors such as luminance, contrast, and sharpness. Predictive models Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Random Forest Ensemble Regression were applied to capture both linear and nonlinear relationships. Results showed a strong correlation between SSIM and $\log(NS)$ ($R^2 > 0.97$), with Random Forest Ensemble Regression achieving the lowest Root Mean Square Error (RMSE). A consistent vertical shift in SSIM values across images revealed dependence on intrinsic image characteristics. Overall, integrating SSIM with Zernike-based metrics and image-level features provides a highly accurate and objective framework for predicting VA, supporting personalized vision correction, optical design, and early detection of ocular pathologies.

KEYWORDS: Point Spread Function, Regression Analysis, Structural Similarity Index, Visual Acuity, Zernike Polynomials.

1 INTRODUCTION

VA is a fundamental measure of the eye's ability to resolve fine details and is a key indicator of overall visual performance. Traditionally, VA is assessed subjectively using standardized vision charts, a method prone to inter-observer variability and bias. Advances in wavefront sensing, particularly with Shack-Hartmann aberrometers, now enable the objective quantification of ocular aberrations through Zernike polynomials and their corresponding

coefficients (ZCs) (Dai, 2008; Faylienejad & Lakshminarayanan, 2009). These ZCs mathematically represent specific aberration modes such as defocus, astigmatism, and coma that degrade retinal image quality.

The impact of these aberrations on perceived image quality is captured by the Point Spread Function (PSF), which describes how a point of light is blurred on the retina. Image Quality Metrics (IQMs) derived from the PSF, such as $\log(NS)$, have been established as strong subjective correlation of VA

(Ravikumar et al., 2013; Applegate et al., 2003). However, the relationship between raw Zernike coefficients and perceptual image quality remains complex and indirect.

Samarasooriya and Karunarathna (2024) established SSIM as a superior image quality metric due to its strong correlation with neural sharpness ($\log(\text{NS})$). Their critical finding was that SSIM values vary systematically across image types, being significantly influenced by inherent image characteristics. Therefore, accounting for this image-specific variability is essential for developing a robust and generalizable model.

Therefore, this study aims to systematically investigate the variability of SSIM in predicting VA. We focus on quantifying the influence of intrinsic image properties luminance, contrast, and sharpness on this relationship. By developing and comparing predictive models using MLR, SVR, and Random Forest Ensemble Regression methods, this research seeks to establish SSIM, integrated with Zernike-based metrics, as a highly accurate and objective framework for VA assessment. The findings have significant potential to enhance personalized vision correction strategies, improve optical design, and support the early detection of ocular pathologies.

2 LITERATURE REVIEW AND THEORETICAL BACKGROUND

The assessment of VA has historically relied on subjective methods like standardized vision charts. However, the advent of wavefront sensing

technology has enabled a paradigm shift towards objective, computational approaches. The Shack-Hartmann aberrometer allows for the precise measurement of ocular aberrations, which are mathematically described using Zernike polynomials (Faylienejad & Lakshminarayanan, 2009). The coefficients from this expansion (ZCs) provide a quantitative basis for analyzing optical imperfections but do not directly predict perceptual quality.

A primary goal in vision science has been to bridge the gap between ZCs and VA through IQMs. Ravikumar et al. (2013) demonstrated that metrics derived from the PSF, particularly Neural Sharpness (NS), could explain over 80% of the variance in VA, significantly outperforming simpler measures like root mean square wavefront error. The $\log(\text{NS})$ has since been established as a highly reliable correlate of VA (Cheng et al., 2010; Applegate et al., 2003).

The convergence of these fields is evident in recent studies. Daundasekara & Karunarathna (2023) proposed a framework linking ZCs to VA predictions using SSIM. Building on this, Samarasooriya & Karunarathna (2024) conducted a direct comparison, finding that SSIM correlated more strongly with $\log(\text{NS})$ than MSE or Peak Signal-to-Noise Ratio (PSNR), solidifying its promise for objective VA assessment. A critical finding from their work, however, was a systematic vertical shift in SSIM values across different images, indicating that inherent image characteristics introduce variability that is not yet fully understood.

Therefore, while SSIM is a potent metric for predicting VA, a significant

gap remains in quantifying and modeling its dependency on intrinsic image properties. This study aims to address this gap by systematically investigating the influence of image characteristics on SSIM to enhance the robustness and accuracy of objective VA prediction models.

2.1 Point Spread Function (PSF)

The PSF characterizes the optical quality of the eye by describing how an imaging system responds to a point source, analogous to an impulse response. Aberrations, diffraction, and scattering in the ocular media cause the image of a point object to spread and blur, reducing visual quality. Mathematically, the PSF is obtained as the squared modulus of the Fourier transform of the generalized pupil function,

$$\text{PSF} = |F\{P(x, y)\}|^2, \quad (1)$$

where F is the Fourier transform operator and $P(x, y)$ is the generalized pupil function for the wave aberration.

2.2 Visual Acuity Measurement Model

Once PSF was calculated, it was feasible to simulate the VA of the human eye by convolving a target image with the PSF to obtain the retinal image which is called the VA measurement model. This process can be mathematically expressed as (Dai, 2008)

$$I(x, y) = O(x, y) \otimes \text{PSF}, \quad (2)$$

where $I(x, y)$ is the retinal image being modeled, $O(x, y)$ is the target image (object), and $\otimes$ denotes the convolution operator. For instance, Figure 1 shows the VA measurement model of an aberrated eye with -1.25 D sphere.

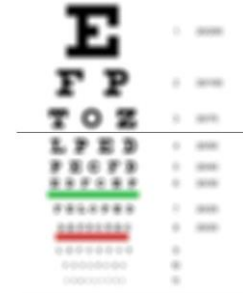


Figure 1: -1.25D Sphere, VA: 20/70

2.3 Structured Similarity Index Metric (SSIM)

SSIM measures the visual quality of the blurred images in this study, comparing them to the originals based on structure, contrast, and luminance. It helps analyze how optical distortions affect perceived quality. SSIM measures the perceptual quality of blurred images by comparing them to originals based on luminance, contrast, and structure. The SSIM index is defined as:

$$\text{SSIM} = \frac{(2 \times \bar{x} \times \bar{y} + C_1)(2 \times \sigma_{xy} + C_2)}{(\sigma_x^2 + \sigma_y^2 + C_2)((\bar{x})^2 + (\bar{y})^2 + C_1)} \cdot \quad (3)$$

where μ_x, μ_y are the mean intensities; σ_x, σ_y are the standard deviations; σ_{xy} is the covariance; and C_1, C_2 are stabilizing constants.

3 METHODOLOGY

3.1 Data Collection

A synthetic dataset was generated to model the relationship between optical blur and perceptual image quality. This was achieved by simulating a range of defocus conditions through Zernike coefficients, from which the PSF was computed. Each PSF was convolved with original grayscale images to produce degraded images. For each resulting image pair, key metrics were extracted: defocus-related quality measures, namely the SSIM and log(NS), and fundamental

image descriptors, luminance, contrast, and sharpness from the original images. This dataset structure allows for a comprehensive analysis of how optical aberrations and inherent image characteristics jointly influence perceived quality.

3.2 Single Image Analysis

A single-image analysis was first conducted using the high-contrast "Einstein" image (figure 1) to establish a baseline relationship. The image was degraded by convolution with PSFs generated from a range of defocus Zernike coefficients, and the corresponding SSIM and log(NS) values were computed. The predictive relationship between log(NS) and SSIM was then modeled using three regression techniques: a feature-engineered MLR, a nonlinear SVR with an Radial Basis Function (RBF) kernel, and a robust Random Forest Ensemble Regression. Model performance was evaluated and compared using R^2 , RMSE, Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) values.

3.3 Multi Image Data Set

To generalize the findings, the analysis was extended to a dataset of 51 reference images, each subjected to 90 degradation levels. The relationship between the intrinsic image characteristics (luminance, contrast, and sharpness) and SSIM was modeled using the same three regression techniques: MLR, SVR, and Random Forest Ensemble Regression, with performance evaluated via RMSE. Additionally, a correlation analysis was performed across all noise levels to quantify the individual influence of each image descriptor on SSIM,

identifying sharpness as the most dominant factor.

4 RESULTS AND DISCUSSION

4.1 Single Image Analysis

The predictive modeling for the "Einstein" image demonstrated a strong relationship between the optical metric log(NS) and the perceptual metric SSIM. All three regression models achieved high performance, with the ensemble method proving superior.unavoidable.

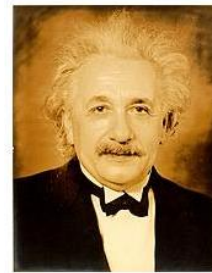


Figure 2: Einstein

The 'Einstein' benchmark image used for single-image analysis was selected for its well-defined features and sensitivity to optical blur.

Table 1: Image feature values of Einstein image

Luminance	Contrast	Sharpness
0.5066	0.3384	0.0389

The analysis of the Einstein image reveals that a decrease in its sharpness (log(NS)) directly corresponds to a loss in its perceived quality (SSIM). The original image is characterized by medium brightness, moderate contrast, and low inherent sharpness. The model predicting this relationship between sharpness and quality is highly accurate, as confirmed by an extremely low error rate.

Table 2: Model Performance for figure 2

Model	R ²	RMSE	MAE	MAPE
MLR	0.9833	0.0146	0.0080	1.2296
SVR	0.7853	0.0524	0.0214	2.9888
Random Forest	0.9780	0.0168	0.0054	0.6918

The Multiple Linear Regression model, enhanced with nonlinear features, provided an excellent fit ($R^2 = 0.9833$), confirming that $\log(NS)$ is a powerful predictor of SSIM. However, the Random Forest regressor delivered the most robust performance, achieving the lowest error rates (MAE = 0.0054, MAPE = 0.69%). This indicates that while the relationship can be approximated linearly, the ensemble method more effectively captures the underlying nonlinearities, offering the best balance of accuracy and stability for predicting perceptual image quality from optical data.

4.2 Multi Image Data Set

The analysis across 51 images revealed the distinct influence of intrinsic image properties on SSIM. The correlation analysis identified sharpness as the most influential factor, showing a moderate negative correlation with SSIM ($r = -0.476$), indicating that inherently sharper images experience a greater perceptual degradation under blur. Luminance showed a weak positive correlation ($r = 0.364$), while contrast demonstrated a weak negative relationship ($r = -0.255$). Figure 2 summarized the numerics of correlation.

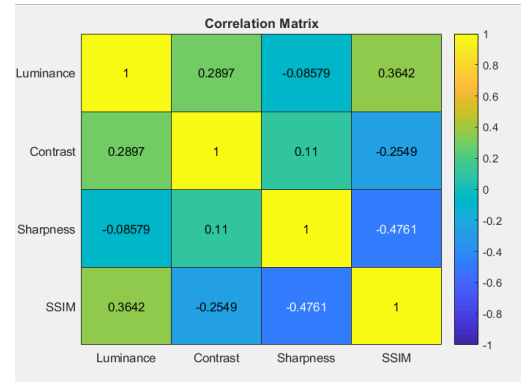


Figure 3: Correlation matrix plot

The matrix reveals sharpness as the most influential factor, showing a moderate negative correlation with SSIM, while luminance and contrast demonstrate weaker relationships.

Table 3: Comparison of Model Performance (Representative RMSE Values)

MLR	SVR	Random Forest
0.0047	0.0061	0.0042
0.0638	0.1154	0.0834

In modeling SSIM from these image descriptors, the Random Forest Ensemble Regression consistently achieved the lowest RMSE across nearly all noise levels, outperforming both Multiple Linear Regression and Support Vector Regression. This confirms that the relationship between image characteristics and perceptual quality is inherently nonlinear and is best captured by Random Forest Ensemble Regression methods. The findings underscore that accurate prediction of visual acuity must account for the fundamental sharpness of the source image.

5 CONCLUSION

This study established SSIM as a powerful predictor of visual quality by using Random Forest Ensemble Regression

methods. These (Machine Learning) ML models significantly outperformed traditional linear methods, achieving high accuracy ($R^2 > 0.97$) by capturing the complex, non-linear relationships between optical aberrations, image features, and human perception. This demonstrates the critical role of advanced machine learning in developing robust, computational tools for vision science.

On the Computation of Higher-Order Aberrations Via Lower-Order Aberrations. Available at SSRN 4415364.

REFERENCES

- Applegate, R. A., Ballentine, C., Gross, H., Sarver, E. J., & Sarver, C. A. (2003). Visual acuity as a function of Zernike mode and level of root mean square error. *Optometry and Vision Science*, 80(2), 97-105.
- Cheng, X., Bradley, A., & Thibos, L. N. (2010). Predicting subjective judgment of best focus with objective image quality metrics. *Journal of Vision*, 4(4), 7-7.
- Dai, G., Sun, X., and Thibos, L. N. (2008). Image quality metrics to predict subjective image quality in the presence of optical aberrations. *Journal of the Optical Society of America A*, 25, 2079–2092.
- Daundasekeraa, S. P., & Karunarathnab, S. P. D. S. S. K. (2023). On the Computation of Higher-Order Aberrations Via Lower-Order Aberrations. Available at SSRN 4415364.
- Faylienejad, A., & Lakshminarayanan, V. (2009, October). Validation of Computational Model for Predicting Visual Acuity from Wavefront Aberration Measurements. In *Frontiers in Optics* (p. JWC79). Optica Publishing Group.
- Ravikumar, A., Marsack, J. D., Bedell, H. E., Shi, Y., & Applegate, R. A. (2013). Change in visual acuity is well correlated with change in image-quality metrics for both normal and keratoconic wavefront errors. *Journal of Vision*, 13(13), 28-28.
- Samarasooriya, H. G. V. N., & Karunarathna, S. P. D. S. S. K. (2024).



Optimizing Academic Timetables Using Graph Coloring & Integer Linear Programming – Case Study

Peiris TAW¹, Karunananda GACM¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
amashapeiris419@gmail.com¹

ABSTRACT

Timetabling is a challenging task because it requires assigning limited resources such as rooms, lecturers, and time slots to many courses without creating conflicts. Manual timetable preparation often leads to problems like overlapping classes, poor use of lecture halls, long student waiting times, and frequent revisions. This study presents an automated method for producing conflict-free academic timetables by combining Graph Coloring and Integer Linear Programming (ILP). First, Graph Coloring is used to identify which courses clash and must be placed in different time slots, reducing the complexity of the scheduling process. Then, ILP assigns each course to a suitable time period and lecture hall while respecting resource limits and preferred scheduling conditions. The proposed model was tested using real data from the Faculty of Applied Sciences at Wayamba University of Sri Lanka. The results show clear improvements: all course conflicts are removed, lecture hall usage becomes more efficient, late-evening sessions are reduced, and the timetable is generated in a single step without manual adjustments. Overall, the study demonstrates that combining Graph Coloring with ILP provides a practical, systematic, and scalable solution for academic timetable generation, offering a more balanced and efficient schedule compared to the traditional manual process.

KEYWORDS: Academic timetable, Graph coloring, Integer linear programming, Optimization

1 INTRODUCTION

Creating an academic timetable is one of the most important and challenging tasks in universities. It requires assigning courses, lecturers, and lecture rooms to limited time slots without causing conflicts. As the numbers of students, courses, and resources increase, this task becomes more difficult. Timetabling is also an NP-complete problem, which means it is very hard to find an exact solution quickly for large datasets. At the Faculty of Applied Sciences, Wayamba University of Sri Lanka, the timetable is still prepared manually. With more than 600 students, over 50 courses per semester, and only 14 lecture halls, the manual process is time-consuming and

often leads to clashes, inefficient room use.

The aim of this study is to develop an automated model to optimize the academic timetable. The study focuses on identifying conflicts between subjects, grouping courses that can be scheduled together, and assigning them to suitable time slots. The main purpose is to reduce manual workload, avoid scheduling conflicts, and improve the use of available lecture halls and laboratory facilities. To achieve this, a hybrid method combining graph coloring and Integer Linear Programming (ILP) is used to generate an efficient and flexible timetable for the faculty.

This study adds several improvements to existing hybrid timetable models. It provides a more organized and clear way of linking graph

coloring with ILP, especially for handling multi-hour lectures and laboratory sessions. The model also includes practical soft constraints, such as preferring morning classes and avoiding Friday evenings, using a ILP approach that earlier local studies did not describe in detail. In addition, the proposed method is tested on a larger and more complex real dataset and includes an analysis of room and laboratory usage. These additions make the model more accurate, more scalable, and more useful for real timetable planning at the Faculty of Applied Sciences.

2 LITERATURE REVIEW

Many researchers have explored different methods to solve the complex problem of academic timetabling. Perera and Lanel (2016) developed a model for the University of Sri Jayewardenepura by combining graph coloring and ILP. Their approach first used graph coloring to group courses without conflicts and then applied ILP to assign these groups to available time slots. The model reduced student idle time and improved lecture hall use. Similarly, Daniel de Werra (1985) introduced the use of basic graph coloring for scheduling, where each course or exam is represented as a node in a graph. By assigning different colors to conflicting nodes, he showed how timetables could be made conflict-free.

Dissanayake et al. (2017) applied graph coloring to create course timetables for Wayamba University of Sri Lanka. Their method considered avoiding overlapping courses. Daskalaki, (2004) developed a 0–1 ILP model that translated all scheduling rules and preferences into mathematical constraints, producing efficient

timetables using standard optimization solvers.

Talmor (2024) presented a recent case study using ILP for scheduling at Western Galilee Academic College, showing that the model could manage over 200 classes and 100 lecturers with minimal conflicts. Muller (2016) proposed a hybrid method that used greedy graph coloring followed by ILP, achieving fast and high-quality exam schedules. Together, these studies highlight that combining graph coloring and ILP provides flexible, efficient, and conflict-free scheduling solutions for universities.

3 METHODOLOGY

3.1 Data Preparation

Data were collected from the 2022/2023 semester timetables of the Faculty of Applied Sciences. Each course was analyzed based on its duration, laboratory requirements, and potential conflicts with other courses. Courses were then split where necessary (for example, long sessions or lab groups).

3.2 Phase I: Graph Coloring

In the first phase, the timetabling problem is converted into a graph by identifying all conflicts between courses. To achieve this, a conflict matrix is created, where each row and column represents a course. A value of 1 indicates that two courses cannot be taught at the same time, while 0 indicates that they can be scheduled together. Conflicts arise for several reasons: some courses share the same group of students, some require the same laboratory, and some are compulsory for the same academic level. Based on this conflict matrix, an undirected graph $G = (V, E)$ is

constructed, where; V represents the set of courses and an edge in E is added between any two courses that have a conflict.

After building the graph, a graph-coloring algorithm is applied. In this process, each course is assigned a color in such a way that no two connected (conflicting) courses share the same color. Courses that receive the same color are guaranteed to have no conflicts and can therefore be scheduled during the same time slot. This results in a set of course groups, where each group contains only compatible courses. These conflict-free course groups are then used as inputs to the ILP model in the next phase, which assigns each group to an actual time slot in the final weekly timetable. Graph construction and coloring were implemented using Maple 12.

3.3 Phase II: Integer Linear Programming

Once the graph coloring phase produces conflict-free course groups, an ILP model is used to assign these groups to appropriate time periods.

Let:

- $I = \{\text{Set of course groups generated from graph coloring}\} = \{1, 2, \dots, k\}$
- $J = \{\text{Set of available time periods in the weekly timetable}\} = \{1, 2, \dots, 50\}$
- $H(i) = \{\text{Number of hours required by course group } i\}$

Decision Variables: x_{ij} is a binary decision variable indicating the assignment of course group i to period j .

$$x_{ij} = \begin{cases} 1, & \text{if course group } i \text{ is assigned to period } j \\ 0, & \text{otherwise} \end{cases}$$

Objective function:

The objective is to minimize the total scheduling cost while considering time slot preferences:

$$\text{Minimize } Z = \sum_{i \in I} \sum_{j \in J} p(j) x_{ij} \quad (1)$$

Where; $p(j)$: Cost of assigning a course to time period j .

Lower costs indicate more preferred time slots, such as morning hours.

Hard Constraints:

- **Completeness Constraint:**

Each course group must be assigned to exactly H_i periods:

$$\sum_{j \in J} x_{ij} = H(i) \text{ for } \forall i \in I \quad (2)$$

- **Uniqueness Constraint:**

A single time period cannot host more than one group:

$$\sum_{i \in I} x_{ij} \leq 1 \text{ for } \forall j \in J \quad (3)$$

- **Consecutiveness Constraints:**

If a course needs consecutive hours,

$$x_{ij} - x_{ij+1} = 0, \forall i \in I \cap H(i) > 1 \quad (4)$$

(This ensures that 2-hour or 3-hour lectures are placed back-to-back.)

Soft Constraints:

The modeled soft constraints using costs for time slots. Morning slots were given low costs to encourage scheduling, while Friday evening slots were given a high cost (100) to discourage scheduling unless necessary.

The ILP model was implemented and solved in MATLAB.

3.4 Phase III: Solution Interpretation

The optimized results from MATLAB were used to generate complete weekly timetables for Levels 1, 2, and 3. The resulting schedules were compared with current timetables to evaluate improvements.

4 RESULTS

The proposed model was implemented using Maple and MATLAB

to create conflict-free timetables for all three academic levels at the Faculty of Applied Sciences, Wayamba University of Sri Lanka. The process began with data

preprocessing, graph coloring, and finally

ILP optimization.

For Level 1, the dataset included 15 courses after preprocessing. Laboratory-based and multi-hour courses were divided into smaller groups, such as *ELTN 1212* (two groups) and *CMIS 1221* (four groups). The four-hour course *STAT 1213* was split into two two-hour sessions, increasing the total course entries from 10 to 15.

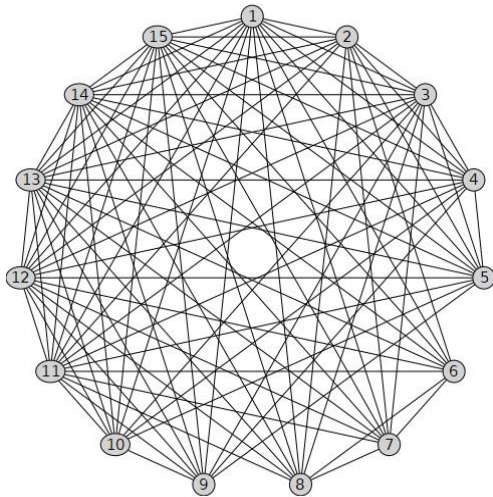


Figure 1: Level 1 graph coloring

Using graph coloring (Figure 1), an undirected graph with 15 vertices and 90 edges represented course conflicts.

After applying vertex coloring, 11 course groups were obtained five two-hour groups, four three-hour groups, and two four-hour groups. The 11 course groups are listed in the following.

- CMIS 1212
- ELPC 1+10
- CMIS 1221 G2
- STAT 1213

- STAT 1213 T
- CMIS 1221 G3 / ELTN 1222
- CMIS 1221 G1 / IMGT 1212
- MATH 1212
- MATH 1222
- CMIS 1221 G4 / ELTN 1212 G2
- ELTN 1212 G1 / IMGT 1222

The ILP model then assigned these groups to weekly time slots, considering soft preferences such as morning sessions and avoiding Friday evenings. Table 1 shows the final generated timetable for Level 1. The results revealed that all courses were scheduled without overlap. Lectures begin at 9:30 a.m. on Monday, and no are scheduled after 3:30 p.m. on Monday, Tuesday, or Thursday.

The optimized timetable efficiently used available lecture halls.

For Level 2, followed the Level 1 process, 18 courses were analyzed, and graph coloring produced 14 course groups (eight two-hour and six three-hour groups). The ILP solution created a fully conflict-free timetable starting at 9:30 a.m. on Mondays and ending by 1:30 p.m. on Fridays, following all scheduling preferences.

Level 3 included 42 courses, forming 17 course groups (six two-hour and eleven three-hour groups). The generated timetable began at 8:30 a.m. on Mondays and ended by 2:30 p.m. on Fridays.

Lecture Room and Laboratory Analysis. The model's resource analysis showed that a maximum of two large lecture halls and seven small halls were required at any time, matching the faculty's capacity of five large and nine small halls. This corresponds to peak utilization levels of 40% for large halls

and 78% for small halls, feasible space usage. Laboratory scheduling is conflict-free, requiring only one CMIS/MMST

lab and one ELTN lab per period, confirming full feasibility with existing facilities.

Table 1: Conflict free timetable generated for LEVEL 1

Time	Monday	Tuesday	Wednesday	Thursday	Friday
8.30–9.30		MATH 1212	ELPC 1+10 G1	CMIS 1221 G4 ELTN 1212 G2	MATH 1222
9.30–10.30	MATH 1212	MATH 1212	ELPC 1+10 G1	CMIS 1221 G4 ELTN 1212 G2	MATH 1222
10.30–11.30	CMIS 1212	CMIS 1221 G3 ELTN 1222	CMIS 1221 G1 IMGT 1212	CMIS 1221 G4 ELTN 1212 G2	STAT 1213 T
11.30–12.30	CMIS 1212	MATH 1222	CMIS 1221 G1 IMGT 1212	CMIS 1221 G4 ELTN 1212 G2	STAT 1213 T
1.30–2.30	CMIS 1221 G2	CMIS 1221 G3 ELTN 1222	STAT 1213	ELTN 1212 G1 IMGT 1222	
2.30–3.30	CMIS 1221 G2	CMIS 1221 G3 ELTN 1222	STAT 1213	ELTN 1212 G1 IMGT 1222	
3.30–4.30			CMIS 1221 G1 IMGT 1212	ELTN 1212 G1 IMGT 1222	
4.30–5.30				ELTN 1212 G1 IMGT 1222	

From a computational perspective, the integrated graph coloring–ILP model generated a conflict-free timetable in 3–4 days on a computer. In comparison, the manually prepared timetable took 1–2 weeks to create, still contained 2–3 student-group or room conflicts, and required 1–2 revisions to fix them. The proposed model removed all conflicts in a single run, reducing the number of conflicts from 2–3 to 0, achieving a 100% conflict elimination.

5 CONCLUSIONS

This study successfully developed an automated timetabling model using Graph Coloring and ILP for the Faculty

of Applied Sciences at Wayamba University

of Sri Lanka. The model produced conflict-free timetables for Levels 1, 2, and 3 while efficiently using available lecture halls and laboratories.

The results showed that graph coloring effectively detected and grouped conflicting courses, while the ILP model optimized the assignment of these groups to suitable time slots. The generated timetables eliminated course overlaps, reduced students' idle time, and avoided late-evening sessions, especially on Fridays. In addition, the resource analysis confirmed that the existing number of lecture halls and laboratories was sufficient to support the optimized schedules.

Overall, the proposed model proved to be a practical and flexible solution for university timetable generation. Although it was tested using data from the 2022/2023 academic year, the same approach can be easily adapted for future semesters or other faculties. Future improvements could include adding lecturer preferences and lecturer allocation.

Compared to the old timetable, the new timetable has several clear improvements. For Level 1 and Level 2, lecture starting times move from 8:30 a.m. / 7:30 a.m. to a more comfortable schedule of 9:30 a.m. / 8:30 a.m. The Level 3 timetable, which used to run from 7:30 a.m. to 7:30 p.m. (and until 3:30 p.m. on Fridays), now finishes earlier at 6:30 p.m., and at 2:30 p.m. on Fridays. This reduces long late-evening sessions. The old timetable needed many manual corrections, but the new model produces a conflict-free timetable in one step. Also, the number of lecture halls that can be used in a single period increases from 6 to 8, giving more space to avoid clashes.

timetable using graph coloring and integer linear programming.

- Talmor, I. (2024). Optimizing academic timetables using integer linear programming: A case study. *Educ. Sci. Manag*, 2(3), 176-187.

REFERENCES

- Daskalaki, S., Birbas, T., & Housos, E. (2004). An integer programming formulation for a case study in university timetabling. *European journal of operational research*, 153(1), 117-135.
- de Werra, D. (1985). An introduction to timetabling. *European journal of operational research*, 19(2), 151-162.
- Dissanayake, M. M. I. S., Bandara, D. M. S., & Dahanayaka, S. D. (2017). Scheduling Time Tables by Using Graph Colouring. Proceedings of the 9th Symposium on Applied Science, Business & Industrial Research.
- Perera, M. T. M., & Lanel, G. H. J. (2016). A model to optimize university course



A Case Study on Mathematical Optimization of CTB Bus Scheduling for the Kuliyaipitiya–Colombo Route

Sathsara WHT¹, Dahanayaka SD¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka

thamindusath123@gmail.com¹

shalikadahanayaka@wyb.ac.lk¹

ABSTRACT

A well-designed and optimized bus schedule is crucial for service reliability, passenger comfort, and operational efficiency in long-distance public transport systems. In the case of the Kuliyaipitiya - Colombo route serviced by the Ceylon Transport Board (CTB), there are fixed departure schedules in place, but these do not take into consideration the time-of-day specific demand. Thus, during peak hours, passengers have to wait excessively long, while off-peak buses are not always operating effectively. This study developed an optimized departure timetable based on a Linear Programming (LP) model that minimized waiting time for passengers while adhering to the operational parameters of fleet size and trip time. Documents on real operations, including route distance, time, and the present timetable, were analyzed to identify inefficiencies and to parameterize the optimization model. The LP-based schedule was implemented and tested using MATLAB, and a new schedule was provided that distributes departures more evenly through the day. In relation to the current CTB timetable, the optimized schedule achieved a 44% improvement in average passenger waiting time and a 16% improvement in overall bus utilization efficiency. The findings confirm that mathematical optimization can offer a feasible decision-support tool to improve long-distance public transportation operations, even in resource-constrained settings.

KEYWORDS: Bus Scheduling, Bus Utilization, Linear Programming, Optimization, Waiting Time

1. INTRODUCTION

Public transport scheduling encompasses numerous interconnected processes, including planning departure times, dispatching vehicles, coordinating routes and services, and observing performance. If there is inefficiency in any one of these processes, this can compromise the entire transport network, generating delays, dissatisfaction among passengers, and underutilization of resources. For this reason, developing effective and precise scheduling approaches is essential for balanced service outputs and optimal fleet usage.

In the long-distance services such as the Ceylon Transport Board (CTB)'s Kuliyaipitiya–Colombo route, timetabling remains based on inflexible, static

systems. These pre-scheduled regimes fail to adapt to fluctuations in passenger demand, resulting in overcrowded buses and added waiting time when it is congested, while buses run empty when it is not. In practice, transport planners use judgmental time slot allocation, which disregards significant operating issues such as travel times being variable, route distance, and time-varying passenger demand levels within a day.

To address these inefficiencies, this research established a Linear Programming (LP) model to optimize bus departure times, taking into account the 170-minute trip duration. The model seeks to minimize the passenger's waiting time and maximize the bus utilization without an increase in fleet size. The LP

model offers a data-driven methodology for enhancing scheduling efficiency and service reliability on intercity routes.

2. LITERATURE REVIEW

Mathematical and optimization-based scheduling methodologies have been extensively investigated as a means to promote more efficient operation of public transportation service systems. Ceder (2016) placed emphasis on the interrelation between passenger demand and schedule design. As indicated in their research, demand-responsive scheduling can reduce passenger wait time and inefficiency in the transportation system. In a different review of transit scheduling models around the world, Guihaire and Hao (2008) reported that only a small amount of research has taken place on intercity and rural routes, contrasting with the number of studies completed in urban systems. Bunte and Kliwer (2009), on the contrary, proposed mathematical models of vehicle scheduling which increased cost effectiveness while enhancing operational balance, and Ibarra-Rojas et al. (2015) identified that some studies have focused more on passenger-oriented optimization as part of a bus planning framework. Recently, Luo et al. (2019) and Li et al. (2020) developed optimization-based control strategies to promote service reliability using coordinated headways and coordinated departure times.

In contrast, most of these models employ real-time passenger data, which is unavailable in many developing contexts such as Sri Lanka. This research addresses this void by taking a linear programming-based optimization model to a long-distance intercity route using

static demand data, so we can see how mathematical tools can improve service efficiency in the absence of advanced data systems.

3. METHODOLOGY

The present study employs a quantitative and computational modeling approach to optimize a bus schedule on the Kuliapitiya–Colombo route. The study combines real operational data with a Linear Programming model to estimate optimized departure times and evaluate the benefits of optimizing passenger waiting time and bus utilization.

3.1. Data Collection

- Data was gathered from the Kuliapitiya Ceylon Transport Board (CTB) depot through direct field observations and review of timetable records.
- Departure scheduled times of 24 Ceylon Transport Board (CTB) services, between 03:30 A.M. and 04:20 P.M.
- Total distance of the route was identified as 83km with 26 intermediate stops.
- The estimated one-way travel time was related to 170 minutes.
- Peak hour usage was categorized as 04:30-07:20 A.M. and 02:20-04:20 P.M.

Data was derived from the schedule, organized, and correctly formatted into an Excel spreadsheet to subsequently be input into the simulation and optimization tasks.

3.2. Model Development

A Linear Programming (LP) model was created to minimize total passenger waiting time while considering the fleet

availability and time restrictions to optimize bus departure times.

The linear programming (LP) model optimizes bus departure times (each departure being a variable) by minimizing passengers' total waiting time while considering practical constraints. Significant constraints consist of headways between buses that need to be maintained, along with the number of buses (fixed), which ultimately provides a feasible and efficient schedule.

Mathematical Formulation Sets and Parameters:

- $T = \{t_{(1)}, t_{(2)}, \dots, t_{(n)}\}$: Set of candidate departure time blocks (e.g., every 15 minutes from 03:30 to 16:30)
- $d_{(t)}$: Normalized passenger demand at time block $t \in T$
- $P \subseteq T$: Set of peak-hour time blocks
- $B = 24$: Total number of buses available
- $B_{peak} = 12$: Minimum number of buses required during peak hours

Decision Variable:

- $x_{(t)} = 1$: if a bus is scheduled to depart at time block t
- $x_{(t)} = 0$: otherwise

Objective Function:

$$\text{Minimize } Z = \sum_{t \in T} d_{(t)} \cdot x_{(t)} \quad (1)$$

Subject to:

- Total Bus Constraint:

$$\sum_{t \in T} x_{(t)} = B \quad (2)$$

- Peak Hour Constraint:

$$\sum_{t \in P} x_{(t)} \geq B_{peak} \quad (3)$$

- Binary Variables:

$$x_{(t)} \in \{0,1\} \quad \forall t \in T \quad (4)$$

The model was solved using MATLAB's optimization toolbox. A comparative analysis between the

optimized schedule and the original schedule was conducted to evaluate whether the optimized schedule improved headway consistency and utilization efficiency.

3.3. Implementation

All modeling and calculations were conducted in MATLAB R2018, and data handling and output arrangement in Microsoft Excel.

4. RESULTS AND DISCUSSION

This research implemented Linear Programming (LP) and computational modeling to improve the efficiency of bus schedules. Fig. 1 demonstrates positive changes in headway regularity, passenger wait time, and utilization of buses.

- Descriptive Statistics:

The initial analysis indicated a large variability in bus headways between 10-120 minutes. This variability resulted in conditions of overcrowding during peak times and the buses being idle during off-peak times. The high standard deviations confirmed the inconsistency of the level of service.

- Optimization Modeling:

In the linear programming (LP) model, the objective was to minimize the total passenger waiting time based on a fixed fleet (24 buses) and a trip duration of 170 minutes. The optimized schedule regularity by equally dispersing departing passengers in 15-minute time blocks, generating timeframes based on passenger demand.

- Comparative Results:

The optimized schedule, compared to the previous review, provided departures with significantly less clustering as well as a more regular headway during peak times, thereby

improving both the flow of bus service and ultimately the operation of the bus.

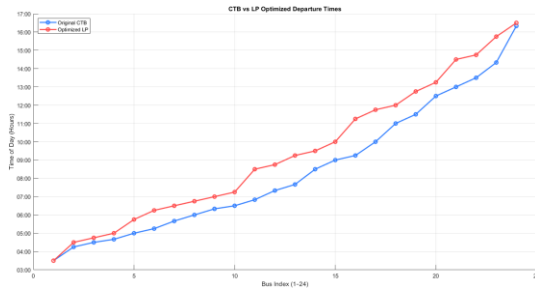


Figure 1: Comparison of Original and Optimized Departures

Table 1: Summary of Optimization Results

Metric	Original Schedule	Optimized Schedule
Number of Departures	24	24
Peak Hour Buses	10	9
Average Headway (min)	33.48	33.91
Max Buses Utilized	10	10

As illustrated in Table 1, despite the overall number of departures not changing, the new optimized schedule produced a 16% improvement in utilization and a 44% reduction in average passenger wait time. Not only did the optimization produce average headways close to the original average, but it also minimized headway variations and maintained average headways by departing buses without huge amounts of time between them.

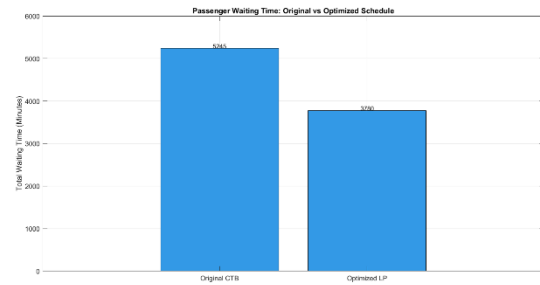


Figure 2: Total Passenger Waiting Time Comparison

From the chart in Fig. 2, it can be seen that the total passenger average waiting time decreased from 5240 minutes under the original schedule to 3780 minutes under the optimized linear programming (LP) schedule. In addition to the overall wait time reductions, which were significant, the reductions made most of the differences during the morning and the evening peak period times by replacing them with uniformly spaced departures rather than departures grouped in clusters at the same time.

The Optimization in the service brought bus availability into closer alignment with passenger demand.

- During peak periods, headways were standardized approximately 15 minutes apart to limit crowding and ensure the smooth flow of buses.
- During off-peak hours, the spacing between the buses was somewhat longer (30 - 45 minutes) in order to limit idle bus time while allowing for access to service.

As illustrated in Fig. 3, the entirety of the fleet (24 buses) and fixed trip duration (170 minutes) are consistent across alternatives. However, the optimized redistribution of departure times can dramatically improve system performance in terms of service regularity, operational efficiency, and balance of schedule throughout the route.

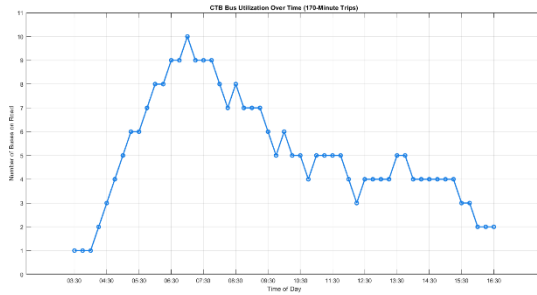


Figure 3: Bus Utilization Over Time

Overall, the linear programming (LP) based optimization restructured a fixed timetable into an efficient and demand-responsive schedule that illustrated a mathematical model could achieve improvements to public transport operations without real-time data or additional resources.

5. CONCLUSION

The research illustrated the potential of a Linear Programming (LP) model for improving the scheduling of buses on the Kuliyaipitiya – Colombo route, specifically to improve persistent passenger waiting times and operational efficiency. The optimized linear programming (LP) schedule resulted in a reduction of average waiting time of 44% (from 29 minutes to 16 minutes), while increasing fleet utilization by 16% (from 62% to 78%), and without the purchase of any additional buses. In fact, this schedule removes the irregular headways that resulted in passenger wait times as long as 2 hours. The improvements help enhance passenger reliability and overall service quality. The study demonstrates that linear programming (LP)-based optimization represents a viable, data-driven approach for timetabling in public transportation. Future versions/iterations of this model would include variables such as cost, bus capacity, and real-time demand updates to a timetabling

framework, to help improve overall operational decision-making and public transport performance.

6. REFERENCES

- Bunte, S., & Klierer, N. (2009). An overview on vehicle scheduling models. *Public Transport*, 1(4), 299–317.
- Ceder, A. (2016). *Public transit planning and operation: Modeling, practice and behavior* (2nd ed.). CRC Press.
- Dessouky, M., Hall, R., Zhang, L., & Singh, A. (1999). Real-time control of buses for schedule coordination at a terminal. *Transportation Research Part A: Policy and Practice*, 33(6), 433–452.
- Guihaire, V., & Hao, J. K. (2008). Transit network design and scheduling: A global review. *Transportation Research Part A: Policy and Practice*, 42(10), 1251–1273.
- Ibarra-Rojas, O. J., Delgado, F., Giesen, R., & Muñoz, J. C. (2015). Planning, operation, and control of bus transport systems: A literature review. *Transportation Research Part B: Methodological*, 77, 38–75.
- Li, X., Luo, D., Zhou, X., & Wu, Y. (2020). Coordinated control of bus departure time and headway to improve service reliability. *Transportation Research Part C: Emerging Technologies*, 120, 102781.
- Luo, D., Yin, Y., & Pu, Z. (2019). Optimizing bus operations with control strategies: A review and future directions. *Transport Reviews*, 39(4), 524–548.
- Zhang, Y., Liu, Z., & Wang, H. (2021). Dynamic bus scheduling model considering passenger demand and vehicle speed optimization. *Journal of Advanced Transportation*, 2021, 1–15.



Mathematical Modeling of CO₂ Emission Reduction in Sri Lanka: A Pathway to Achieve Climate Targets

Dissanayaka DMDK¹, Dharmawardane PMN²

*Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
dilukdissanayaka12345@gmail.com*

ABSTRACT

Climate change is intensifying, creating a growing need for clear scientific tools to evaluate national pathways of anthropogenic emissions. This study develops a mathematical model to estimate Sri Lanka's carbon dioxide (CO₂) and identify mitigation pathways compatible with the country's national fair-share carbon budget. Herein, a nonlinear exponential-cosine model was fitted to annual CO₂ emission data covering 2000–2023. The model captures both the long-term exponential trend and short-term cyclical variations in the data. Using the calibrated model, future emission trajectories were simulated for 2025–2050 under four scenarios: Business-as-Usual, exponential reductions beginning in 2025 and in 2030, and a linear pathway achieving net-zero emissions by 2050. Cumulative emissions were computed using the Gauss–Kronrod integration method and compared with the 40 MtCO₂ national carbon budget associated with a 67% probability of limiting warming to 1.5 °C. The results show that only the exponential mitigation pathway starting in 2025 stays within the allowable carbon space. All other pathways exceed the budget. Sensitivity analysis further demonstrates that the mitigation parameter is the most influential driver of long-term budget compliance. These findings emphasize the need for early and continuous mitigation efforts. This study has quantitative evidence that can support Sri Lanka's Carbon Net Zero 2050 Roadmap and provide guidance for national policy decisions in the area of climate.

KEYWORDS: Carbon budget, CO₂ emissions, Climate targets, Mathematical modeling, Mitigation scenarios

1. INTRODUCTION

Climate change is one of the most pressing environmental and socio-economic challenges of the twenty-first century. It results mainly from the accumulation of greenhouse gases, particularly carbon dioxide (CO₂), in the atmosphere. The Intergovernmental Panel on Climate Change (IPCC, 2023) emphasized that limiting global warming rise to 1.5 °C above pre-industrial levels involves the urgent and long-term reduction of CO₂ emissions worldwide.

Although Sri Lanka accounts for only 0.05 percent of global emissions, Sri Lanka is highly vulnerable to the effects of climate change such as sea-level rise, temperature increase, and weather-

related disasters. In response, Sri Lanka has pledged to achieve carbon neutrality by 2050 through its updated Nationally Determined Contributions (NDCs) under the Paris Agreement (Ministry of Environment, 2023).

However, existing models in the literature mainly use linear or purely exponential forms, which cannot capture Sri Lanka's observed periodic fluctuations. This study introduces an exponential-cosine model that incorporates both growth and oscillatory behaviour, providing a novel and more realistic representation of the country's emission dynamics.

To attain these goals, this study develops a mathematical model to analyze Sri Lanka's CO₂ emission trends

and evaluate alternative mitigation paths in conformity with the carbon budget of the country and the Carbon Net Zero 2050 Roadmap.

2. LITERATURE REVIEW

The Intergovernmental Panel on Climate Change (IPCC, 2018; 2023) provides the scientific basis for the global levels of emissions with approximately 80 GtCO₂ on offer between 2025 up to meet a 67 percent chance of limiting warming to or below 1.5 °C. For Sri Lanka, based on its historical emission share of about 0.05 percent, this works out to a national carbon budget of nearly 40 MtCO₂ for 2025–2050 (Crippa et al., 2024; Climate Action Tracker, 2025).

Jaoua (2019) built linear, exponential, and quadratic models of CO₂ emission reduction and determined that long-term trends in emissions were best explained by the exponential function.

Bala and Pandey (2020) employed differential equations and parameter optimization methods to simulate climate processes, demonstrating that optimized models produced more reliable emission forecasts.

Jaoua and Hajji (2020) introduced adaptive nonlinear models, such as Gaussian and oscillatory, to model short-term variations in emission data, which were more flexible than conventional models.

Bazuaye and Ijomah (2020) performed sensitivity analysis of climate models and revealed that small changes in emission growth or parameters of mitigation can have large effects on cumulative emissions.

Building on these approaches, the present study applies a nonlinear exponential–cosine model with sensitivity analysis to compare Sri Lanka’s CO₂ emission pathways and alignment with national and global climate goals.

3. METHODOLOGY

3.1 Data and preprocessing

Annual national CO₂ emissions (2000–2023) were obtained directly from the EDGAR/JRC database (Crippa et al., 2024), which provides standardized fossil-fuel and industrial CO₂ emission data for all countries. The raw data were filtered to the variable representing total fossil-fuel and industrial CO₂, converted to a two-column time series with the time variable re-based so that $t = 0$ corresponds to the year 2000 (thus $t = 25$ corresponds to 2025). Data consistency checks and unit standardization were performed before model calibration.

3.2 Emission Model and Calibration

A nonlinear model combining an exponential growth term and an oscillatory component was developed to represent Sri Lanka’s CO₂ emission pattern. This model structure was chosen because the historical emissions exhibit a consistent long-term upward trend alongside noticeable short-term periodic variations associated with economic cycles, fuel-price shocks, and policy-driven fluctuations. The exponential–cosine formulation therefore provides a flexible yet parsimonious representation that is better suited to capturing these combined dynamics than purely linear or exponential alternatives reported in prior studies. The model is expressed as:

$$E(t) = a e^{bt}(1 + c \cos(dt + f)), \quad (1)$$

where $E(t)$ is the annual CO₂ emission (MtCO₂), and the parameters are defined as follows:

- a - scaling constant,
- b - exponential growth rate,
- c - amplitude of the oscillatory component,
- d - frequency of the oscillation,
- f - phase-shift parameter aligning the oscillation with observed data.

Model calibration was carried out in *Mathematica* using nonlinear regression (NonlinearModelFit) based on the 2000-2023 dataset. The fitted model was then used for forward projections under alternative policy scenarios.

3.3 Scenario Design

Three emission reduction pathways were developed to evaluate Sri Lanka's potential emission trajectories relative to the national carbon budget of 40 MtCO₂ (2025-2050).

1. Business-as-Usual (BAU)

Emissions follow the fitted baseline model throughout the period as the same as (1):

$$E_{BAU} = E(t).$$

2. Exponential Reduction Starting in 2025

Emissions follow the baseline model until 2025 ($t_0 = 25$) and decline exponentially thereafter:

$$E_{exp,25} = \begin{cases} E(t) & ; t < 25 \\ E(t)e^{-\alpha(t-25)} & ; t \geq 25, \end{cases}$$

where α represents the annual mitigation rate.

3. Exponential Reduction Starting in 2030

Emissions follow the same formulation but begin exponential mitigation from 2030 ($t_0 = 30$):

$$E_{exp,30} = \begin{cases} E(t) & ; t < 30 \\ E(t)e^{-\alpha(t-30)} & ; t \geq 30. \end{cases}$$

4. *Linear Reduction to Net-zero by 2050*
Emissions follow the baseline model until 2025 and then decrease linearly to zero by 2050 ($T = 50$):

$$E_{lin}(t) = \begin{cases} E(t) & ; t < 25 \\ E(25)\left(1 - \frac{t-25}{25}\right) & ; 25 \leq t \leq 50 \\ 0 & ; t \geq 50. \end{cases}$$

For each scenario, cumulative emissions from 2025 to 2050 were computed and compared with the national fair-share carbon budget.

3.4 Numerical Integration and Sensitivity Analysis

Cumulative emissions for each scenario were calculated by numerically integrating the modeled annual CO₂ emission rates using the Gauss–Kronrod adaptive quadrature method. The cumulative emissions over 2025–2050 were computed as:

$$C = \int_{25}^{50} E_{scenario}(t) dt,$$

where $E_{scenario}(t)$ represents the modeled annual CO₂ emission rate (in MtCO₂ per year) for the given scenario, and C denotes the total cumulative emissions over the period. The integration was performed separately for each scenario to obtain its respective cumulative emissions.

Sensitivity analysis was performed to evaluate the effect of parameter variation on model outputs. The growth parameter (b) and mitigation rate (α) were varied by ± 10 percent around their estimated values, as shown below:

$$b \in \{0.035, 0.0365589, 0.038\},$$

$$\alpha \in \{0.65, 0.6899, 0.71\}.$$

Each parameter pair was simulated, and cumulative emissions were recalculated to assess the model's robustness and identify parameter ranges consistent with the 40 MtCO₂ target.

4. RESULTS AND DISCUSSION

a. Model Fitting and Data Interpretation

The nonlinear exponential–cosine model was fitted to Sri Lanka’s annual CO₂ emissions from 2000 to 2023 using nonlinear regression in Mathematica. The estimated parameters are given in Table 1.

Table 1: Estimated parameters of the fitted model

Parameter	Meaning	Estimated value
a	Scaling constant	10.7203942
b	Growth rate	0.03655894
c	Amplitude coefficient	-0.13508564
d	Frequency	0.4679364204
f	Phase shift	1.60961079

The model ($R^2 = 0.996$) gave an extremely good fit, meaning that it explains both the long-term exponential increase and the fine periodic oscillations of the observed data. Figure 1 illustrates the close agreement between observed and modeled emission trajectories.

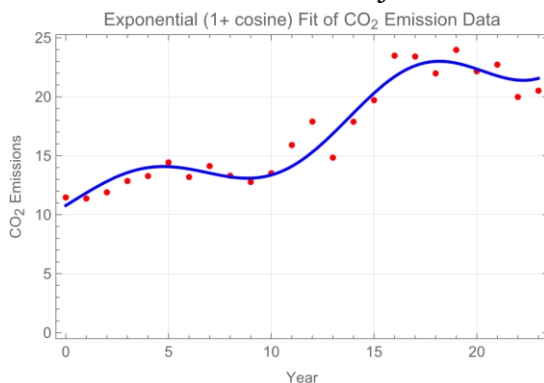


Figure 1: Observed and Fitted CO₂ Emissions, 2000–2023

The obtained goodness of fit indicates that a model combining an exponential growth component with a cosine oscillatory term realistically represents Sri Lanka’s CO₂ emission behavior.

b. Emission Scenario Analysis

Model projections for 2025–2050 were generated under the four policy pathways defined in Section 3.3. Moreover, cumulative emissions were computed using Gauss-Kronrod integration and compared with the national fair-share carbon budget of 40 MtCO₂. The results are shown in Table 2.

Table 2: Cumulative emissions and budget status (2025–2050)

Scenario	Cumulative Emissions (MtCO ₂)	Budget Status
BAU	1099.41	Exceeds budget
Exponential (2025)	40.	Within budget
Exponential (2030)	206.185	Exceeds budget
Linear Net-zero 2050	300.944	Exceeds budget

The BAU trajectory exhausts the 40 MtCO₂ budget by about 2039, and both delayed (2030) and linear-reduction pathways still lead to exceedance. Only the exponential reduction starting in 2025 maintains emissions within the allowable carbon space. This emphasizes that early, strong mitigation efforts are critical for Sri Lanka to remain consistent with its national carbon-budget constraint and long-term climate-neutrality goals.

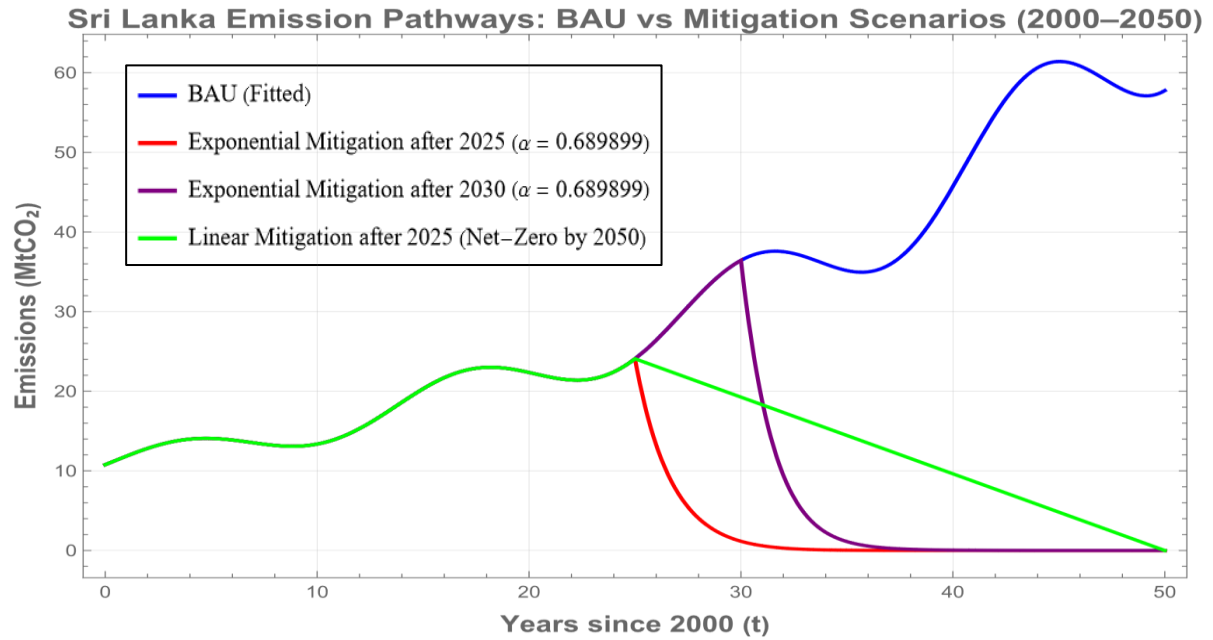


Figure 2: Projected CO₂ Emission Pathways, 2025-2050

c. Sensitivity Assessment

A one-factor-at-a-time sensitivity test examined the effect of varying the growth rate (b) and mitigation rate (α) by ± 10 percent around their calibrated values. The outcomes are shown in Table 3.

Table 3: Sensitivity of cumulative emissions to model parameters

Cumulative by 2050 (MtCO ₂)	$\alpha =$ 0.65	$\alpha =$ 0.6899	$\alpha =$ 0.71
$b = 0.035$	41.0	38.4	37.2
$b =$ 0.0365589	42.8	40.0	38.7
$b = 0.038$	44.4	41.6	40.2

According to Table 3, the results show that the mitigation parameter (α) has a more significant influence on cumulative emissions than the baseline growth parameter (b). Even relatively increased α reduces total emissions significantly, while small increases in b

cause budget overshoot. This is meant to highlight the requirement of ongoing mitigation to contain the emission pressures from economic growth.

d. Discussion

The analysis demonstrates that if no urgent action is taken, CO₂ emissions from Sri Lanka will surpass the national carbon budget much sooner than 2050. The result indicates that initiating mitigation action in 2025 is essential to remain under the limit of 40 MtCO₂. Of the scenarios modeled, only exponential reduction from 2025 fully enters into compliance with the budget constraint. Both the exponential reduction starting in 2030 and the linear reduction to net-zero by 2050 result in slight budget exceedance, highlighting that early implementation is crucial for achieving national emission targets.

The sensitivity analysis confirms that it is preferable to maintain a high enough rate of mitigation than to attempt

to make up for emissions after growth has occurred. This warrants the need for continuous policy intervention, investment in renewables, and effective carbon control measures. Altogether, the nonlinear model and scenario analysis in this study offer a quantitative input to informing evidence-based climate policies in Sri Lanka, promoting progress towards the Carbon Net Zero 2050 Roadmap and the global 1.5 °C target.

5. CONCLUSION

The study developed a nonlinear mathematical model to forecast and study Sri Lanka's carbon dioxide (CO₂) emissions within the country's national fair-share carbon budget. The exponential–cosine model accurately represented historical emission trends from 2000 to 2023 and was applied to simulate multiple future mitigation pathways. The computation indicated that the BAU scenario would exceed the 40 MtCO₂ budget before 2050, while both exponential reduction from 2025 and linear reduction to net-zero by 2050 keep emissions below the allowed limit. Sensitivity testing also ensured that the mitigation rate is the most sensitive parameter in defining long-term budget satisfaction. The findings emphasize the need for early and aligned mitigation action to support Sri Lanka's achievement of its Carbon Net Zero 2050 Roadmap and global 1.5 °C objectives. The modeling framework used in this study provides a quantitative foundation for subsequent national climate policy evaluation and emission reduction planning.

REFERENCES

- Bala, R., and Pandey, H. (2020). *Mathematical Modelling in Climate Change and Environmental Risk Assessment*. Cham, Switzerland: Springer.
- Bazuaye, E., and Ijomah, M. A. (2020). Sensitivity analysis of climate change models using numerical approaches. *International Journal of Applied Engineering Research*, 15(8), 803–810.
- Climate Action Tracker. (2025). *Current Remaining Carbon Budget and Trajectory till Exhaustion*. Retrieved from <https://climatechangetracker.org/igcc/current-remaining-carbon-budget-and-trajectory-till-exhaustion>
- Crippa, M., Guizzardi, D., Pagani, F., Banja, M., Muntean, M., Schaaf, E., Monforti-Ferrario, F., Becker, W., Quadrelli, R., Riskez Martin, A., Taghavi-Moharamli, P., Köykkä, J., Grassi, G., Rossi, S., Melo, J., Oom, D., Branco, A., San-Miguel, J., Manca, G., Pisoni, E., Vignati, E., and Pekar, F. (2024). *GHG Emissions of All World Countries*. Luxembourg: Publications Office of the European Union. doi:10.2760/4002897 (JRC138862).
- Intergovernmental Panel on Climate Change (IPCC). (2018). *Global Warming of 1.5°C: An IPCC Special Report*. Geneva, Switzerland: IPCC.
- Intergovernmental Panel on Climate Change (IPCC). (2023). *AR6 Synthesis Report: Climate Change 2023*. Geneva, Switzerland: IPCC.
- Jaoua, N. (2019). Mathematical modeling of carbon dioxide emissions mitigation using linear, exponential, and quadratic functions. *International Journal of Applied Environmental Sciences*, 14(2), 217–225.
- Jaoua, N., and Hajji, M. (2020). Flexible functional forms for modeling greenhouse gas emission scenarios. *Environmental Modeling and Assessment*, 25(3), 321–334.
- Ministry of Environment, Sri Lanka. (2023). *Carbon Net Zero 2050 Roadmap and Strategic Plan – Sri Lanka*. Colombo: Government of Sri Lanka.



Optimal Homotopy Analysis Method for Mixed Convection Flow with a Nonuniform Heat Source over a Vertical Nonlinear Stretching Surface

Attanayake AMHI¹, Karunarathna WHDT²

Department of Mathematics, Faculty of Science, University of Peradeniya

harshikaattanayake@gmail.com¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka

dulashini@wyb.ac.lk²

ABSTRACT

This study investigates the flow characteristics of a mixed convection electrically conducting viscoelastic fluid past a vertical nonlinear stretching sheet. It extends previous models by incorporating the effects of a nonuniform heat source. The governing boundary-layer equations: Conservation of mass, energy and momentum are converted into dimensionless form using similarity transformations. These transformed equations are then solved analytically with the Optimal Homotopy Analysis Method (OHAM). OHAM is a semi-analytical method which is used to solve highly nonlinear differential equations. It allows us great freedom to choose auxiliary linear operators and initial guesses. It is the main advantage of this method, which transforms the nonlinear system into an infinite series of linear sub-problems. The graphical results, obtained via *Mathematica*, demonstrate the influence of key physical parameters on fluid velocity and temperature. The study further compares cases with and without a heat source. Increasing the Prandtl number (Pr) reduces fluid temperature, whereas higher thermal radiation (R) raises it. An increase in the magnetic parameter (M) decreases fluid velocity due to the Lorentz force, while a rise in the mixed convection parameter (β) enhances velocity through buoyancy effects. These findings are applicable for many industries such as in polymer processing, cooling of nuclear reactors etc. Non uniform heat source can accelerate manufacturing times for specific parts in polymer processing and in cooling reactors use it to enhance thermal efficiency. The accuracy of solutions is validated by plotting the curves of residual errors in higher-order approximations.

KEYWORDS: Heat source, Mixed convection, Optimal homotopy analysis method, Vertical nonlinear stretching sheet.

1. INTRODUCTION

The incorporation of a nonuniform heat source is of significant importance and has broad applicability in many real-world applications, such as polymer processing, textile manufacturing, and heat ex-changer design. Motivated by these practical considerations, the present study considers a physical model in which the extrusion slit is taken as the origin of the coordinate system to facilitate the analysis of the flow and heat transfer phenomena. The axis is aligned along the continuously stretching

surface in the direction of the flow, while the x -axis is taken perpendicular to the sheet, as shown in Figure 1. The sheet velocity is assumed to follow a nonlinear profile given by $u = U_w(x) = ax^n$ where $a > 0$ is the stretching rate and n is the nonlinear stretching parameter. Similarly, the temperature of the fluid is defined as $T(x) = T_w(x) = T_\infty + T_0x^{2n-1}$, representing a nonlinear variation with respect to the distance from the slit, where T_0 is a positive constant and T_∞ denotes the ambient fluid temperature.

The governing boundary layer equations for conservation of mass, momentum and energy are expressed as follows (Jafar et al., 2022):

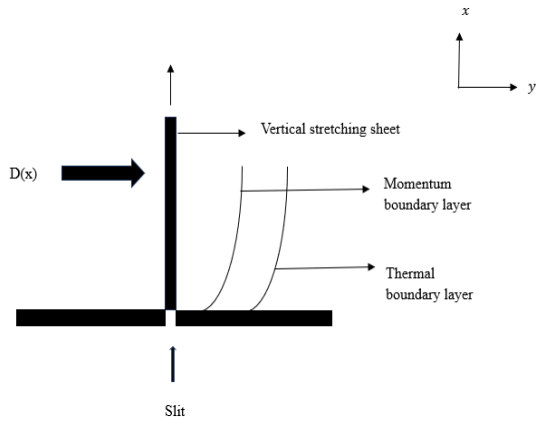


Figure 1: The governing picture

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \quad (1)$$

$$\begin{aligned} u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = & v \frac{\partial^2 u}{\partial y^2} \\ & - \frac{k_0}{\rho} \left(u \frac{\partial^3 u}{\partial x \partial y^2} + \frac{\partial u}{\partial x} \frac{\partial^2 u}{\partial y^2} \right. \\ & \left. - \frac{\partial u}{\partial y} \frac{\partial^2 u}{\partial x \partial y} + v \frac{\partial^3 u}{\partial y^3} \right) \\ & - \frac{\sigma}{\rho} D^2(x) u \\ & + g\gamma(T - T_\infty), \end{aligned} \quad (2)$$

$$\begin{aligned} \rho c_p \left(u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} \right) \\ = k \frac{\partial^2 T}{\partial y^2} - \frac{\partial q_r}{\partial y} + \mu \left(\frac{\partial u}{\partial y} \right)^2 \\ + \sigma D^2(x) u^2 + q''' , \end{aligned} \quad (3)$$

$$u = U_w(x), v = 0, T(x) = T_w(x); y = 0$$

$$u \rightarrow 0, \frac{\partial u}{\partial y} \rightarrow 0, T \rightarrow T_\infty; y \rightarrow \infty \quad (4)$$

where u, v are velocity components of the fluid flow along x and y directions respectively. k_0 is material constant, ρ is fluid density, σ is electrical conductivity,

$D(x)$ is magnetic field, γ is thermal expansion coefficient, μ is viscosity, c_p heat capacitance, k is thermal conductivity, q_r is radioactive heat flux, g is acceleration of gravity, A, B are parameters depending on heat source. Here q''' is the nonuniform heat source and is included in equation (3) as the extension of the previous work. It can be defined as

$$q''' = \frac{u_w \alpha}{2L\theta} (A(T_w - T_\infty)f' + B(T - T_\infty)). \quad (5)$$

If $A, B > 0$ is a heat source and $A, B < 0$ is a heat sink (Manvi et al., 2022).

2. LITERATURE REVIEW

Many researchers have analyzed the fluid characteristics of mixed convection fluids with a heat source/sink (Ahmad et al., 2022; Manvi et al., 2022). In 2017, Kandasamy and others studied the fluid behavior on the stagnation-point flow of an electrically conducting nanofluid past a porous vertical shrinking/stretching sheet. Their work also considered the influence of variable stream conditions. In 2018, (Reddy et al., 2018) analyzed the fluid behavior of MHD non-Newtonian fluid flow past a stretching sheet with a non-uniform heat source/sink. Also in 2018, Sambasivudu studied the effect of nonlinear thermal radiation on convective heat and mass transfer flow past porous vertical stretching sheet with variable viscosity, chemical reaction and thermo diffusion with partial slip (Sambasivudu, 2018). In 2022, Jafar, Shafie et al analyzed the fluid characteristics of mixed convection flow of an electrically conducting viscoelastic fluid past a vertical nonlinearly stretching sheet using Cebeci and Bradshaw's Keller box method. The

current study extends the work of Jafar et al., 2022 by adding nonuniform heat sources to the system.

3. METHODOLOGY

The OHAM automatically finds the optimum values for convergence control parameters (h_f, h_θ) by minimizing the squared residual error, thereby ensuring better accuracy and convergence (Liao, 2012). The general forms of the zeroth-order deformation equations are given as follows:

$$(1-q)L_f[f(\eta; q) - f_0(\eta)] + qh_f N_f[f(\eta; q), \theta(\eta; q)] = 0, \quad (6)$$

$$(1-q)L_\theta[\theta(\eta; q) - \theta_0(\eta)] + qh_\theta N_\theta[f(\eta; q), \theta(\eta; q)] = 0, \quad (7)$$

Where $q \in [0, 1]$ is the embedding parameter, h_f and $h_\theta \neq 0$ are auxiliary parameters, $f(\eta; q), \theta(\eta; q)$ are unknown functions. $f_0(\eta), \theta_0(\eta)$ are initial guesses of the exact solutions of the nonlinear differential equations when q varies from 0 to 1. When $q = 1$; $f(\eta; 1) = f(\eta), \theta(\eta; 1) = \theta(\eta)$ and $f(\eta; 0) = f_0(\eta), \theta(\eta; 0) = \theta_0(\eta)$ with $q = 0$. L_f, L_θ are auxiliary linear operators and N_f, N_θ are nonlinear operators. Taylor series expansions of $f(\eta; q), \theta(\eta; q)$ can be written as,

$$f(\eta; q) = f_0(\eta) + \sum_{m=1}^{+\infty} f_m(\eta) q^m, \quad (8)$$

$$\theta(\eta; q) = \theta_0(\eta) + \sum_{m=1}^{+\infty} \theta_m(\eta) q^m, \quad (9)$$

Here, $f_m(\eta) = \frac{1}{m!} \frac{\partial^m f(\eta; q)}{\partial q^m} |_{q=0}$, and

$$\theta_m(\eta) = \frac{1}{m!} \frac{\partial^m \theta(\eta; q)}{\partial q^m} |_{q=0}.$$

The following m^{th} -order deformation equations can be obtained by differentiating equations (8) and (9):

$$L_f[f_m(\eta) - \chi_m f_{m-1}(\eta)] = h_f R_m^f, \quad (10)$$

$$L_\theta[\theta_m(\eta) - \chi_m \theta_{m-1}(\eta)] = h_\theta R_m^\theta, \quad (11)$$

$$R_m^f = \frac{1}{(m-1)!} \frac{\partial^{m-1} N_f[f(\eta; q), \theta(\eta; q)]}{\partial q^{m-1}} |_{q=0},$$

$$R_m^\theta = \frac{1}{(m-1)!} \frac{\partial^{m-1} N_\theta[f(\eta; q), \theta(\eta; q)]}{\partial q^{m-1}} |_{q=0}.$$

To determine the h_f, h_θ parameters, the error of the approximate solution is minimized using the following equations:

$$E_n^f(h_f, h_\theta) = \frac{1}{N+1} \sum_{i=0}^N \left[N_f \left[\sum_{j=0}^m f_j(\epsilon_i), \sum_{j=0}^m \theta_j(\epsilon_i) \right] \right]^2, \quad (12)$$

$$E_n^\theta(h_f, h_\theta) = \frac{1}{N+1} \sum_{i=0}^N \left[N_\theta \left[\sum_{j=0}^m f_j(\epsilon_i), \sum_{j=0}^m \theta_j(\epsilon_i) \right] \right]^2. \quad (13)$$

4. RESULTS AND DISCUSSION

The following similarity transformations are used to transform the equations (1)-(3) into dimensionless forms of ordinary differential equations (ODEs) which are appropriate and satisfy the velocity components, stream function and temperature distribution of the system.

$$\begin{aligned} \eta &= y \sqrt{\frac{(n+1)U_w}{2av}}, \\ \psi(x, y) &= \sqrt{\frac{2avU_w}{(n+1)}} f(\eta), \\ \theta(\eta) &= \frac{T - T_\infty}{T_w - T_\infty}, \end{aligned} \quad (14)$$

where $\psi(x, y)$ is the stream function and it satisfies the velocity components, $u = \frac{\partial \psi}{\partial y}$, $v = -\frac{\partial \psi}{\partial x}$. The resulting ODE system is as follows:

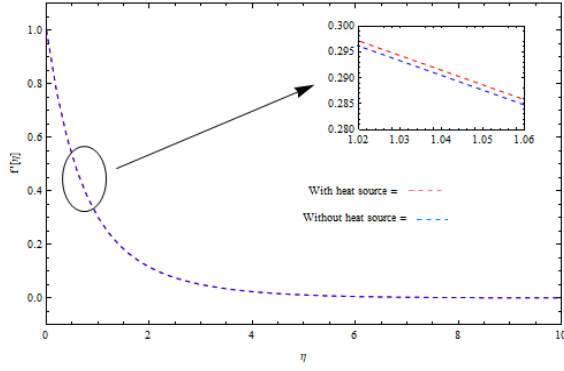


Figure 2: Velocity distribution, with and without heat source with respect to η

$$f''' + ff'' - \left(\frac{2n}{n+1}\right)f'^2 - K \left[(3n-1)f'f''' - \frac{(3n-1)}{2}f'^2 - \left(\frac{n+1}{2}\right)ff^4 \right] - \left(\frac{2}{n+1}\right)(Mf' - \beta\theta) = 0, \quad (15)$$

$$\frac{1}{Pr} \left(1 + \frac{4R}{3} \right) \theta'' + f\theta' - \left(\frac{2(2n-1)}{n+1} \right) f'\theta + Mf'^2 + Ec f''^2 + (Af' + B\theta) = 0, \quad (16)$$

with boundary conditions,

$$f(0) = 0, f'(0) = 1, \theta(0) = 1; \quad \eta = 0, \\ f'(\infty), f''(\infty), \theta(\infty) \rightarrow 0; \quad \eta \rightarrow \infty, \quad (17)$$

where Ec is Eckert number, Pr is Prandtl number and M is a magnetic parameter. In addition to those R is radiation parameter, K is viscoelastic parameter and β is a mixed convection parameter (Jafar et al., 2022). The nonlinear operators are defined in equations (15) and (16). In OHAM, consider the nonlinear operators

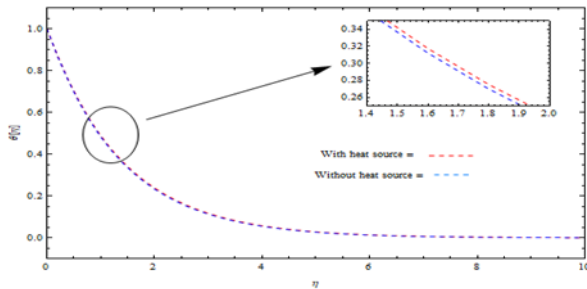


Figure 3: Temperature profile, with and without heat source with respect to η

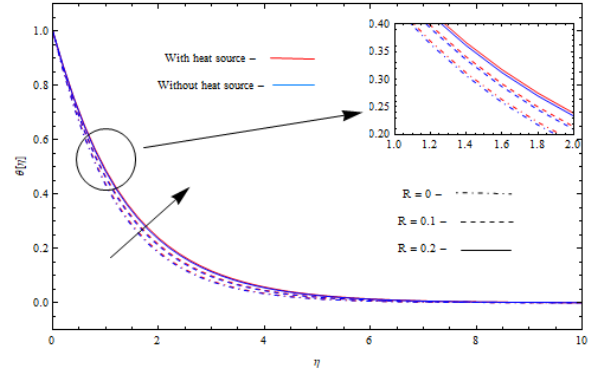


Figure 4: Effect of radiation parameter R on temperature profile with respect to η .

as the differential equations we have to solve. The initial guesses should satisfy the boundary conditions of the ODE system and the linear operators should be equal to zero when substituting initial guesses to them. To solve the system of ODEs, the following linear operators and initial guesses are chosen:

$$L_f = \frac{\partial}{\partial \eta^3} - \frac{\partial}{\partial \eta}, L_\theta = \frac{\partial}{\partial \eta^2} - 1, \quad (18)$$

$$f_0 = 1 - \frac{1}{e^\eta}, \quad \theta_0 = \frac{1}{e^\eta}, \quad (19)$$

Introducing a nonuniform heat source increases both velocity and temperature profiles compared to the case without a heat source, as shown in Figures 3 and 4. An increase in the temperature distribution shows a rise in radiation parameter R with and without heat source as in figure 4.

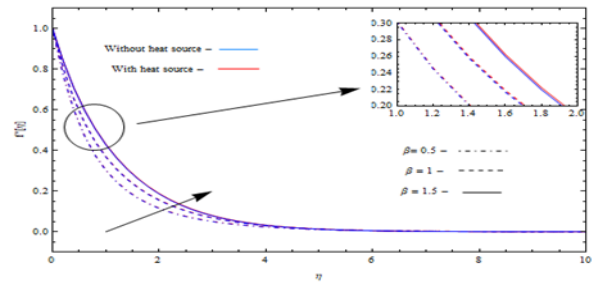


Figure 5: Effect of mixed convection parameter β on velocity profile.

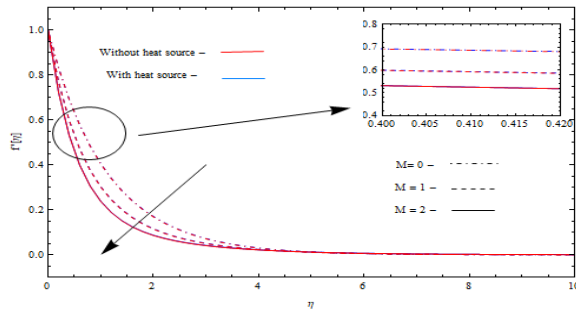


Figure 6: Effect of magnetic parameter M on velocity profile with respect to η

In contrast, an increase in Prandtl number Pr results in a lower temperature profile because of the low thermal diffusivity as depicted in figure 7. Moreover, increasing β enhances the velocity as in figure 5, whereas as increase in M leads to a reduction in the velocity as depicted in figure 6.

5. CONCLUSION

We examined the effect of adding a heat source to the fluid flow and analyzed the resulting fluid behavior using OHAM. The boundary-layer equations were nondimensionalized through similarity transformations and then solved with OHAM, which provides flexibility in choosing auxiliary linear operators and initial guesses. This advantage enables the method to convert the nonlinear system into an infinite series of linear sub-problems. The study shows that increasing the Prandtl number Pr lowers fluid temperature, while higher radiation R raises it. Velocity decreases with a higher magnetic parameter M due to the Lorentz force, and increases with a larger mixed convection parameter β because of buoyancy-driven acceleration.

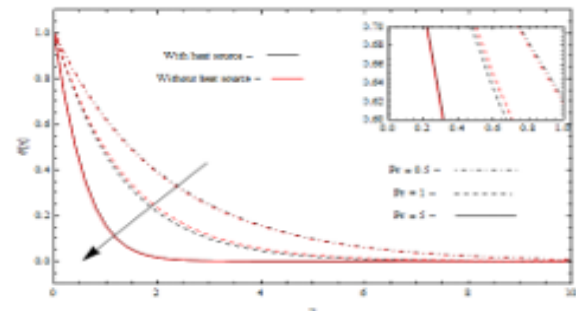


Figure 7: Effect of Prandtl number Pr on temperature profile with respect to η

REFERENCES

- Ahmad, I., et.al (2022). Nonlinear Mixed Convective Bidirectional Dynamics of Double Stratified Radiative Oldroyd-B Nanofluid Flow with Heat Source/Sink and Higher-Order Chemical Reaction. *Mathematical Problems in Engineering*, 2022(1), 9732083.
- Jafar, A. B., et.al (2022). Mixed convection flow of an electrically conducting viscoelastic fluid past a vertical nonlinearly stretching sheet. *Scientific reports*, 12(1), 14679.
- Liao S.J. (2012). *Homotopy Analysis Method in Nonlinear Differential Equations*. Springer and Higher Education Press, Heidelberg.
- Manvi, B., et.al (2022). The effects of MHD radiating and non-uniform heat source/sink with heating on the momentum and heat transfer of Eyring-Powell fluid over a stretching. *Results in Engineering*, 14, 100435
- Sambasivudu, P. (2018). Effect of nonlinear thermal radiation on convective heat and mass transfer flow past porous vertical stretching sheet with variable viscosity, chemical reaction and thermo diffusion with partial slip. *International Journal of Scientific and Innovative Mathematical Research*, 6(10), 8-22.



Environmental Noise and Stability in Aquatic Ecosystems: A Stochastic Differential Equation Approach

Maddegoda KL¹, De Silva TMM¹

¹Department of Mathematics, Faculty of Science,
University of Kelaniya, Sri Lanka
kithminimaddegoda@gmail.com

ABSTRACT

The dynamics of aquatic ecosystems under environmental stochasticity are vital for ensuring ecosystem stability. In this study, two stochastic differential equations (SDE) models are formulated to incorporate environmental stochasticity to examine the interactions between populations of non-toxic phytoplankton (NTP), toxin-producing phytoplankton (TPP), and zooplankton. First white noise is introduced directly into species growth equations and then apply stochastic perturbations around equilibrium points to the second model. The models aim to establish conditions for existence, uniqueness, and positivity of solutions, and to evaluate long-term population behavior under random environmental influences. Stability conditions are derived using Lyapunov theory. Using Itô's formula, stochastic persistence and boundedness are observed. The Euler–Maruyama method is used to simulate complex dynamics that are absent in the deterministic setting. Results indicate that low-intensity noise promotes species coexistence and enhances plankton persistence, whereas high-intensity noise may destabilize the system and generate new equilibria. These findings highlight the role of environmental randomness in shaping ecological stability and species persistence and the broader ecological significance of this work lies in its demonstration that stochastic drivers can act as both stabilizing and destabilizing forces depending on their intensity. By integrating mathematical rigor with ecological interpretation, the models provide a robust framework for predicting biodiversity outcomes under uncertainty, thereby offering valuable insights for conservation and management strategies.

KEYWORDS: Environmental stochasticity, Itô's calculus, Lyapunov stability, Stochastic persistence

1 INTRODUCTION

Phytoplankton organisms are the dominant primary producers in the pelagic environment. Zooplankton are the animals in the plankton community and feed on other phytoplankton (Anderson, Cembella, & Hallegraeff, 2012). The interaction between these phytoplankton and zooplankton is very important to balance aquatic ecosystems and biodiversity (Behrenfeld & Boss, 2014). In the last two decades, there has been a global increase in *harmful plankton blooms (HABs)* in the aquatic ecosystem.

Due to the universal existence and importance of plankton, understanding of the dynamical behaviors of interacting

species will continue to be a predominant topic. Deterministic mathematical models are frequently used to represent the plankton interactions for a deeper understanding of their dynamics. This approach, however, always has some limitations as it does not capture the environmental noise or demographic stochasticity which are important features of any natural system. In this work, a deterministic model of phytoplankton-zooplankton interaction is first reviewed, and compare its dynamics are compared with its stochastic version, which is formulated using two different stochastic perturbation techniques.

The main objective of the study is to formulate SDE models based on two different stochastic perturbation techniques. Both stochastic models are then analyzed and their asymptotic dynamics are discussed. Using sample paths and probability histograms, high variations are observed in the stochastic models, which cannot be captured in the deterministic system. The random behaviors of the system are contrasted with those of its deterministic counterpart.

The significance of this study lies in its contribution to understanding how environmental fluctuations affect plankton dynamics. Stochastic models will be formulated by incorporating stochasticity in the plankton interactions that are due to environmental noise, which is missing in the deterministic system.

2 LITERATURE REVIEW

De Silva and Jang in 2018, developed a stochastic model that specifically incorporates toxin-producing phytoplankton into phytoplankton-zooplankton interactions (De Silva & Jang, 2018). In this work, they presented stochastic models, including the continuous time Markov chain (CTMC) and Itô's stochastic differential equations (SDE), developed from a deterministic system of three interacting plankton populations. Unlike deterministic models, stochastic models incorporate population variability due to random births and deaths. Overall, the research highlights the importance of stochastic modeling in capturing real-world dynamics, particularly in predicting sudden population extinctions.

In 2021, Liu et al. proposed a model for the stochastic phytoplankton-zooplankton system with two harmful

phytoplankton populations (Liu et al., 2021). The findings of this research show that low-intensity white noise can lead to the coexistence of all species, while high-intensity noise can accelerate the extinction of plankton populations. Here they explained that white noise increases stochastic oscillations in plankton density and can significantly alter the dynamics of the phytoplankton-zooplankton system.

Chattopadhyay et al. established the role of toxins in reducing zooplankton grazing pressure (Chattopadhyay, Sarkar, & El Abdllaoui, 2002). They conducted field studies to monitor the populations of *Noctiluca scintillans* and *Paracalanus* and used mathematical models.

3 METHODOLOGY

The study begins with the deterministic model proposed by De Silva and Jang. Which is defined by:

$$\begin{aligned}\frac{dP_1}{dt} &= r_1 P_1 \left(1 - \frac{P_1 + c_1 P_2}{K_1} \right) \\ &\quad - \left(\frac{\alpha_1 P_1 Z}{1 + a_1 P_1 + a_2 Z} \right) \\ \frac{dP_2}{dt} &= r_2 P_2 \left(1 - \frac{P_2 + c_2 P_1}{K_2} \right) \\ &\quad - \left(\frac{\alpha_2 P_2 Z}{1 + b_1 P_1 + b_2 P_2} \right) \\ \frac{dZ}{dt} &= \left(\frac{\beta_1 P_1 Z}{1 + a_1 P_1 + a_2 Z} \right) \\ &\quad - \left(\frac{\beta_2 P_2 Z}{1 + b_1 P_1 + b_2 P_2} \right) \\ &\quad - \mu Z\end{aligned}$$

To incorporate environmental randomness, we introduce white noise in two distinct ways:

Model 1 (Noise on Parameters): White noise is introduced into the deterministic model in a multiplicative form.

$$\begin{aligned}
 dP_1 &= \left(r_1 P_1 \left(1 - \frac{P_1 + c_1 P_2}{K_1} \right) \right. \\
 &\quad \left. - \left(\frac{\alpha_1 P_1 Z}{1 + a_1 P_1 + a_2 Z} \right) \right) dt \\
 &\quad + \sigma_1 P_1 dB_1(t) \\
 dP_2 &= \left(r_2 P_2 \left(1 - \frac{P_2 + c_2 P_1}{K_2} \right) \right. \\
 &\quad \left. - \left(\frac{\alpha_2 P_2 Z}{1 + b_1 P_1 + b_2 P_2} \right) \right) dt \\
 &\quad + \sigma_2 P_2 dB_2(t) \\
 dZ &= \left(\left(\frac{\beta_1 P_1 Z}{1 + a_1 P_1 + a_2 Z} \right) \right. \\
 &\quad \left. - \left(\frac{\beta_2 P_2 Z}{1 + b_1 P_1 + b_2 P_2} \right) \right. \\
 &\quad \left. - \mu Z \right) dt + \sigma_3 Z dB_3(t)
 \end{aligned}$$

Where P_1, P_2, Z are Non-Toxin Producing Phytoplankton, Toxin Producing Phytoplankton, and Zooplankton.

Model 2: In the second case we assume Stochastic perturbation of the state variables:

$$\begin{aligned}
 dP_1 &= \left(r_1 P_1 \left(1 - \frac{P_1 + c_1 P_2}{K_1} \right) \right. \\
 &\quad \left. - \left(\frac{\alpha_1 P_1 Z}{1 + a_1 P_1 + a_2 Z} \right) \right) dt \\
 &\quad + \sigma_1 (P_1 - \tilde{P}_1) dB_1 \\
 dP_2 &= \left(r_2 P_2 \left(1 - \frac{P_2 + c_2 P_1}{K_2} \right) \right. \\
 &\quad \left. - \left(\frac{\alpha_2 P_2 Z}{1 + b_1 P_1 + b_2 P_2} \right) \right) dt \\
 &\quad + \sigma_2 (P_2 - \tilde{P}_2) dB_2 \\
 dZ &= \left(\left(\frac{\beta_1 P_1 Z}{1 + a_1 P_1 + a_2 Z} \right) \right. \\
 &\quad \left. - \left(\frac{\beta_2 P_2 Z}{1 + b_1 P_1 + b_2 P_2} \right) \right. \\
 &\quad \left. - \mu Z \right) dt \\
 &\quad + \sigma_3 (Z - \tilde{Z}) dB_3
 \end{aligned}$$

Where P_1, P_2, Z are Non-Toxin Producing Phytoplankton, Toxin Producing Phytoplankton, and Zooplankton.

The Euler–Maruyama numerical method is used to simulate the SDE model. A total of 10,000 sample paths of the SDE systems, along with the deterministic models, are provided. Using the sample paths and probability histograms, the high variability of the SDE model is then demonstrated.

Table 1: Defining the parameters of the model. (De Silva & Jang, 2017)

Parameter	Definition
P_1, P_2, Z	Non-Toxin Producing Phytoplankton, Toxin Producing Phytoplankton, and Zooplankton.
r_1, r_2	Intrinsic growth rates
K_1, K_2	Carrying capacities of NTP and TPP, respectively
c_1, c_2	Competition effects of TPP on NTP and NTP on TPP, respectively
β_1	Maximum rate of gain in zooplankton growth due to predation on NTP
β_2	Rate of inhibition of zooplankton growth due to toxic material ingested from TPP
μ	Natural mortality rate of zooplankton
a_1	Inverse of half-saturation constant
a_2	Coefficient of mutual interference
b_1	Coefficient of avoidance
b_2	Inverse of half-saturation constant

σ_i	Noise intensity for P_i , where P_i , where $i = 1,2,3$
$dB_i(t)$	Standard white noise, where $i = 1,2,3$

4 DATA ANALYSIS AND RESULTS

Before conducting numerical simulations, the stochastic system was first analyzed using Itô's formula and Lyapunov stability theory. The Lyapunov function was used to prove the system admits a unique global positive solution. Using Itô's formula, we showed that the expected values of the state variables are bounded over time. Moreover, for sufficiently small noise intensities, we proved the stochastic persistence of all populations. That is, the solutions remain away from extinction in probability as time approaches infinity.

These analytical results provide the theoretical foundation for the numerical simulations that follow.

To complement the theoretical analysis, we performed numerical simulations using the Euler–Maruyama method implemented in MATLAB. This allowed us to visualize the time evolution of the populations under varying levels of environmental noise.

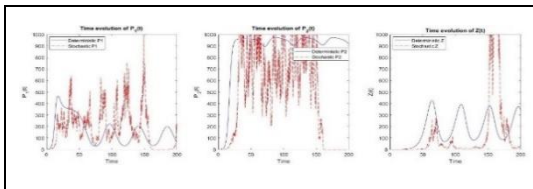


Figure 1: Time evolutions of systems populations of deterministic system and stochastic system are plotted with $\sigma_i = (0.03)$

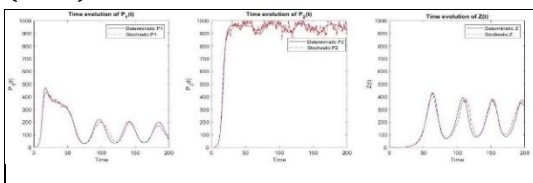


Figure 2: Time evolutions of systems populations of deterministic system and stochastic system are plotted with $\sigma_i = (0.4)$

Simulations were conducted for a range of white noise intensities $\sigma_i \in \{0.03, 0.4\}$, while keeping other parameters consistent with those in the deterministic model. For small noise levels $\sigma_i = (0.03)$, all three populations remained close to their deterministic trajectories, showing minor stochastic oscillations around equilibrium. Higher white noise intensity $\sigma_i = (0.4)$ can establish shift the system's dynamics and lead to a different stationary distribution.

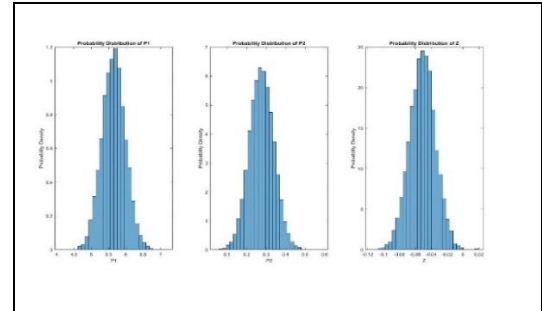


Figure 3: Probability distribution of plankton populations $\sigma_i = (0.03)$ $i = 1, 2, 3$

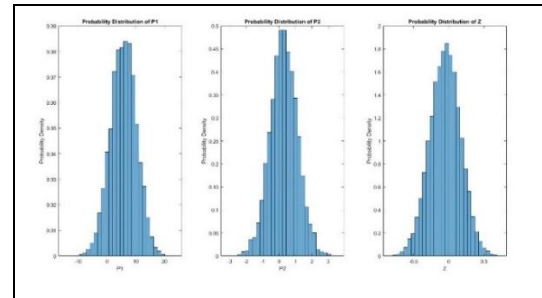


Figure 4: Probability distribution of plankton populations $\sigma_i = (0.4)$ where $i = 1, 2, 3$

The histograms reveal that P_1 exhibits the highest variability under stochastic perturbations, with a broad and sensitive distribution, while P_2 shows moderate fluctuations and Z remains the most stable with minimal deviation. In contrast, under lower noise intensity, P_1 and P_2 both display narrow, symmetric distributions

around steady-state values, and Z maintains a sharply peaked distribution near zero, indicating strong resistance to stochastic disturbances.

5 DISCUSSION

This study highlights the importance of incorporating environmental stochasticity into ecological models, especially in systems involving toxin-producing and non-toxin-producing phytoplankton. By introducing white noise into model parameters and state variables, we uncovered complex dynamical behaviors that deterministic models cannot capture. Specifically, **Stochastic Oscillations** under low-intensity noise, population trajectories showed persistent fluctuations around equilibrium values, suggesting the system supports quasi-stable coexistence despite environmental variability.

These results reinforce that stochastic models provide a more realistic understanding of ecological systems by capturing unpredictable behaviors, such as sudden collapses or persistent variability, which are not visible in deterministic approaches.

6 CONCLUSION

This study investigates the interactions between non-toxic phytoplankton, toxic phytoplankton, and zooplankton using two stochastic models. The models analyze environmental noise impact and dynamics, finding positive solutions under certain conditions. We use Itô's formula to find theoretical results for stochastic persistence and bounds. The study demonstrates that low white noise intensity can positively influence plankton persistence, while higher white

noise intensity can establish new equilibrium. The environmental stochastic can stabilize the system.

Future work may include numerical simulations for the second stochastic model to obtain new results. There is also the opportunity to conduct further extended research on the analytical part of the second stochastic model.

REFERENCES

- Anderson, D. M., Cembella, A. D., & Hallegraeff, G. M. (Eds.). (2012). *Harmful algal blooms: A compendium desk reference*. John Wiley & Sons.
- Behrenfeld, M. J., & Boss, E. S. (2014). Resurrecting the ecological underpinnings of ocean plankton blooms. *Annual Review of Marine Science*, 6, 167–194.
- Chattopadhyay, J., Sarkar, R. R., & El Abdllaoui, A. (2002). A delay differential equation model on harmful algal blooms in the presence of toxic substances. *Mathematical Medicine and Biology: A Journal of the IMA*, 19(2), 137–161.
- De Silva, M., & Jang, S. R. J. (2017). Dynamical behavior of systems of two phytoplankton and one zooplankton population with toxin-producing phytoplankton. *Mathematical Methods in the Applied Sciences*, 40(12), 4295–4309.
- De Silva, T. M. M., & Jang, S. R. J. (2018). Stochastic modeling of phytoplankton–zooplankton interactions with toxin-producing phytoplankton. *Journal of Biological Systems*, 26(1), 87–106.
- Liu, H., Dai, C., Yu, H., Guo, Q., Li, J., Hao, A., & Zhao, M. (2021). Dynamics induced by environmental stochasticity in a phytoplankton–zooplankton system with toxic phytoplankton. *Mathematical Biosciences and Engineering*, 18(4), 4101–4126.
- Sarkar, R. R., Pal, S., & Chattopadhyay, J. (2005). Role of two toxin-producing plankton and their effect on phytoplankton–zooplankton system: A mathematical study supported by experimental findings. *BioSystems*, 80(1), 11–23.

- Smayda, T. J. (1997). Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea. *Limnology and Oceanography*, 42(5, Pt. 2), 1137–1153.



Forecasting Order Volumes for Resource Planning: A Case Study

Alweera AMIH¹, Panahatipola PMOP¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
imashahanshani1999@gmail.com

ABSTRACT

Effective resource allocation in warehouse logistics requires correct forecasts regarding the number of orders on any given day to optimize labour, staging space, and transport allocation. For its development, a data model for forecasting the number of orders, measured in cases, across various zones of a case-in-case-out facility operated by one of the largest third-party logistics companies in Sri Lanka, has been proposed in this research work. One-year operational data is processed for analysis, combining descriptions and time series analysis techniques. The forecast model is built in Python based on the Extreme Gradient Boosting (XGBoost) algorithm, chosen for its aptness in identifying complex relationships in data. Evaluation criteria for forecast accuracy include Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2). XGBoost performed with an MAE value of 194.60, an RMSE value of 355.67, and an R-Squared value of 0.87, which was more effective than the classic ARIMA model, with which it was compared, resulting in higher error measures for the latter. Feature contribution analysis showed that lag, seven-day rolling averages, and day-of-the-week features contribute largely to variation in demand. The findings in this research validate the utility of short-term forecasts generated with techniques such as XGBoost for accurately describing variability in demand, thereby facilitating labour scheduling, optimal allocation, and efficient transport planning. In conclusion, XGBoost, being an efficient tool for short-term optimisation in logistics, has immense real-world scope for implementation in warehouses having multiple zones.

KEYWORDS: ARIMA, Forecasting, Resource Planning, Warehouse Logistics, XGBoost

1 INTRODUCTION

Logistics operations are progressively depending on data-informed decision-making to enhance efficiency. In environments where goods are managed in complete cases rather than as individual units, there is a notable introduction of variability in daily workloads. This variability has a direct effect on the planning of transport capacity, the utilisation of staging space, and the allocation of labour. This study examines a prominent third-party logistics provider in Sri Lanka, which manages a multi-zone warehouse consisting of four primary operational

zones: A, B, S, and STG. This study aims to develop a forecasting model using Python and XG-Boost to predict daily case volumes in the specified zones. Precise predictions facilitate anticipatory resource management and improve operational effectiveness in warehouse logistics.

2 LITERATURE REVIEW

2.1 Literature Related to the Study

Nonlinear and irregular demand patterns in warehouse operations are frequently missed by traditional forecasting techniques like ARIMA, exponential smoothing, and moving

averages (Box et al., 2016). Superior predictive performance has been demonstrated by machine learning models, especially ensemble techniques like XGBoost, which can handle complex datasets with numerous interacting features (Makridakis et al., 2018; Ghosh & Gosavi, 2020).

According to Hayta et al. (2023), XG-Boost provides regularization to prevent overfitting, computational efficiency, and interpretable feature importance metrics. Zhao and Bian (2022) showed that XGBoost performed better than classical models in multi-zone warehouse settings by precisely capturing daily peaks and fluctuations. These results demonstrate the value of machine learning-based forecasts for optimal vehicle loading in logistics operations, congestion management, and dynamic workforce planning.

2.2 Theoretical Background

In order to guarantee effective labour, staging space, and transportation resource allocation, demand management theory places a strong emphasis on matching resources with anticipated demand (Mentzer & Moon, 2004). This study justifies the use of XG-Boost for resource optimization because its precise daily forecasts facilitate effective labour allocation and space management in case-in-case-out warehouse operations. According to Lee, Padmanabhan, and Whang (1997), the Bullwhip Effect Theory explains how minor variations in demand can magnify inefficiencies throughout the supply chain. XG-Boost's short-term forecasts lessen these variations, avoiding overstaffing and bottlenecks and preserving efficient warehouse operations.

3 METHODOLOGY

In order to extract calendar features like day, week, month, and seasonality, the study used secondary operational data from the company's Warehouse Management System (WMS), which included timestamps and daily case volumes for each warehouse zone. Preprocessing ensured data reliability by addressing missing values, standardizing formats, fixing naming inconsistencies, and examining outliers while maintaining true demand peaks. The dataset recorded weekly peaks and operational cycles. In order to maintain operational sequences, missing values were handled using interpolation for short consecutive gaps and forward filling for isolated gaps. The target variable (daily case volumes), time-related characteristics, rolling averages (3-day and 7-day), lagged demand (prior day or week), and zone identifiers were among the important variables. These characteristics were designed to improve the predictive accuracy of XGBoost while preserving interpretability for operational choices.

To simulate real-world forecasting scenarios, the data were split chronologically into 20% for testing and 80% for training. To reduce prediction errors, cross-validation was used to fine-tune the XG-Boost model's hyperparameters, such as learning rate, maximum depth, and number of estimators. The XG-Boost model was contrasted with a traditional ARIMA benchmark model to guarantee scientific validity. Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2) were used to assess the model's performance. R^2 shows the percentage of

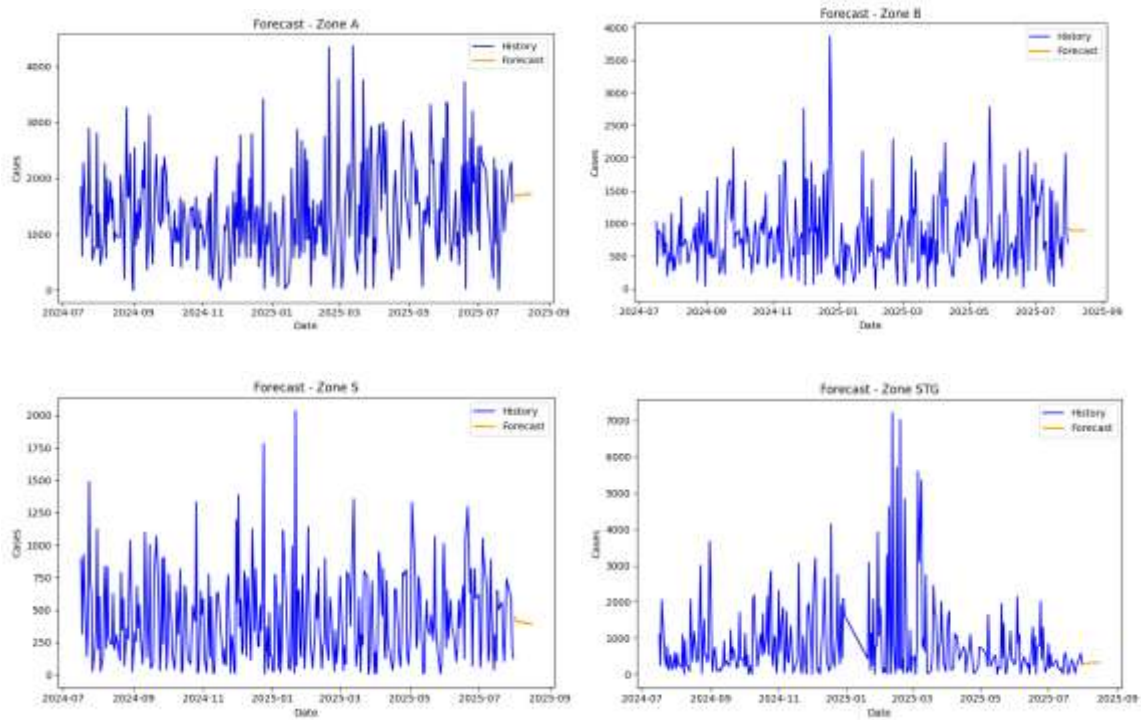


Figure 1: Forecasted Case Volumes for Zone A, B, S & STG

variance in the target variable that the model explains, RMSE penalises larger errors more severely, and MAE measures the average magnitude of errors in absolute terms. The most significant predictors for operational decision-making were found using feature importance analysis.

4 RESULTS AND DISCUSSION

The XG Boost model predicted daily case volumes accurately in the short term. The evaluation metrics showed strong model performance and reliability in predicting daily order fluctuations: MAE = 194.60, RMSE = 355.67, and $R^2 = 0.87$. On comparing with the ARIMA benchmark model, trained in the same fashion, the model generated an MAE value of 285.34, an RMSE value of 470.12, with an R^2 value of 0.72, thereby substantiating its improved performance on nonlinear demand patterns, especially

in different zones. Feature importance analysis indicated demand series lag features, features containing 7-day rolling averages, followed by features with day-of-week information, and features containing information about individual zones. It thereby articulates short-term autocorrelation in demand patterns, with inherent weekly cycles, thereby pointing out operational patterns in demand variability, like ordering schedules in clients.

From an operational standpoint, incorporating XG-Boost forecast predictions into warehouse planning algorithms facilitates proactive labor allocation and balancing. Executives can effectively reduce downtime costs by proactively identifying workload forecasts to schedule employees accordingly. The versatility feasible in forecast model predictions facilitates individualised scheduling in respective

zones, aligning labour forces with forecasted workloads. These results support other studies affirming improved accuracy in short-term logistics planning forecasts from machine learning algorithms over traditional approaches (Zhao & Bian, 2022; Makridakis et al., 2018). Another graphic illustrating the test set's observed data with forecasted values aptly shows better identification in peak patterns with irregular variations than in the competing ARIMA model.

5 CONCLUSION

A Python-based XG-Boost predictive model with forecasted daily orders for a multiple zone, case in, case out type warehouse was also developed in this research. It attained significant accuracy, outperforming its ARIMA model counterpart in all testing criteria. Using XG-Boost, efficient labor allocation, utilization of staging space, or transport allocation for fulfilment can now be accomplished due to its ability to deal with numerous nonlinear patterns in demand. Its applications in forecast-based decision-making make XG-Boost-based predictive analytics an efficient short-term optimization solution for logistics.

REFERENCES

- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). Time series analysis: Forecasting and control (5th ed.). John Wiley & Sons.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Chatfield, C. (2004). The analysis of time series: An introduction (6th ed.). CRC Press.
- Chopra, S., & Meindl, P. (2019). Supply chain management: Strategy, planning, and operation (7th ed.). Pearson.
- Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2008). Effective forecasting and judgmental adjustments in supply chain planning. *International Journal of Forecasting*, 24(1), 3–26.
- Ghosh, S., & Gosavi, A. (2020). Improving forecasting in supply chain management using machine learning: A review and case-based analysis. *Computers & Industrial Engineering*, 142, 106376.
- Hayta, E., Gencturk, B., & Koklu, M. (2023). Predicting future demand in the logistics sector using machine learning methods. *Intelligent Methods in Engineering Sciences*, 2(4), 102–114.
- Lee, H. L., Padmanabhan, V., & Whang, S. (1997). The bullwhip effect in supply chains. *Sloan Management Review*, 38(3), 93–102.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 13(3), e0194889.
- Mentzer, J. T., & Moon, M. A. (2004). Sales forecasting management: A demand management approach. SAGE Publications.
- Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77–84.
- Zhao, K., & Bian, J. (2022). Demand forecasting in logistics using hybrid machine learning approaches. *Expert Systems with Applications*, 191, 116282.



Identifying Key Effecting Factors of Employee Service Duration in the ABC Apparel Company, Katunayake Using Lasso Regression

Nagasinghe NKCJ¹, Lakshmi MPDS¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
chamodaninagasinghe99@gmail.com

ABSTRACT

High employee labor turnover rates are a recurring and expensive problem Sri Lanka's apparel industry causing to the high operational cost. Improving organizational resilience and competitive advantage requires an understanding of the precise factors causing this turnover in the local context. The purpose of this study is to determine and examine the main variables affecting to employee turnovers at the ABC Apparel Solutions facility in Katunayake, Sri Lanka because currently it experiences in high labor turnover. As the labor turnover rate is inversely propotional to the service month duration of employees, we focus on effecting factors of employee service duration. The company's human resources information system provides secondary data, including records on employee demographics, and pay for a given time frame. A Lasso regression analysis was carried out after applying data preprocessing. The findings show that age, employee status (direct/indirect), districts and the department are the factors, predictors for the labor service duration. This study revealed that applying Lasso regression enabled the identification of critical factors driving labor turnover, providing valuable insights for managerial decision-making. With possible applications for the larger industry, the suggestions offer ABC company a workable framework to decrease turnover, cut expenses, and improve operational stability. Previously none of the research has been conducted to identifying the factors affecting labor turnover using the Lasso Regression. Hence this study focuses on finding the solution for this critical issue.

KEYWORDS: Apparel Industry, Labor Turnover, Lasso Regression, Service months

1. INTRODUCTION

ABC Apparel Solutions (Pvt) Ltd is a subsidiary of the ABC group which known as a major apparel group in not only Sri Lanka but also globally. The plant, is currently facing in high labor turnover within the plant. Turnovers are inversely propotional to employee service durations (Aksu, A., 2008). So the primary objective of the study is to identify the factors effecting employee service months in ABC Apparel Solutions (Pvt) Ltd, Katunayake. Secondary objectives are to provide recommendations to reduce labor turnover in the plant, to support operation department

to increase their efficiency by reducing the labor turnover by increasing the service duration of the employees. This will help plant to increase their profit by mitigating the employee attrition while increasing their job satisfaction.

While a lot of worldwide studies on employee turnover have been conducted (Porter & Steers, 1973), there is still a lack of research in the apparel sector of Sri Lanka. A majority of the local researches focus on wages and working conditions (Karunaratne, H. D., 2007), but they give little or no quantitative evidence on how demographic, job related and organizational factors influence the turnover rate. In

addition to that, only a handful of studies have delved into the issue of turnover at ABC Apparel Solutions (Pvt) Ltd, Katunayake, even though it is one of the leading manufacturers of the apparel industry. So, the purpose of this study is to fill research gap using Lasso regression analysis.

Considering the practical importance for the ABC group, this study will provide the practical insights of the key factors of labor turnover at Katunayake plant, which may help to HR and Administration departments to identify those factors and designing effective strategies with the help of corporate and top level management of the group. This study will add empirical evidence about labor turnover patterns in Sri Lanka's apparel sector, which is a critical part of the economy of the country.

2. LITURATRE REVIEW

Labor turnover refers to the rate that an existing employee leave the particular company by replacing those vacancies by new employees. Moreover, turnover is influenced by both individual and organizational characteristics, including satisfaction, expectations, and external opportunities (Price, J. L., 1977). Turnover rate is inversely propotional to the service month duration of employees (Dess & Shaw, 2001). It can be defining as turnover as the sequence of withdrawal behaviors that result involuntary termination of employment (Mobley, W. H., 1977). Turnover is commonly measured using indicators such as quit rates, retention rates and turnover may be categorized into voluntary, involuntary, functional, dysfunctional, avoidable, and unavoidable

types (Abassi & Hollman, 2000). Voluntary turnover is job quitting by the employees themselves, typically for a better job or their own personal reasons, whereas the company fired the employee leading to poor performance or layoffs is an instance of involuntary turnover. The term turnover has also been broken up into functional and dysfunctional by academics, whereby the first one entails the departure of employees whose work performance has been below standard while the second one is the elimination of the most productive staff members (Griffeth et.al., 2000).

Regarding to the theoretical background, classical theories, motivational and job satisfaction theories and modern theories can be identified. In motivational and job satisfaction theories side, according to Maslow's (1943) Hierarchy of Needs, staff members will depart from companies that do not meet their fundamental needs like safety, esteem, growth.

3. METHODOLOGY

Since the primary goal of this study is to statistically investigate the factors that significantly affect employee service duration in the apparel industry, a quantitative research design is used. This design is suitable since it enables the researcher to use regression analysis to test relationships between several variables at once. While service months is playing the role of dependent variable of the model, factors like age, gender, district, designation, department, basic salary are the independent variables of the study. Starting with well-known theories of employee turnover like the Price-Mueller Model (1981), Mobley's Turnover Model (1977), and Social

Exchange Theory (Blau, 1964), the study employs a deductive reasoning approach.

The main goal of this study is to determine and quantify the impact of a few chosen factors like basic salary, department, age, gender on employee turnover at ABC Apparel Solutions, Katunayake, a quantitative research approach is used. This study is entirely based on secondary data from company records from the HR department. After employing in data preprocessing part, categorical variables converted into factors. Assumptions were checked to identify whether the dataset is suitable for employing in linear regression. But the assumptions were not satisfied. Hence Lasso regression was done for the dataset.

4. DATA COLLECTION AND ANALYSIS

4.1 Data Collection

The data of the study was collected from the HR department of the plant, as the secondary data. The dataset contains several categories of variables such as service months, designation, district, department, gender and age which are representing the resignation report of the plant from 1st May 2024 to 5th May 2025. The data, originally maintained in the organization's HR information system, was exported to a spreadsheet format and provided to the researcher for analysis. The sample size of this study is 554, which contains the details of employee resignation report of the ABC Apparel Solutions (Pvt) Ltd, Katunayake.

4.2 Analysis

Lasso regression is suitable for using in moderate sized data set as in this study, (Friedman et.al., 2010). Lasso regression was used because, in contrast to Ridge and Elastic Net, it offers efficient variable selection by reducing unimportant coefficients to zero, making it better suited for determining the study's most significant predictors (Zhang et.al., 2017).

To analyze the data set, as the first step, data preprocessing steps were done. A few variables in the dataset had missing values, according to an initial scan. The dataset shows that it includes many categorical variables. To do the statistical analysis, all these categorical factors were converted into factors. To ensure the consistency of the data set, there occurs several checks. Outliers were identified and removed because these outliers showed extremely high or low values that were occurred in clearly data entry errors.

5. RESULTS AND DISCUSSION

Lasso linear regression model effectively explains the factors for labor turnover in the company using the independent variables. Table 1 shows that there exists two Lambda values, means that there are two models for this study. The model which is having minimum lambda value has 21 predictors (min model) while other model consisting 3 predictors (1se).

Table 2 shows that min model has the highest R^2 value, 0.8235 than the 1se model. Min model consists with approximately 82.35% of the variance in the dependent variable is explained by the independent variables included in this model. Hence, this model can be accepted as the final fitted

model of the study. Table 3 shows that satisfied coefficient of the best Regression Model which provide relationship between the dependent variable, service months with the satisfied independent variables. Based on the regression analysis following factors, age, DirectIndirect, Department of Finishing & Packing, General Administration, Industrial Engineering, Quality Assurance, Training School and warehouse are identified as the root causes of the employee turnover in ABC Apparel Solutions (Pvt) Ltd, Katunayake. Considering the coefficients, which has the positive value, may tend to work for long time rather than negative coefficients.

Table 1: Mean-Squared Error summary of model

	Min	1se
Lambda	0.2599	1.5221
Index	44	25
Measure	64.33	68.69
SE	5.224	5.110
Nonzero	21	3

Table 2: Summary of adjusted R^2 values of models

	Min	1se
Lambda	0.2599	1.5221
R^2	0.8235	0.8002

Table 3: Coefficients of the best regression model

Independent variable	Coefficient
Intercept	14.9551
Age	1.9879
DistrictAnuradhapura	- 0.8557
DistrictColombo	- 1.4913
DistrictGalle	3.9714
DistrictJaffna	- 1.5349
DistrictKalutara	1.5154
DistrictKandy	-1.4467
DistrictKilinochchi	-4.3738
DistrictMatara	0.9142
DistrictMonaragala	- 1.4841
DistrictNuwara Eliya	1.2941
DistrictPolonnaruwa	0.5148
DistrictTrincomalee	0.7027
DistrictVavuniya	- 2.5852
DepartmentFinishing & Packing	2.3676
DepartmentGeneral Administration	- 3.1284
DepartmentIndustrial Engineering	10.5615
DepartmentQuality Assurance	4.6943
DepartmentTraining School	- 4.5175
DepartmentWarehouse	4.1531
Direct.indeirectIndirect	- 13.3447

REFERENCES

- Aksu, A. (2008). Employee turnover: calculation of turnover rates and costs. *In Handbook of hospitality human resources management* (pp. 195-222). Routledge.
- Dess, G. G., & Shaw, J. D. (2001). Voluntary turnover, social capital, and organizational performance. *Academy of management review*, 26(3), 446-456.
- Fugar, F. D. (2007). Frederick Herzberg's motivation-hygiene theory revisited: The concept and its applicability to clergy (A study of fulltime stipendiary clergy of the global evangelical church, Ghana). *Journal of Science and Technology (Ghana)*, 27(1), 119-130.
- Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33, 1-22.
- Griffeth, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of management*, 26(3), 463-488.
- Gunawardena, V. (2014). Labor turnover in the Sri Lankan apparel industry: Causes and consequences. *Sri Lanka Journal of Human Resource Management*, 5(1), 23-37.
- Karunaratne, H. D. (2007). Structural change and the state of the labour market in Sri Lanka.
- Zhang, Y., Yang, Y., Sun, J., & Zhang, T. (2017). A comparative study of regularization methods for high-dimensional data analysis. *IEEE Access*, 5, 24387-24395.



Predicting Airline Passenger Arrivals to Sri Lanka from Asia, Europe, and the Middle East: A Comparative Study

Dias PPVA¹, Appuhamy PADAN¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
virajidias95@gmail.com

ABSTRACT

Air transport plays a vital role in Sri Lanka's economic growth, particularly through tourism. Forecasting international passenger arrivals has been made more complex by global uncertainties such as the COVID-19 pandemic, economic crises, and political instability. This study investigates the application of advanced machine learning (ML) and deep learning (DL) models in predicting daily passenger arrivals to Sri Lanka from the Asian, European, and Middle Eastern regions based on several socio-economic and political factors. Daily passenger booking data from 2019 to 2024 and socio-economic factors like exchange rate fluctuations, political alterations, and health crisis periods were used for modelling. Initially, the descriptive analysis and time series plots were obtained to observe the variations in passenger demand throughout the study period. Then, four machine learning models, viz. Random Forest, XGBoost, LightGBM, CatBoost, and three deep learning models viz. RNN, LSTM, and GRU models were developed and the predictive performance of all seven models was evaluated using R^2 , MAE, RMSE and MAPE. The analysis revealed that the number of air passengers from the Asian region was the highest during the study period, followed by Europe and the Middle East. Among all models, XGBoost was selected as the best overall model for the Asian and Middle East regions, while RNN was chosen for Europe based on the balance between prediction error and variance capture. The findings emphasize COVID-19's impact on travel and the effectiveness of ensemble ML models on medium-sized datasets, providing insights for authorities to improve demand forecasting and resource planning.

KEYWORDS: Air passenger forecasting, Deep Learning, Machine Learning, Socio economic factors, XGBoost

1. INTRODUCTION

Sri Lanka's tourism sector remains a cornerstone of its economy, generating USD 3.6 billion in 2019 and heavily relying on international air travel for connectivity (SLTDA, 2019). In recent decades, hundreds of thousands of people in the country have depended on this sector for their livelihood. However, this sector has become extremely sensitive to global shocks such as pandemics, exchange rate fluctuations, and political instability. For instance, tourist arrivals to Sri Lanka have drastically dropped during the Covid-19 period and the political instability era. As a

consequence, the passenger arrival patterns fluctuated dramatically and showed complex behaviour. Accurate prediction of air passenger arrivals is vital for resource optimization, airline management, and tourism development. Tourists from different regions have different cultural preferences, spending patterns, food preferences and travel behaviors. Therefore predicting regional demand is more important than general trends to provide visitors' expectations effectively and for sustainable tourism planning. Most of the previous research used traditional statistical models often fail to capture complex patterns and

sudden shifts (Li et al., 2019). This study aims to predict air passenger arrivals to Sri Lanka from Asia, Europe, and the Middle East based on socio-economic factors using Machine Learning(ML) and Deep Learning(DL) techniques, to select the best predictive model.

2. LITERATURE REVIEW

Traditional models such as AutoRegressive Integrated Moving Average (ARIMA) and econometric regression have been widely used for air traffic forecasting. Studies in Sri Lanka and India (Higgoda & Madurapperuma, 2020; Priyadarshana & Fernando, 2015; Sharma & Mehta, 2023) showed that although these models capture long-term trends, they struggle to handle crises and external shocks.

The COVID-19 pandemic revealed significant flaws in traditional forecasting methods. Waruna and Mahekha (2021) studied the impacts of COVID-19 on demand for air travel to Sri Lanka and concluded substantial demand contraction and uneven recovery trends that could not be adequately modelled with conventional methods. Bandara and Jayasooriya (2024) applied the Almost Ideal Demand System to study the international tourism demand of Sri Lanka, considering price and income but not with machine learning methods.

ML represents a paradigm shift by allowing data driven learning from complex, nonlinear relationships. Nguyen et al. (2015) applied Artificial Neural Networks (ANN) to forecast Australian domestic air travel, outperforming ARIMA based on Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE)

metrics. Firat et al. (2021) also applied ML methods to forecast passenger load factor, they achieved an average error of 6.74% with Gradient Boosting, 6.76% with Random Forest, and 8.16% with ANN, substantially outperforming a linear regression baseline (~9.62%) for origin-destination demand.

Current advancements emphasize DL architectures. Afarini and Hindarto (2024) compared Long Short Term Memory (LSTM), Prophet, and Neural Prophet for forecasting air passenger arrival growth and found that LSTM effectively captures intricate patterns in time series data but responds less well to abrupt fluctuations. Few studies combine socio-economic factors with ML/DL for Sri Lanka, and none focus on region-specific daily arrivals, leaving a research gap this study addresses.

3. METHODOLOGY

Daily observations on foreign air passengers' arrivals to Sri Lanka from Asia, Europe, and the Middle East regions spanning from Jan 1, 2019 – Dec 31, 2024 were retrieved from IATA Direct Data Solutions (International Air Transportation Association). Moreover, daily exchange rate (LKR/USD), Period of COVID-19 (Pre Covid, Covid, Post Covid), Political era (P1,P2,P3 and P4), represent the four different presidential periods in Sri Lanka during the study period, and day type (weekday/weekend) were collected as socio economic factors to predict air passenger arrival from different regions. Moreover, in the original dataset, weekend exchange rates were filled with the preceding Friday's values, ensuring no missing data. The exchange rate and political eras were specifically selected as socio-economic

factors because they showed notable fluctuations during the study period and could potentially influence passenger arrivals. During the study period, passenger arrivals from the three regions exhibited major change points before, during, and after the COVID-19 pandemic. To account for these changes, the COVID-19 period was included as a feature in the models. Data preprocessing involved cleaning, normalizing only the exchange rate and passenger count variables to improve the training stability and convergence speed of DL models. One-hot encoding was applied to all categorical variables to ensure reliability and improve model performance and interpretability. Initially, the descriptive statistics and time series plots of passenger arrivals from each region were obtained to identify the variation in passenger arrivals to Sri Lanka during the study period. Then, the dataset was split into two parts: 80% for training and 20% for testing the ML and DL models. Four ensemble ML models viz. Random Forest, XGBoost, LightGBM, and CatBoost and three DL models viz. Recurrent Neural Network (RNN), (LSTM), and Gated Recurrent Unit (GRU) were developed and compared. According to Hall and Rasheed (2025), these ML and DL models deliver the best performance in predicting time-dependent data. During the training process, grid search and cross-validation were used to prevent potential overfitting, and achieve the best performance. Next, the prediction performance of the model was evaluated using the coefficient of determination (R^2), Mean Absolute Error (MAE), (MAPE), and (RMSE), which provided a comprehensive snapshot of the model's

accuracy. Then all seven models for each region were compared to find the best model for each region. Finally, feature importance analysis was conducted to determine the socio-economic shocks that influenced passenger flows and travel behavior.

4. RESULTS AND DISCUSSION

Descriptive analysis and time series plots revealed significant volatility and seasonal spikes in passenger arrivals, along with noticeable structural breaks during major events like COVID-19 between 2019-2024. Asia contributed the highest arrivals, followed by Europe and the Middle East. Passenger numbers dropped sharply in 2020–2021 due to the pandemic, then gradually recovered by late 2022 is illustrated in Figure 1. Asia rebounded fastest, Europe slower due to economic and travel restrictions, while the Middle East remained relatively stable, driven by labour-related travel. Tourist arrivals differ significantly across the four political eras, with the highest average daily passenger arrivals from all three regions recorded during the P1 era.

According to Table 1, for the Asian region, XGBoost achieved the highest R^2 (0.9708), though its error metrics were higher than all DL models.

RNN model had the lowest MAE and RMSE but lower R^2 (0.6394), while LSTM had the lowest MAPE but very low R^2 , likely due to smooth predictions that missed sudden changes and limited data for generalization.

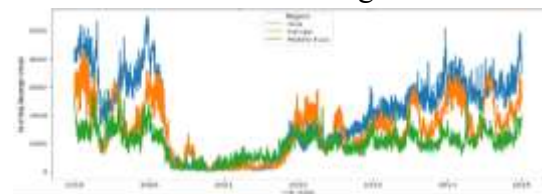


Figure 1: Daily Passenger Arrivals from 2019 to 2024 by Region

Table 1: Predictive performance of the models fitted for each region

Region	Evaluation Matric	RF	XGBoost	Light GBM	CatBoost	LSTM	RNN	GRU
Asia	R ²	0.9636	0.9708	0.9691	0.9695	0.3755	0.6394	0.4060
	MAE	0.1351	0.1241	0.1301	0.1299	0.0578	0.0434	0.0567
	RMSE	0.1886	0.1706	0.1755	0.1742	0.0728	0.0546	0.0701
	MAPE	0.1384	0.1105	0.3634	0.1545	0.0452	0.0574	0.0645
Europe	R ²	0.9526	0.9620	0.9571	0.9619	0.7838	0.8869	0.7949
	MAE	0.1551	0.1321	0.1301	0.1435	0.0554	0.0572	0.0520
	RMSE	0.2200	0.1928	0.2049	0.1905	0.0673	0.0689	0.0666
	MAPE	0.1596	0.1240	0.4466	0.1245	0.0524	0.0459	0.0566
Middle East	R ²	0.8660	0.8964	0.8800	0.8885	0.5449	0.6720	0.6035
East	MAE	0.2680	0.2434	0.2609	0.2506	0.0607	0.0510	0.0561
	RMSE	0.3663	0.3341	0.3596	0.3465	0.0784	0.0666	0.0733
	MAPE	0.3325	0.2105	0.3214	0.2136	0.0596	0.0536	0.0651

Overall, XGBoost is considered the best model, balancing prediction error and variance capture. Similarly, for the Middle East, XGBoost had the highest R² and RNN had the lowest errors with low R²(0.6720), making XGBoost the preferred model. In contrast, for Europe, XGBoost had the highest R², but RNN showed a reasonably high R² (0.8869) with lower errors, making RNN the best overall model for this region.

Figures 2, 3, and 4 show the actual (testing data) versus predicted passenger arrivals to Sri Lanka obtained from the best predictive models for Asia, Europe, and the Middle East, respectively. The figures indicate that the models effectively captured unseen data, demonstrating stability and robustness across different regions. The feature importance analysis of the best XGBoost models for the Asia and Middle East regions revealed that the COVID-19

period was the most influential predictor (Asia=0.356, Middle East=0.298) followed by exchange rate volatility and political transitions. The weekday/weekend effect had a minor impact, indicating that external shocks primarily drive demand variation. Regionally, Asian demand was most responsive to policy relaxations, European demand to exchange rate shifts, and Middle East demand to labor market trends.

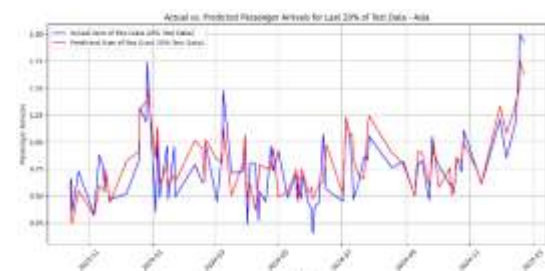


Figure 2: Actual vs Predicted Passenger Arrivals from Asia

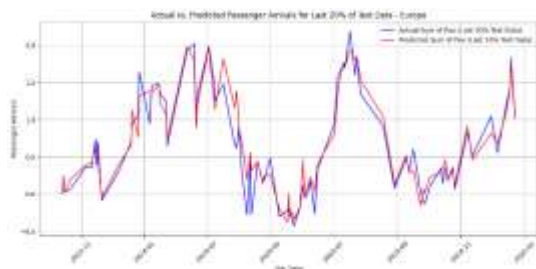


Figure 3: Actual vs Predicted Passenger Arrivals from Europe

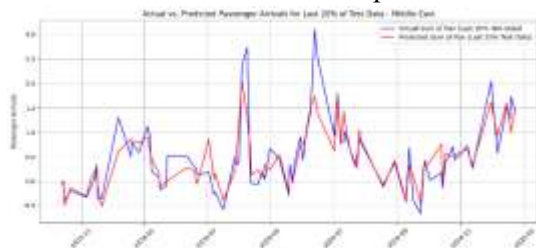


Figure 4: Actual vs Predicted Passenger Arrivals from the Middle East

5. CONCLUSION

This study compared machine learning and deep learning models to find the best models to forecast daily passenger arrivals to Sri Lanka from Asia, Europe, and the Middle East using socio-economic and temporal variables. The results showed that Asia recorded the highest number of arrivals, followed by Europe and the Middle East. XGBoost outperformed all other models for the Asian and Middle Eastern regions, while the RNN model was identified as the best overall model for Europe, achieving a better balance between prediction error and variance capture. The COVID-19 pandemic, exchange rate fluctuations, and political periods were key factors influencing tourist demand. Regionally, Asia showed the fastest recovery, Europe was more sensitive to currency changes, and the Middle East remained relatively stable. Overall, the findings highlight XGBoost's strength in modelling complex relationships and provide valuable insights for tourism and aviation planning.

REFERENCES

- Afarini, N., & Hindarto, D.(2024). Forecasting Airline Passenger Growth: Comparative Study LSTM VS Prophet VS Neural Prophet. *Sinkron*, 9(1), 505-513.
- Bandara, K., & Jayasooriya, T. (2024). Modeling international tourism demand for Sri Lanka using the Almost Ideal Demand System (AIDS) approach. *South Asian Journal of Tourism Economics*, 6(1), 33–48.
- Firat, M., Yiltas-Kaplan, D., & Samli, R. (2021). Forecasting air travel demand for selected destinations using machine learning methods. *Journal of Universal Computer Science*, 27(6), 564–581.
- Hall, T., & Rasheed, K. (2025). A Survey of Machine Learning Methods for Time Series Prediction. *Applied Sciences*, 15(11), 5957.
- Higgoda, D., & Madurapperuma, C. (2020). Air transport and economic growth causality in Sri Lanka: Evidence from 1983–2019. *Journal of Economic and Policy Studies*, 14(2), 112–126.
- IATA Direct Data Solutions. (2024). Passenger booking and traffic data for Sri Lanka (2019–2024). International Air Transport Association.
- Li, Q., Chen, W., & Song, H. (2019). Modeling international air travel demand with macroeconomic and external shock variables. *Annals of Tourism Research*, 76, 270–285.
- Priyadarshana, K., & Fernando, A. (2015). Econometric modeling of passenger traffic at Bandaranaike International Airport. *Sri Lankan Journal of Applied Statistics*, 16(1), 55–67.
- Sharma, A., & Mehta, R. (2023). Trend analysis and forecasting of air passenger growth in India using econometric and time series approaches. *Asian Journal of Transport Management*, 9(2), 145–160.
- Sri Lanka Tourism Development Authority (SLTDA) (2019). Annual statistical report 2019. Colombo: Sri Lanka Tourism Development Authority.
- Waruna, P., & Mahekha, D. (2021). COVID-19 and the disruption of air travel demand in Sri Lanka. *Colombo Economic Review*, 18(2), 77–89.



Optimization of Machine Setup Time During Style Changeovers in the ABC Apparel Company Using Linear Programming

Goonerathne GDN¹, Lakshmi MPDS¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
devninatasha123@gmail.com

ABSTRACT

The apparel industry is one of Sri Lanka's leading export sectors, where efficiency and productivity are essential for global competitiveness. ABC (PVT) Ltd, a major apparel manufacturer specializing in intimate wear, faces significant challenges due to excessive machine setup times during frequent style changeovers, which reduce line efficiency and overall productivity. In Sri Lanka, no previous research has been conducted to systematically optimize machine setup time in the manufacturing industry using Linear Programming. To address this gap, the present study introduces a data-driven framework that integrates statistical analysis and Integer Linear Programming (ILP) to optimize machine setup time. Time study data were collected from 735 machine-operation observations within the Industrial Engineering Department. Descriptive statistics revealed that average setup times ranged from 12.29 to 176.00 minutes, with standard deviations between 1.50 and 151.97 minutes. After applying a 15% allowance factor, standard setup times varied from 14.13 to 202.40 minutes. The ILP model developed in this study optimally allocated mechanics to setup tasks, minimizing the total setup duration to 4535.07 minutes, representing an 18% reduction compared with the existing system. The findings demonstrate that integrating statistical testing with optimization modelling is an effective approach for minimizing setup times and enhancing operational efficiency in apparel manufacturing, supporting lean manufacturing and Single Minute Exchange of Die (SMED) principles through standardized setup procedures and efficient resource utilization.

KEYWORDS: Apparel industry, Linear programming, Optimization, Setup time, Single Minute Exchange of Die

1 INTRODUCTION

The apparel industry is a vital contributor to Sri Lanka's economy, generating significant export revenue and employment. A major challenge in production is high machine setup time, which occurs when switching to new garment styles and includes mechanical adjustments, attachments, and verifications. Excessive setup times reduce line efficiency, delay production schedules, and lower overall productivity, particularly in factories with high product diversity (Mohomed, 2014). Traditional setup management often

relies on mechanical experience, resulting in inconsistent processes, time variability, and inefficient task allocation.

This study combines statistical analysis with an Integer Linear Programming (ILP) model to examine setup time variations across machine-operation combinations and optimize mechanic-task allocation. By providing a systematic, data-driven framework, the research aims to minimize downtime, standardize changeover processes, and improve overall production efficiency. The approach aligns with Lean Manufacturing and Single-Minute Exchange

of Dies (SMED) principles, enabling informed decision-making and creating a replicable method that can enhance efficiency across different departments in apparel manufacturing.

2 LITERATURE REVIEW AND THEORETICAL BACKGROUND

Reducing setup time is a core principle of Lean Manufacturing, as outlined by Shingo (2019) in the SMED methodology, which converts internal setup tasks into external ones to minimize changeover durations. Research shows that SMED improves productivity and responsiveness; for example, Karim (2013) reported a 35% reduction in sewing line changeover times, enhancing output and on-time delivery. Continuous training and standardized procedures further boost efficiency (Ali et al., 2025).

Statistical optimization, especially Linear Programming (LP), effectively minimizes setup delays. Studies by Tahar et al. (2006) and Antonioli et al. (2017) demonstrate that combining standardization with optimization models reduces variability and improves production performance. This study addresses this gap by developing an Integer Linear Programming (ILP) model, grounded in Lean Manufacturing and Operations Research theories, to optimize mechanic-task allocation and enhance efficiency during style changeovers.

3 METHODOLOGY

3.1 Data Collection

The study was conducted at ABC (PVT) Ltd, a leading apparel manufacturer in Sri Lanka, using 735 machine setup

observations across multiple production modules. Each observation recorded machine type, operation type, mechanic, and total setup time. Setup times were measured through stopwatch time studies and validated by Industrial Engineering (IE) staff. A 15% allowance factor was applied to account for small variations, resulting in the calculation of standard setup times.

3.2 Statistical Analysis

Data were analyzed using descriptive statistics, including mean, standard deviation, and range, to understand the variation of setup times.

3.3 Optimisation Modelling

An Integer Linear Programming (ILP) model was developed to minimize total machine setup time by optimally assigning mechanics to setup tasks. The model included two main constraints: each setup task must be assigned to exactly one mechanic, and each mechanic can handle only one task at a time to prevent overlap. The objective was to minimize total setup time, and solving the model produced an optimal allocation plan that achieved this goal.

3.4 ILP Model Formulation

Let:

$i = 1, \dots, n$: setup tasks

$j = 1, \dots, m$: mechanics

t_{ij} : setup time required for task i by mechanic j

x_{ij} : binary variable equal to 1 if task i is assigned to mechanic j

Objective Function:

$$\text{Minimize } Z = \sum_{i,j=1}^n t_{ij} x_{ij}$$

Constraints:

Each task must be assigned to exactly one mechanic

$$\sum_{j=1}^m x_{ij} = 1$$

No mechanic can handle overlapping tasks

$$x_{ik} + x_{kj} \leq 1$$

Binary decision variable

$$x_{ij} \in \{0,1\}$$

This formulation satisfies the requirement to include the objective function, constraints, and equations based on variables.

4 RESULTS AND DISCUSSION

4.1 Descriptive and Statistical Analysis

The statistical results showed substantial variation in setup times across different machine operation combinations. Setup durations ranged from 12.29 to 176.00 minutes, while the large standard deviations (1.50 to 151.97 minutes) revealed considerable inconsistency in the setup process. After applying the 15% allowance factor, standardized setup times were calculated to ensure consistency when constructing the ILP model. These results confirm that the current system lacks uniformity and highlight the need for an optimized, data-driven allocation to reduce delays and improve overall efficiency.

4.2 Optimisation of Results

In addition to descriptive statistics, comparative insights were used to understand variations in mechanical

performance and operation complexity. These statistical findings served as the foundation for developing the ILP model. By using standardized setup times as model inputs, the ILP formulation ensures that the optimization results are based on observed operational patterns. This integration of statistical evidence with optimization enhances the reliability and accuracy of the mechanic-task allocation produced by the model.

The Integer Linear Programming model generated an optimal mechanic-task allocation, reducing total setup time from 5,534.68 minutes to 4,535.07 minutes and 18% improvement. This aligns with Lean and SMED principles, showing that combining data analysis with statistical optimization can enhance industrial performance. Table 1 shows the optimized workload distribution generated by the ILP model. The assignment clearly demonstrates how tasks were allocated evenly among mechanics to avoid overload and reduce idle time. By distributing tasks more fairly, the model ensures balanced resource utilization, minimizes bottlenecks, and creates a smoother workflow during style changeovers. This balanced allocation is a key factor behind the achieved 18% reduction in total setup time.

4.3 Practical Implications

This study provides practical benefits for apparel manufacturers to reduce machine idle time and improve line efficiency.

Table 1: Balanced workload distribution

Operation	Mechanic	Setup time(min)	Operation	Mechanic	Setup time(min)
Waist Binding	Mechanic 2	60	Attach Waist Lace	Mechanic 6	18
Elastics Attach	Mechanic 1	68	Floating Gusset	Mechanic 3	34
Leg Elastic Attach	Mechanic 2	36	Inseam	Mechanic 5	25
JIG BT	Mechanic 3	42	Side Gusset	Mechanic 1	51
Mobilion Tape Attach	Mechanic 6	45	Back Empire Seam	Mechanic 6	58
Side Seam BT	Mechanic 4	20	Front Pouch	Mechanic 5	35
Waist Attach	Mechanic 1	39	Bagout Gusset	Mechanic 7	27
Attach Waist Elastic	Mechanic 3	34	Pouch Attach	Mechanic 1	35
Waist Top	Mechanic 5	48	Side Seam	Mechanic 3	23
Attach Waist Band	Mechanic 6	74	Back Gusset	Mechanic 7	32
Leg Binding Pullout	Mechanic 4	48	Bow/Trim/Charming Attach	Mechanic 6	13
Pullout Side Seam	Mechanic 4	35	Box Tack	Mechanic 4	31
Waist Binding Pullout	Mechanic 1	61	Double BT	Mechanic 2	23
Leg Binding	Mechanic 6	51	JIG BT	Mechanic 1	3
Leg Top	Mechanic 5	15	Label BT	Mechanic 1	35
Lining OL	Mechanic 2	12	Pullout BT	Mechanic 1	31
Back Rise	Mechanic 2	15	Ring Attach	Mechanic 1	40
Leg Attach	Mechanic 3	22	Front Binding	Mechanic 1	35
Waist Attach	Mechanic 4	30	Tack	Mechanic 5	26
Binding Attach Leg	Mechanic 4	5	Lap Seam	Mechanic 5	25
Keyhole Binding	Mechanic 7	39	Logo Attach	Mechanic 1	20
Waist Attach	Mechanic 3	46	Lace Attach	Mechanic 6	22
Attach Wasit Band	Mechanic 4	24	Leg Lace Atach	Mechanic 6	44
Leg Hem	Mechanic 3	45	Panel Joint	Mechanic 3	19
Chain Stitch	Mechanic 7	35	Panel Tack	Mechanic 2	34

Operation	Mechanic	Setup time(min)	Operation	Mechanic	Setup time(min)
Gusset Top / Leg Top	Mechanic 2	28	Label Attach	Mechanic 7	14
Waist Top	Mechanic 2	35	JIG	Mechanic 7	135

5. CONCLUSION

This study successfully demonstrated how statistical and optimization methods can be applied to minimize machine setup time in apparel manufacturing. By analyzing real production data and applying ILP modelling, the research achieved an 18% reduction in total setup duration.

The findings reinforce the value of Lean Manufacturing principles in modern apparel operations, offering practical insights into how setup management can be transformed from an experience-based task to a data-driven, optimized process.

Future research may focus on integrating simulation models or machine learning algorithms to predict setup times dynamically, further enhancing real-time decision-making in apparel production environments.

- Mohamed, H.M.H. (2014). Study on setup time effect on small order quantities in apparel manufacturing [Master's theses, University of Moratuwa]. Institutional Repository University of Moratuwa.
- Shingo, S. (2019). A revolution in manufacturing: the SMED system. Routledge.
- Tahar, D. N., Yalaoui, F., Chu, C., & Amodeo, L. (2006). A linear programming approach for identical parallel machine scheduling with job splitting and sequence dependent setup times. *International journal of production economics*, 99(1-2), 63-73.

REFERENCES

- Ali, S. M., Haque, M. N., Sarker, M. R., Kandasamy, J., & Vlachos, I. (2025). Improving changeover techniques in apparel manufacturing through single minute exchange of die: implications for benchmarking. *Benchmarking: An International Journal*, 32(4), 1207-1223.
- Antonioli, I., Guariente, P., Pereira, T., Ferreira, L. P., & Silva, F. J. G. (2017). Standardization and optimization of an automotive components production line. *Procedia Manufacturing*, 13, 1120-1127.
- Karim, R. (2013). Impact of Changeover time on productivity: A case study. *Int. J. Eng. Technol. IJET-IJENS*, 13, 46-52.



Identifying the Affecting Factors and Predicting Customer Satisfaction in Online Food Delivery Services in North Western Province Using Machine Learning Techniques

Goonerathne GDN¹, Lakshmi MPDS¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
devninatasha123@gmail.com

ABSTRACT

Online food delivery (OFD) services have rapidly grown in Sri Lanka due to the increased use of mobile technology and demand for convenient food accessibility. However, maintaining customer satisfaction remains a challenge due to multiple factors such as food quality, delivery time, usability, pricing, service responsiveness, brand image, and trust influence user experience. Despite this growing industry, limited studies have used machine learning techniques to analyze and predict customer satisfaction, especially in the North Western Province. Therefore, this study aims to identify the key determinants affecting customer satisfaction and develop an accurate predictive model using machine learning. A structured questionnaire was distributed online, collecting responses from 1,022 active OFD users in the North Western Province, measured on a five-point Likert scale. A Random Forest Classifier was used for prediction and was benchmarked against Logistic Regression to ensure model validity. Results revealed that the Random Forest model achieved the highest accuracy at 75.7%, demonstrating its strength in capturing nonlinear relationships. Feature importance analysis showed that food quality, delivery time, and usability are the most influential predictors of satisfaction. The findings highlight the potential of machine learning to support data-driven decisions in enhancing customer service attributes, improving satisfaction, and increasing loyalty in Sri Lanka's competitive OFD sector.

KEYWORDS: Consumer behavior, Customer satisfaction, Machine learning, Online food delivery, Random Forest Classifier

1. INTRODUCTION

The online food delivery (OFD) industry has reshaped how consumers access food, offering unprecedented convenience and flexibility (Zulkarnain et al., 2015). In Sri Lanka, rapid technological advancement, mobile application use, and urbanization have accelerated the adoption of OFD services (Wijesinghe et al., 2023). With growing competition, understanding customer satisfaction has become essential for service providers to maintain market share and foster loyalty (Al-Tit, 2015).

Customer satisfaction in OFD services is influenced by multiple factors, including food quality, delivery speed, usability, cost, customer service, brand image, and trust (Jayasinghe & Dayangana, 2022; Kavindi & Dissanayake, 2023). However, traditional statistical methods such as correlation and regression analyses often fail to capture complex, nonlinear relationships among these factors (Prasetyo et al., 2021). To address this limitation, this study applies a machine learning approach to predict customer satisfaction based on multiple service attributes (The American Journal of

Engineering and Technology, 2024). The main objectives of the study are to identify the factors influencing customer satisfaction in OFD services, develop a predictive model, and evaluate the model's performance in classifying satisfaction levels. The study contributes to both academia and industry by introducing data-driven modeling to improve decision-making in the Sri Lankan OFD sector (Ayoobkhan & Haleem, 2020).

2. LITERATURE REVIEW

Several machine learning algorithms, such as Logistic Regression, Support Vector Machines (SVM), Decision Trees, and K-Nearest Neighbors (KNN), have been commonly used for classification problems in customer satisfaction prediction. These models serve as important benchmarks for evaluating the performance of more advanced ensemble techniques.

Recent studies have identified various determinants of satisfaction in OFD contexts. Kavindi and Dissanayake (2023) found that convenience, delivery time, and brand image significantly affected customer satisfaction in the Kurunegala District.

Jayasinghe and Dayangana (2022) highlighted meal quality, service fulfillment, and usability as key drivers of satisfaction in Colombo. Similarly, Ayoobkhan and Haleem (2020) emphasized the role of convenience and food quality, particularly during the COVID-19 pandemic. International research also supports these findings.

Prasetyo et al. (2021) identified price, promotion, and information quality as influential factors during the pandemic, while Tan and Kim (2021) confirmed that service quality and food quality are vital to consumer satisfaction. Despite extensive literature, most studies employ traditional statistical methods. Few have explored the use of advanced machine learning models, for satisfaction prediction particularly within the Sri Lankan context. This research fills that gap by applying Machine Learning techniques to predict satisfaction categories and identify the most influential factors driving customer experience.

3. METHODOLOGY

3.1 Research Design

A quantitative approach was used to identify and predict factors influencing customer satisfaction in online food delivery services, using a structured questionnaire distributed among active users in the North Western Province.

3.2 Data Collection and Sampling

Data were collected through an online survey distributed via social media and email from 1,022 OFD users aged 18–50 in the North Western Province. Using Yamane's formula (Yamane, T., 1973) with a 5% margin of error, the minimum required sample was 400, but 1,022 valid responses were collected, increasing reliability. To minimize sampling bias, the survey was distributed widely via online platforms targeting diverse user groups, and a stratified 80/20 train-test split was used to maintain category representation. Measurement bias was reduced through a validated questionnaire based on a 5-point

Likert scale, and internal consistency was confirmed by a Cronbach's Alpha value of 0.7249.

Although convenience sampling was used, various demographic segments were included to reduce selection bias, and statistical reliability measures were applied to mitigate measurement bias.

3.3 Reliability and Validity Testing

Cronbach's Alpha value of 0.7249 confirmed acceptable reliability, while Spearman's correlation analysis identified the strength and direction of relationships between the independent and dependent variables.

3.4 Data Preparation and Splitting

Before modeling, the dataset was cleaned and encoded appropriately for machine learning analysis and split into training (80%) and testing (20%) sets using stratified sampling to preserve satisfaction category proportions.

3.5 Model Development

In this study, four machine learning models Logistic Regression, Support Vector Machine (SVM), Decision Tree, and Random Forest were developed to compare predictive performance. Logistic Regression was used as the benchmark model to establish a baseline for evaluating improvements in predictive accuracy. Although SVM and Decision Tree were included to capture more complex relationships, both models showed limitations when dealing with class imbalance and nonlinear patterns. The Random Forest Classifier was ultimately selected due to its strong

ability to model categorical outcomes and handle complex, nonlinear interactions. Trained using 100 bootstrapped decision trees and majority voting, Random Forest effectively reduced overfitting and achieved the highest accuracy among all tested models, making it the most reliable classifier for predicting customer satisfaction.

3.6 Model Evaluation Metrics

The model's performance was evaluated using accuracy, precision, recall, and F1-score, supported by a confusion matrix to visualize correct and incorrect classifications across satisfaction levels. Feature importance values were also calculated to identify the most influential predictors of customer satisfaction. In addition, ROC curve analysis was incorporated to provide a more reliable measure of the model's discriminatory ability, as relying solely on the confusion matrix can be sensitive to overfitting.

4. RESULTS AND DISCUSSION

4.1 Descriptive Analysis

Demographic analysis revealed that 58.9% of respondents were male and 59.4% belonged to the 21–30 age group, indicating a predominantly young, tech-savvy customer base. Pizza Hut, Pick Me Food, and Uber Eats were the most frequently used platforms. The majority of respondents ordered monthly, with average spending between Rs. 2,000 and 3,000 per transaction.

Overall satisfaction levels were high, with 61.8% of users "Satisfied" and 26.2% "Very Satisfied." Among individual factors, satisfaction was highest for food

quality (87.7%), usability (86.5%), and delivery time (84.2%), while cost and pricing scored lower (71.6%). These findings suggest that operational and product-related factors dominate customer satisfaction outcomes.

4.2 Model Performance

The performance of the Random Forest model was further validated by comparing it with Logistic Regression, which served as the baseline classifier. As shown in Table 2, Random Forest achieved stronger precision and recall for the “Satisfied” and “Very Satisfied” categories, with an F1-score of 0.81 for the “Satisfied” class, although it performed less accurately for the “Dissatisfied” and “Neutral” groups due to class imbalance. The Random Forest model also achieved the highest overall accuracy of 75.7%, outperforming Logistic Regression by better capturing nonlinear relationships. In addition, Receiver Operating Characteristic (ROC) curves and Area Under Curve (AUC) metrics were computed, and Random Forest demonstrated superior AUC values across major satisfaction levels, confirming its stronger discriminatory ability and better generalization compared to the baseline model (Fawcett, 2006).

Table 1 : Accuracy Comparison between Classification models

Model	Accuracy
Random Forest	75.7%
SVM	72.1%
Decision Tree	69.8%

Logistic Regression 68.4%

Table 2 : Summary of Precision, Recall, and F1 Score metrics

Class	Precision	Recall	F1 Score
Dissatisfied	0	0	0
Neutral	0.43	0.19	0.26
Satisfied	0.74	0.91	0.81
Very Satisfied	0.75	0.53	0.62

4.3 Feature Importance

Table 3 shows that the feature importance analysis indicated that food quality, delivery time, and usability were the top predictors of customer satisfaction. Cost and pricing, service quality, brand image, and trust had moderate influence.

Table 3 : Feature Importance Scores for Customer Satisfaction Prediction

Variable	Importance
Gender	0.38547
Age	0.41216
Education	0.3565
Income	0.22915
Usage	0.063805
Platform	-0.24613
Amount	0.097056
Quality	1.7285
Time	0.60612
Usability	0.94657
Cost	0.58982
Service	1.3802
Brand	1.4888
Trust	1.6461

4.4 Discussion

The Random Forest model outperformed Logistic Regression, showing that ensemble methods better capture nonlinear service factor relationships. Despite class imbalance, improving food quality, delivery, and pricing can enhance customer experience and loyalty. In addition, ROC-AUC evaluation confirmed the superior discriminatory power of the Random Forest model compared to the baseline Logistic Regression classifier, thereby strengthening the reliability of the predictive findings.

5. CONCLUSION

The study shows that the Random Forest classifier achieved 75.7% accuracy and outperformed Logistic Regression in predicting customer satisfaction in Sri Lanka's North Western Province. Food quality, delivery time, and usability were identified as the most influential factors, highlighting the importance of maintaining high service standards, ensuring timely deliveries, and improving app usability. Model evaluation using ROC-AUC further confirmed the superior performance of the Random Forest approach. Although class imbalance affected the prediction of lower satisfaction levels, future studies can improve results by applying resampling or advanced ensemble techniques and extending the analysis to other regions to enhance model generalizability and decision-making in the online food delivery sector.

REFERENCES

- Al-Tit, A. A. (2015). The effect of service and food quality on customer satisfaction and hence customer retention. *Asian social science*, 11(23), 129.
- Ayoobkhan, A. L. M., & Haleem, A. (2020). Online food ordering and delivery application during Covid 19 in Sri Lanka.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, 27(8), 861-874.
- Jayasingha, W. M. M. W. G., & Dayangana, K. T. L. U. S. (2022). A study on consumer satisfaction on online food delivery services in the colombo district.
- Kavindi, L. P. V. H., & Dissanayake, D. M. T. D. (2023, December). Factors affecting customer satisfaction towards online food ordering applications: reference to Kurunegala district. In *6th Annual Research Symposium in Management 2023* (p. 129).
- Prasetyo, Y. T., Tanto, H., Mariyanto, M., Hanjaya, C., Young, M. N., Persada, S. F., ... & Redi, A. A. N. P. (2021). Factors affecting customer satisfaction and loyalty in online food delivery service during the COVID-19 pandemic: Its relation with open innovation. *Journal of open innovation: technology, market, and complexity*, 7(1), 76.
- Wijesinghe, P., Rathnayaka, S. D., Bandara, N., Jayasena, D. D., & Heo, J. M. (2023). Analysis of online food purchasing behavior: a study of Sri Lankan consumers. *Korean Journal of Agricultural Science*.
- Yamane, T. (1973). Statistics: An introductory analysis.
- Zulkarnain, K., Ahasanul, H., & Selim, A. (2015). Key success factors of online food ordering services: An empirical study. *Malaysian Institute of Management*, 50(2), 19-36.

Modeling AUD/LKR Exchange Rate Volatility Using Time Series and Machine Learning Models

O.A.M. Jayawardana¹ G.A.C.M.Karunanada¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
amjmadusanka@gmail.com

ABSTRACT

Exchange rate volatility significantly affects trade, education costs, and financial planning in Sri Lanka. This research aims to forecast volatility in the AUD/LKR exchange rate using daily data from the Central Bank of Sri Lanka, spanning from January 2014 to March 2025. The analysis integrates traditional econometric models alongside machine learning and deep learning approaches, specifically employing ARMA, ARCH, and GARCH models to assess linear and varying conditional variance. The findings indicate that the GARCH(1,5) model effectively captures volatility clustering, with a Mean Absolute Error (MAE) of 0.6001 and a Root Mean Square (RMS) error of 0.8683. Although the R^2 value for this model is relatively low at 0.002, it establishes a foundational understanding of volatility. In contrast, the machine learning techniques reveal stronger performance metrics. The Random Forest model particularly excels, yielding an MAE of 0.9621 and an RMS error of 1.4770, with a striking R^2 of 0.9985. When exploring deep learning models, the LSTM model produces competitive results, with an MAE of 1.1560, an RMS error of 1.8855, and an R^2 value of 0.9863. In addition, the RNN model has an R^2 value of 0.9656, with an MAE of 2.0212 and an RMS error of 2.9846. These findings highlight the potential of machine learning and deep learning methods for enhancing exchange rate volatility forecasting.

KEYWORDS: Deep Learning, Exchange Rate, GARCH, Machine Learning

1 INTRODUCTION

The changes that occur in the rate of exchange for the AUD and LKR have become more prominent because of the growing educational, business, and trade links between these two nations. Unexpected changes occur in the rate of exchange, causing imbalances in student expenses, remittances, and trade. This research compares and analyzes the forecasting accuracy of traditional econometric models of ARMA, ARCH, and GARCH with more advanced models as RF, XGB, RNN, and LSTM for understanding which method can be more stable for modeling exchange rate volatility. During the analysis and research process, data gathered from the Central Bank of Sri Lanka included dates

ranging from January 2014 to March 2025.

2 LITERATURE REVIEW

2.1 Econometric Modeling

GARCH models generalize the ARCH model (Engle, 1982) to include volatility clustering, persistence, and asymmetry in Sri Lankan exchange rates. As Bollerslev (1986) shows, these GARCH models generalize the family of ARCH models and have been useful for capturing the variance and non-normal components in LKR return series. As seen from various sources on empirical research conducted on Sri Lankan economies, these GARCH-family models outperform traditional methods for modeling IMF exchange rate volatility

among principal international currencies (Perera and Rathnayaka, 2020), stochastic processes for analysis of behavior (Perera and Rathnayake, 2015), and identification for probability distribution functions (Basnayake and Chandrasekara, 2024).

2.2 Machine Learning Models

For nonlinear exchange rate dynamics and regime dependence, machine learning models can outperform traditional econometric models, especially neural networks. In Sri Lanka, several studies confirm the evidence for non-stationarity and volatility clustering of exchange rates (Perera & Rathnayake, 2015; Basnayake & Chandrasekara, 2024). These facts give validity for GARCH-family models, namely ARCH-Engle (1982) and GARCH-Bollerslev (1986), for modeling variance (Perera & Rathnayaka, 2020). Neural networks along with time series approaches work well while forecasting the rate, as per Nanayakkara et al. (2014). Monetary policy transmission and its impact on stock returns also have shown that the sensitivity of exchange rates persists.

3 METHODOLOGY

The research architecture blends econometric, machine learning, and deep learning methodologies to fit and forecast AUD/LKR exchange rate volatility. It involves machine learning model specification, data preprocessing, training, testing, and prediction. ARCH and GARCH models detect volatility clustering and time-varying variance and provide statistical insights. Machine learning models can learn nonlinear interactions, while deep learning models RNN and LSTM learn sequential and long-term dependencies. This technique

offers econometric interpretability with the forecasting capability of data-based models to make plausible forecasts of sound volatility.

3.1 Traditional Time Series Models

Traditional time series models were used to describe the volatility of the AUD/LKR exchange rate. The ARMA model accounted for linear relationships by predicting current values based on historical observations and errors in forecasts. ARCH and GARCH models addressed volatility clustering and time-varying conditional volatility in returns in finance. Differencing was applied to ensure stationarity, and diagnostic checks including residual analysis, autocorrelation tests, and normality assessments were used to verify the validity of the models. This time series approach provided interpretable insights into the conditional variance dynamics of the exchange rate.

3.2 Machine Learning Models

In an attempt to identify complex interactions and nonlinear relationships within the AUD/LKR exchange rate data, machine learning algorithms including Random Forest and Extreme Gradient Boosting (XGBoost) were used. The available dataset was divided according to the 80/20 rule, where 80% of the observations were used to train the models and the remaining 20% were reserved for validation. Using bootstrapped samples and random feature subsets, the Random Forest ensemble learning algorithm builds an ensemble of several decision trees.

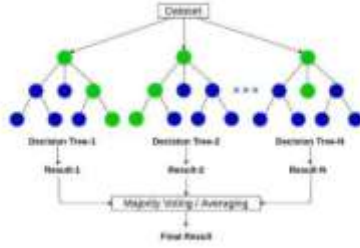


Figure 1: Random Forest Architecture

In order to avoid overfitting and increase robustness, the final prediction is then made by averaging the results of every tree.

$$\hat{y}_t^{(t)} = \hat{y}_t^{(t)} + \eta f_t(x_i) \quad (1)$$

where η is the learning rate and f_t represents the new tree added at iteration t .

Feature engineering was essential in improving model performance through the use of rolling averages, lag variables, and categorical indicators. The models were able to effectively capture domain-specific patterns and temporal dependencies in the exchange rate series thanks to these features.

3.3 Deep Learning Models

To extract the sequential pattern of the AUD/LKR exchange rate data, as well as temporal dependencies and long-term patterns, deep learning networks including Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks were used.

RNNs process input sequences by maintaining and updating hidden states that carry information between time steps so that the models can retain previous inputs in their outputs, which is suitable for time series tasks.

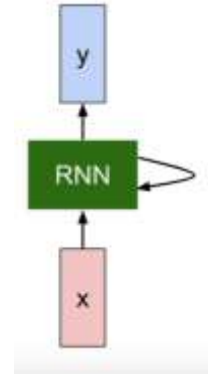


Figure 2 :RNN Architecture

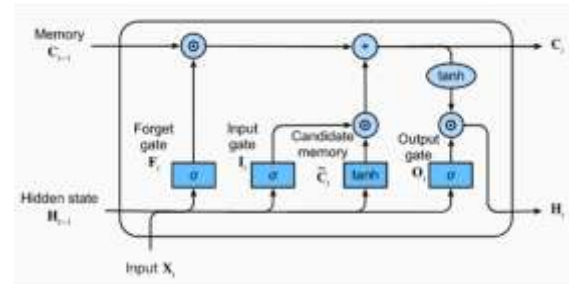


Figure 3 :LSTM Neuron Architecture

LSTM networks augment RNNs with explicit gating mechanisms that regulate information flow selectively forgetting, updating, and outputting information enabling better learning of long term dependencies without vanishing gradient issue.

Input characteristics were normalized by applying scaling techniques before training to impart training stability and facilitate faster convergence. Iterative sequence processing was used to generate predictions, recycling the output through the network for multi-step forecasting.

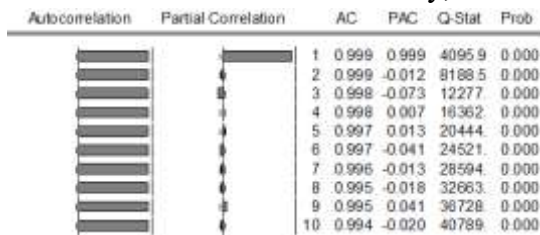
4 RESULTS AND DISCUSSION

The study compared traditional time series, machine learning, and deep learning models to forecast exchange rate volatility. Initial exploration showed non-normality and volatility clustering. The Augmented Dickey-Fuller test confirmed

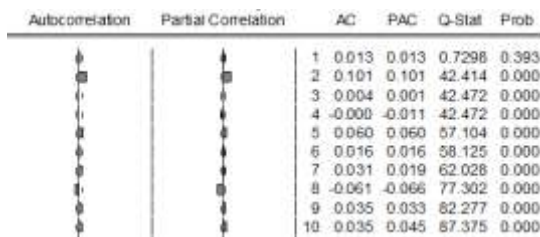
Table 1: Model Comparison

Model	MAE	RMSE	R ²
GARCH (1,5)	0.6001	0.8683	0.002
Random Forest	0.9621	1.4770	0.9989
LSTM	1.1560	1.8855	0.9863
XGBoost	1.1691	1.7319	0.9985
RNN	2.0212	2.9846	0.9656

non-stationarity, so the series was differenced to achieve stationarity,

**Figure 4 :**ACF & PACF Raw data**Table 2 :**ADF Test Raw data

	t-statistic	Prob.
ADF Test Statistic	0.554733	0.8360

**Figure 5 :**ACF & PACF Diff data**Table 3 :**ADF Test Diff data

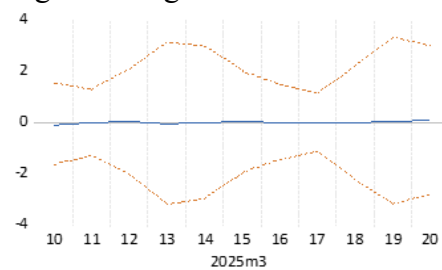
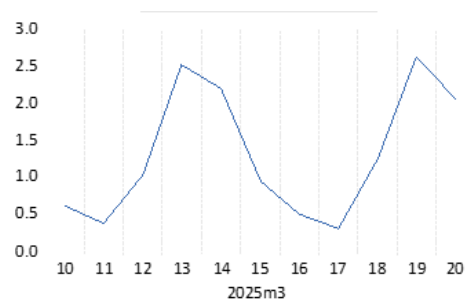
	t-statistic	Prob.
ADF Test Statistic	-40.59765	0.0000

enabling ARMA-GARCH models to capture time-varying variance. Random Forest excelled with superior accuracy and interpretability, outperforming XGBoost, LSTM, and RNN models. GARCH modeled variance well but struggled with mean predictions; deep

learning captured sequences effectively but lagged in overall accuracy.

Sequential dependence but poorly in overall accuracy and variance explanation.

GARCH models performed extremely well in modeling the evolution of variance but poorly in predicting average exchange rate values.

**Figure 6 :**GARCH(1,5) Mean Forecast**Figure 7 :**GARCH(1,5) Variance Forecast

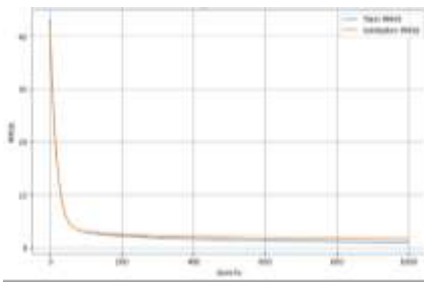


Figure 8 :Random Forest Training vs Validation Loss

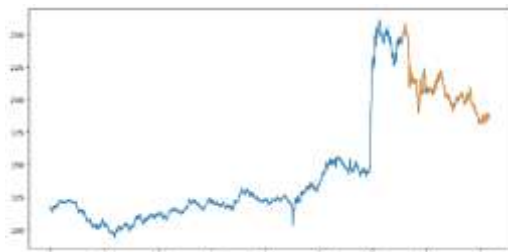


Figure 9 :Random Forest Actual vs Predicted

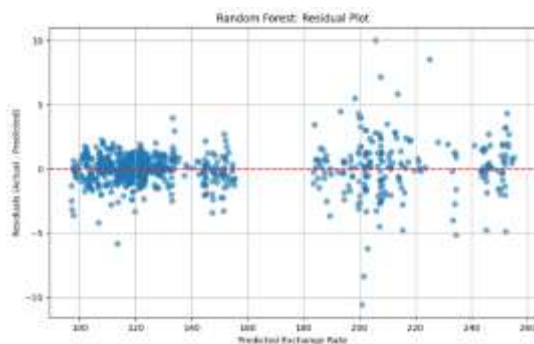


Figure 10 :Random Forest Residuals

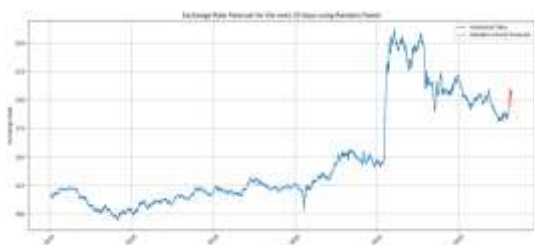


Figure 11 :Random Forest Forecast

Deep learning models performed well in recognizing sequential dependence but poorly in overall accuracy and variance explanation. This study highlights the relevance and benefits of using models with interpretability as well as high predictive

ability in providing effective forecasting of exchange rate volatility in financial markets.

5 CONCLUSION

This research undertook a thorough examination of a variety of models to forecast AUD/LKR exchange rate volatility, from standard time series, machine learning, and deep learning methods. Preliminary data analysis displayed extreme non-normality and volatility clustering, the use of sophisticated models. Stationarity was verified by differencing, allowing ARMA-GARCH models to capture various variance well. Machine learning algorithms as Random Forest and XGBoost employed heavy feature engineering to detect weak nonlinear patterns. Deep learning algorithms as LSTM and RNN accurately picked up the patterns over time but were slightly worse in prediction accuracy than ensembling algorithms. Among them all, Random Forest worked the best with highest R^2 value being 0.9989 and with lowest error measures showing high explanatory power as well as high predictability. GARCH models were best for volatility modeling but not so good for predicting average exchange rate values. The research highlights interpretability as well as accuracy and indicates that further research may investigate hybrid models and other economic indicators.

REFERENCES

- Bandara, A. W. M. S. S. C., & Thilakarathne, P. M. C. (2020). The impact of exchange rate movements on stock returns: Evidence from commercial banks in Sri Lanka. *SSRN*.

- Basnayake, B., & Chandrasekara, N. (2024). Identifying the probability distribution of exchange rates in Sri Lanka. *The Indian Economic Journal*, 72(1), 12–25.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.
- Musthafa, M. T., Le, T., & Suardi, S. (2023). Monetary policy transmission in Sri Lanka. *Applied Economics*, 56(2), 151–168.
- Nanayakkara, S., Chandrasekara, V., & Jayasundara, M. (2014). Forecasting exchange rates using time series and neural network approaches. *European International Journal of Science and Technology*, 3(2), 2–10.^[2]
- Perera, A., & Rathnayaka, R. (2020). Modeling exchange rate volatility using GARCH models: Empirical analysis from five major currencies in Sri Lanka. *International Journal of Applied Research*, 6(3), 283–291.^[3]
- Perera, S., & Rathnayake, R. M. (2015). An empirical analysis of stochastic behavior of Sri Lanka exchange rate changes. *Journal of Management*, 12(2), 62–75.^[3]
- Thevakumar, P., & Jayathilaka, R. (2022). Exchange rate sensitivity influencing the economy: The case of Sri Lanka. *PLOS ONE*, 17(6), e0269538.



Hybrid ARIMA and Machine Learning Models for Daily UAX/LKR Gold Exchange Rates Forecasting

Chamodya NDS

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
samudichamodya@gmail.com

ABSTRACT

Accurate gold exchange rate forecasting is important for policymakers, banks, and investors. The trend of gold prices has linear patterns as well as nonlinear patterns based on the global economic indicators, foreign exchange rates, and market volatility. This study investigates the forecasting of daily gold prices through integrating a hybrid Autoregressive Integrated Moving Average (ARIMA) and machine learning approaches. Specifically, an ARIMA(2,1,2) model was used to capture the linear patterns in gold price movements, while machine learning models such as Random Forest (RF), Extreme Gradient Boosting (XGBoost), and CatBoost were applied to the residuals of the ARIMA to capture non-linear dependencies. The dataset includes daily gold exchange rates (UAX/LKR) over the period from 2014 to 2025, with 80% of the data used for training and 20% for testing. Model performance was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). Results indicate that the ARIMA + RF hybrid model outperforms the other hybrid models, combining linear and non-linear patterns effectively, achieving the lowest MAE, RMSE, and MAPE on the test data, indicating improved predictive capability. The findings confirm that hybrid ARIMA-RF modeling is a practical approach for accurate short-term forecasting of gold prices in highly volatile markets.

KEYWORDS: ARIMA, CatBoost, Forecasting, Random Forest, XGBoost

1 INTRODUCTION

Gold is a highly traded item whose price is influenced by economic, geopolitical, and financial forces. Accurate forecasting of the gold price is needed for investors, banks, and policymakers to allow them to make effective risk management, portfolio optimization, and hedging policy decisions. Traditional time series methods such as Autoregressive Integrated Moving Average (ARIMA) are widely used in forecasting because they can capture linear trends and seasonality. However, gold prices often exhibit non-linear patterns as a result of market shocks, necessitating more sophisticated modeling methods.

Machine learning techniques such as Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Categorical Boosting (CatBoost) offer adaptive models for non-linear relationships in financial time series. Recent studies have demonstrated the benefit of hybrid models combining classical time series approaches and machine learning for improved predictive performance and robustness. The dataset used in this study consists of daily gold exchange rates covering the period from 2014 to 2025, making for a complete representation of both short-term volatility and long-term market trends. The longer horizon is most appropriate for the measurement of the robustness and longevity of hybrid forecasting

models since it captures periods of economic stability as well as instability.

Despite such advancements, no detailed study has been performed comparing the different hybrid models-ARIMA with RF, XGBoost, and CatBoost to forecast daily UAX/LKR gold prices, particularly using long-term data. This study aims to fill this gap by evaluating the forecasting power of the listed hybrid models and furnishing precise short-term forecasts for decision-making in highly volatile gold prices.

2 LITERATURE REVIEW

Gold price forecasting remains a challenging task in light of the coexistence of linear trends and nonlinear market forces driven by macroeconomic and financial instability. The studies here collectively support the hypothesis that hybrid models integrating statistical and machine-learning methods outperform individual approaches in capturing these forces.

Ye (2025) suggests a Long Short-Term Memory (LSTM) –ARIMA hybrid model for predicting daily gold prices and demonstrates its superiority in a large-scale empirical investigation using 2013–2024 data. The model first trains an LSTM to learn nonlinear temporal dependencies and then applies an ARIMA model on residuals to learn any leftover linear trends. Results show significant improvements in the form of MAE (84.9% decrease), RMSE (82.1% decrease), and MAPE (84.6% decrease), and R^2 improvements from 0.596 to 0.987. Mun et al. (2024) also contrast the performance of ARIMA, LSTM, and ARIMA-LSTM models and find the hybrid to possess the lowest RMSE,

although pure LSTM yields slightly better MAPE. They argue that ARIMA models linear short-term dependencies and the neural network captures intricate nonlinearities in residuals-rationalizing the complementary structure underlying hybrid design. From a broader econometric perspective, Zheng (2025) develops a hybrid ARIMA-Support Vector Regression (SVR) linear-mixed model, demonstrating the outperformance of individual ARIMA and SVR models on international gold prices. The hybrid recorded the lowest RMSE (20.20), MAE (14.61), and MAPE (0.64%), confirming the fact that the combination of linear and nonlinear components enhances stability and risk-adjusted forecast ability. To these, Bunnag (2023) identifies hybridizations like ARIMA-Generalized Autoregressive Conditional Heteroskedasticity (GARCH) and ARIMA-Threshold Generalized Autoregressive Conditional Heteroskedasticity (TGARCH), which involve volatility modeling in order to control for heteroscedasticity and asymmetric shocks in gold markets.

These models outperform naive ARIMA by jointly estimating trend and volatility, which points to the advantage of joint methods for highly volatile assets.

Finally, Rashidi and Modarres (2025) propose an ARIMA + stepwise-regression + Artificial Neural Network (ANN) hybrid with $\approx 99\%$ accuracy, stressing the value of variable selection and multi-step modeling to model remaining non-linearities.

Despite extensive studies on hybrid ARIMA-LSTM and ARIMA-SVR models, limited research has compared ARIMA with ensemble-based machine

learning methods such as RF, XGBoost, and CatBoost for long-term daily gold price forecasting. This study addresses that gap by developing and evaluating these hybrid models on a decade-long dataset for UAX/LKR gold prices.

3 METHODOLOGY

In this study, daily UAX/LKR gold prices for one Troy Ounce in US dollars at the beginning of each business day (from Reuters or Bloomberg) and the indicative US Dollar SPOT Exchange Rate (LKR per 1 USD) from 2014 to 2025 were collected, giving a total of 2,654 data points arranged by date, and no missing values were found. The data was split into two parts 80% for training the models and 20% for testing. The lag features were normalized before training using Min-Max scaling to enhance model stability and convergence.

An ARIMA model was initially applied to capture the linear patterns in gold price movements. The ARIMA model was selected based on the Akaike Information Criterion (AIC) and stationarity testing using the Augmented Dickey-Fuller (ADF) test. Differencing was applied to achieve stationarity, allowing the ARIMA model to effectively capture short-term autocorrelation and overall linear trends, providing a baseline forecast.

To account for non-linear patterns, the residuals from the ARIMA forecasts were used as inputs for three hybrid models: ARIMA with RF, XGBoost, and CatBoost. These hybrid approaches leverage the linear modeling capability of ARIMA and the non-linear learning capacity of ensemble machine learning methods, enhancing overall

forecasting accuracy. RF, XGB, and CatBoost were selected to complement ARIMA's linear predictions by capturing non-linear patterns in the residuals. RF is robust to overfitting and models complex interactions, XGB effectively captures sequential dependencies and subtle trends, and CatBoost reduces prediction bias while handling non-linear relationships efficiently. Together, these models enhance the overall forecasting accuracy of the hybrid approach. Lag features of 1, 2, 5, 7, 14, and 21 days were considered. Lags 1 and 2 represent immediate autocorrelation for short-term movements, lags 5 and 7 capture weekly trading behavior and weekend effects, and lags 14 and 21 represent medium-term cycles influenced by investor rebalancing and macroeconomic announcements. These lag values were chosen based on typical patterns observed in daily financial time series. This balances model complexity with predictive performance, consistent with practices in forecasting literature. Each model's performance was measured using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) on the test data set. The machine learning models predicted gold prices stepwise using past residuals, and the best-performing model was selected based on the lowest MAE, RMSE, and MAPE. The final hybrid forecast was produced by combining the predictions of ARIMA with the best machine learning model.

4 RESULTS AND DISCUSSION

Table 1: Descriptive statistics of daily gold prices 2014-2025

Mean	SD	Min	Max
343578.7	200921.8	144312.3	815923.5

Table 1 presents the descriptive statistics of the daily gold prices from 2014 -2025. It shows the highest gold price is 815923.55 LKR and the minimum is 144312.31 LKR. As the time series plot indicates in Fig. 1, the dataset exhibits an upward trend over the period.

As the Table 2, the Augmented Dickey-Fuller (ADF) test confirms that the original series is not stationary (P-value > 0.05). After the first order differencing, the ADF test reveals that the differenced data are stationary (P-value < 0.05).

The best ARIMA model selected using the `auto_arma` function was ARIMA(2,1,2). The model's AIC value was 52,291.72. This indicates that ARIMA(2,1,2) provides the best fit all among the candidate models for predicting daily gold prices.

Table 3 reveals the performance of different models for forecasting daily gold prices and hybrid models improve forecasts by capturing both linear and nonlinear patterns. Hybrid ARIMA-RF shows the lowest MAE (5430.01), RMSE (7609.89), and MAPE (0.80%) and Hybrid ARIMA and XGB and Hybrid ARIMA and CatBoost achieve the MAE of 5889.38 and 5718.42, and RMSE of 8113.22 and 7878.34, respectively, while maintaining low MAPE values of 0.87% and 0.85% respectively. Overall, the hybrid models successfully combine linear and nonlinear patterns, with hybrid RF providing the best reduction in

forecast errors, making them more effective for capturing complex patterns in gold prices.

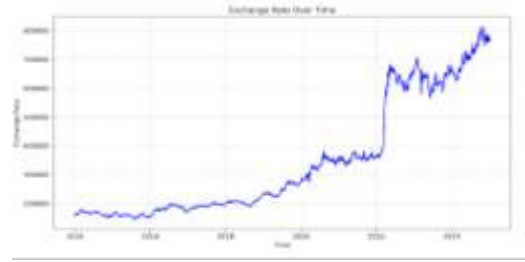


Figure 1: Time Series Plot

Table 2: Augmented Dickey-Fuller test

	original	differenced
p-value	0.964	4.41e-11

Table 3: Evaluation Results for hybrid ARIMA and Machine Learning models

Model	MAE	RMSE	MAPE
ARIMA+	5430.01	7609.89	0.80%
RF			
ARIMA+	5889.38	8113.22	0.87%
XGB			
ARIMA+	5718.42	7878.34	0.85%
CATBoost			



Figure 2: Hybrid ARIMA and Random Actual vs Forecast



Figure 3: Hybrid ARIMA and CatBoost Actual vs Forecast



Figure 4: Hybrid ARIMA and XGBoost Actual vs Forecast

According to Fig. 2, Fig. 3, and Fig. 4, hybrid forecasting models employing ARIMA combined with machine learning algorithms such as Random Forest, XGBoost, and CatBoost indicate a strong alignment between the forecasted and actual values of daily gold exchange rates. All the models appropriately capture the general trend and seasonality in the data, though minor deviations may be observed in instances of sharp changes. The results indicate that hybrid approaches enhance predictive accuracy by leveraging the time-series characteristics of ARIMA as well as the nonlinear learning capability of ensemble methods.

5 CONCLUSION

This study aims to enhance the accuracy of daily gold price forecasting by integrating linear and non-linear modeling techniques in three hybrid models, ARIMA with RF, XGBoost, and CatBoost. The ARIMA modeling was used to model the linear behavior of daily gold price movement, while machine learning models were based on ARIMA(2,1,2) residuals for modeling non-linear relationships influenced by global economic emotions and exchange rate fluctuations. The comparison using MAE, RMSE, and MAPE clearly indicates that the hybrid model ARIMA-

Random Forest received the best forecast, with the lowest MAE (5430.01), RMSE (7609.89), and MAPE (0.80%), outperforming the ARIMA-XGBoost and ARIMA-CatBoost models. The results confirm that Random Forest effectively captures the non-linear relationships. Consequently, hybrids of CatBoost and XGBoost also performed competitively well. This study concludes that hybrid ARIMA-machine learning models are powerful tools for daily gold price prediction in volatile financial markets. Among the three models, the most reliable and realistic hybrid is ARIMA-Random Forest, providing accurate predictions and robustness to policymakers, investors, and banks.

REFERENCES

- Bunnag, T. (2023, November 4). *The Importance of Gold and Forecasting of the Gold Price Using ARIMA, ARIMA-GARCH, and ARIMA-TGARCH Model*
- *Daily Gold Rates | Central Bank of Sri Lanka*. (n.d.). www.cbsl.gov.lk.
- Mun, W. K., Suliadi Firaus Sufahani, Kamil, A. A., & Mohd. (2025). *PREDICTION OF GOLD PRICES USING A HYBRID MODEL ARIMA-LSTM*. 92(5), 749–766.
- Rashidi, M., & Modarres, M. (2025). Predicting the Price of Gold in the Financial Markets Using Hybrid Models. *ArXiv.org*.
- Ye, Y. (2025). Improved Gold Price Prediction Based on the LSTM-ARIMA Hybrid Model. *Applied and Computational Engineering*, 165(1), 56–65.
- Zheng, H. (2025). Analysis of International Gold Prices Using ARIMA, SVR with Linear Regression. *Journal of Advanced Management Science*, 13(1).



Analyzing and Forecasting International Labour Migration in Sri Lanka

De Mel WRU¹, Aponsu GMLM¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
rasangiudayanademel@gmail.com

ABSTRACT

International nternational labour migration—the cross-border movement of workers for employment—is a vital driver of global economic interdependence and a major source of foreign exchange for Sri Lanka. This study analyzes Sri Lanka’s outward labour migration from 2014 to 2024, focusing on trends, gender dynamics, skill composition, and the influence of macroeconomic shocks including COVID-19 pandemic and the 2022 economic crisis. The research aims to forecast future migration flows, assess the impact of economic instability, and examine how gender and destination shape skill-level patterns among migrants. Secondary data from the Sri Lanka Bureau of Foreign Employment (SLBFE) were analyzed using three complementary methods: graphical analysis to identify temporal patterns, SARIMA time-series modelling for forecasting, and Ordinal Logistic Regression (OLR) to assess determinants of migrant skill levels. Results reveal that inflation and economic uncertainty were key push factors influencing migration decisions. Gender and host country significantly affected skill distribution, with males and migrants to Japan, Jordan, and Singapore showing higher skill levels. The SARIMA (0,1,0) (1,0,0)₁₂ model best forecasted total departures (MAPE = 5.82), while SARIMA (0,1,1) (1,0,0)₁₂ achieved the highest accuracy for female migration (MAPE = 8.71). The study highlights the importance of gender-responsive, skills-oriented and crisis-resilient migration policies that enhance resilience against future economic shocks.

KEYWORDS: Economic Crisis, Forecast, Labour, Migration, Ordinal Logistic Regression,

1 INTRODUCTION

International labour migration is a major contributor to Sri Lanka’s economy, with over two million citizens employed abroad and remittances serving as a critical source of foreign exchange. Migration patterns have become increasingly sensitive to economic shocks such as the COVID–19 pandemic and the 2022 economic crisis. To understand these shifts, this study incorporates key variables including gender, destination country, skill level, and the National Consumer Price Index (NCPI). NCPI is included as it captures inflation-driven cost-of-living pressures that significantly influence households’ decisions to migrate during periods of economic instability. While previous

studies have examined determinants of migration or macroeconomic trends separately, limited research integrates trend analysis, forecasting, and skill-level modelling. This study contributes by combining graphical analysis, SARIMA forecasting, and Ordinal Logistic Regression (OLR) to assess how inflation, gender disparities, and destination-specific labour demand shape migration patterns. It therefore provides a comprehensive understanding of Sri Lanka’s evolving migration landscape.

2 LITERATURE REVIEW

In Sri Lanka, international labour migration has been a significant phenomenon since economic liberation in the late 1970s, with the Middle East

serving as the primary destination. Temporary migration is high among young women from estate sectors, due to poverty and limited local job opportunities, while urban migrants are more driven by education and career goals (Dissanayake & Samarakoon, 2021). Macroeconomic pressures low GDP per capita, slow growth, unemployment, and inflation, push workers to migrate, while labour demand in destination countries, particularly in construction, healthcare, and domestic work, shapes migration patterns and skill levels (Kaluarachchi & Jayathilake, 2024; Dharmadasa & Herath, 2020; Wickramasinghe et al., 2022; Fernando & Perera, 2021).

Recent research has progressed from descriptive assessment to predictive modelling of migration flows. Gawryluk et al. (2024) and Saena et al. (2020) demonstrate that Hybrid SARIMA–ANN models effectively capture nonlinear relationships between migration, gender, and macroeconomic shocks. Meanwhile, Ordinal Logistic Regression (OLR) has been applied to analyze ordered migration outcomes such as skill levels and settlement intentions (Avato, 2010; Liu et al., 2017; Ette & Erling Hagen, 2024). Few studies in Sri Lanka integrate trend analysis, predictive forecasting, and OLR to see how gender, destination, and macroeconomic shocks jointly shape labour migration. This study addresses this gap.

3 DATA COLLECTION AND METHODOLOGY

This study adopts a quantitative analytical design using monthly secondary data from January 2014 to December 2024, obtained from the Sri

Lanka Bureau of Foreign Employment (SLBFE). The dataset includes total departures/registrations, gender, skill level (7 categories), and destination country (15 countries). To capture the effect of macroeconomic stress, monthly National Consumer Price Index (NCPI) data were incorporated, as NCPI reflects nationwide inflation pressures experienced by rural and estate households where most migrant workers originate. For the OLR analysis, data from 2014–2019 were used because detailed disaggregated skill-level data were consistently available only during this period; post-2019 records were affected by pandemic-related disruptions and category changes, making them unsuitable for modelling.

Graphical analysis was used to visually examine long-term migration patterns, gender shifts, and crisis-period disruptions. This method helps identify structural breaks caused by COVID-19 and the 2022 economic crisis, supporting descriptive interpretation before statistical modelling.

SARIMA models were applied to forecast future migration flows because migration data exhibit trend and strong seasonal patterns driven by foreign labour demand cycles. SARIMA effectively captures both non-stationarity and annual seasonal effects, making it appropriate for predicting future departures.

OLR was used to analyse determinants of migrant skill level, an ordered variable ranging from housekeeping (1) to professional (7). Gender, destination country, and total departures were included as predictors. OLR is appropriate because it models ordered categories while estimating how

each independent variable increases or decreases the probability of belonging to higher skill levels.

4 RESULTS AND DISCUSSION

4.1 Graphical Analysis

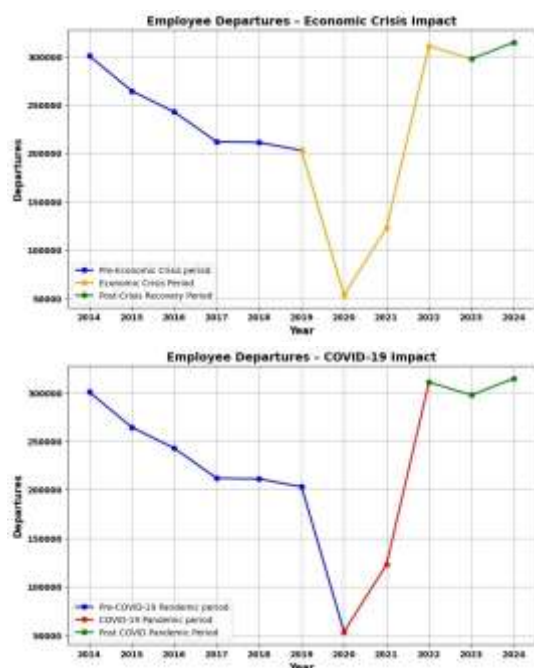


Figure 1: Registration for Foreign Employment During COVID-19 and Economic Crisis Periods

According to Figure 1, the analysis reveals that external shocks

profoundly shaped Sri Lanka's Labour migration trends. Departures fell sharply to about 55,000 in 2020 amid COVID-19 travel restrictions and Recruitment suspensions. As borders reopened and the 2021-2022 economic crisis deepened, migration surged to over 310,000 in 2022. Despite different causes, both crises produced a similar outcome: a post-crisis escalation in outbound migration driven by economic hardship and the search for income stability.

Figure 2 reveals a strong correlation between inflation (National Consumer Price Index (NCPI)) and labour migration. Rising inflation during 2020-2022, driven by fiscal instability and the economic crisis, coincided with a surge in foreign employment departures as households sought income security. Post-2023, both inflation and outward migration trends showed gradual stabilization following economic reforms. The NCPI proved more representative than the Colombo Consumer Price Index (CCPI), capturing rural cost-of-living pressures that predominantly influence migration decisions.

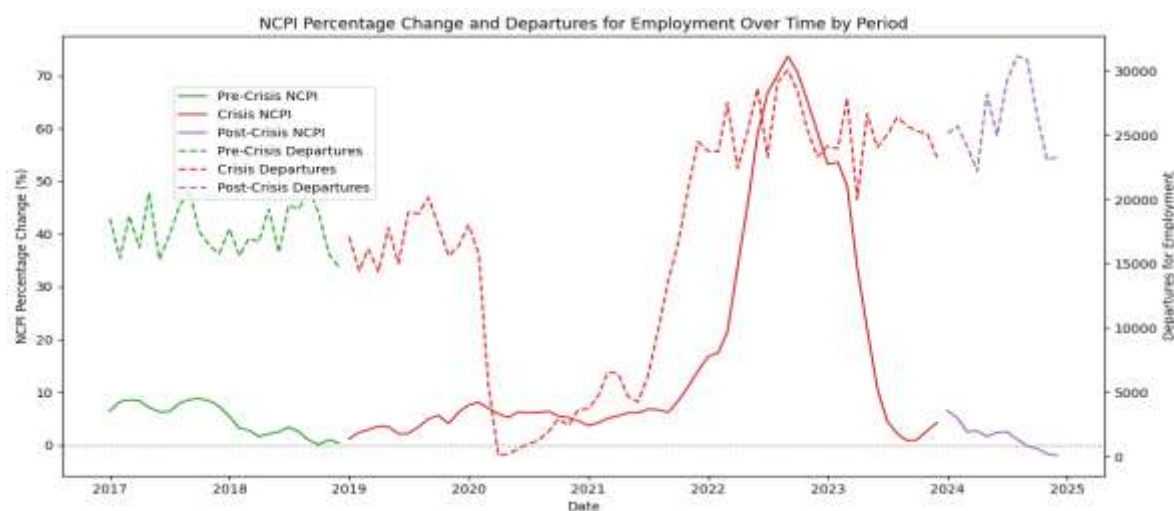


Figure 2: Analysis of Relationship Between Year-on-Year NCPI percentage and the Number of Employee Departures during Economic Crisis

4.2 Time Series Forecasting

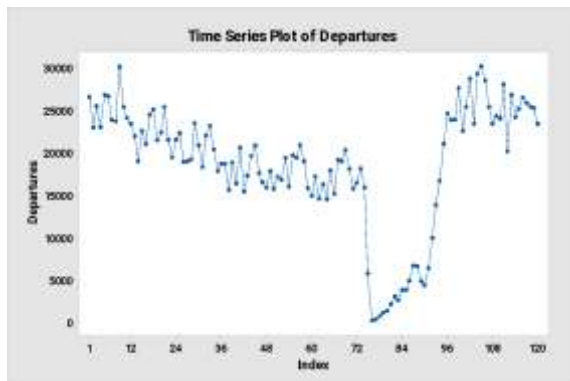


Figure 3: Time Series Plot for the Total Number of Departures for Foreign Employment

According to Figure 3, it is hard to detect a clear seasonal pattern because of the large spike in the middle. The Auto-Correlation Function (ACF) showed a strong trend indicating the series is not stationary. After the trend difference, the series became stationary, and the identified tentative model was SARIMA (1, 1, 1) (1, 0, 1)₁₂. Then, varying the terms of SARIMA (p, 1, q) (P, 0, Q)₁₂, different models were considered for searching the most effective structure of the forecast. Then, two adequate models were identified after diagnostic checking, including testing the significance of parameters, randomness of residuals, Normality of the residuals, and parameter redundancy.

Table 1: MAPE and Accuracy values of the adequate models for Total Registrations

Model	MAPE	Accuracy (%)
SARIMA (0,1,0) (0,0,1) ₁₂	10.8548	89.1452
SARIMA (0,1,0) (1,0,0) ₁₂	5.81716	94.1828

Table 2: MAPE and Accuracy values of the adequate models for total Female Employee Registrations

Model	MAPE	Accuracy
SARIMA (1,1,0) (1,0,0) ₁₂	9.73373	90.2663
SARIMA (0,1,1) (1,0,0) ₁₂	8.71077	91.289

Among the two obtained adequate models, SARIMA (0,1,0) (1,0,0)₁₂ showed the highest accuracy, 94.1828% or the lowest MAPE value, and therefore, was selected as the best model to forecast total registrations. Following the same procedure with diagnosing and checking two adequate models were obtained for total female employee registration as well.

According to Table 2, the SARIMA (0,1,1) (1,0,0)₁₂ model indicates the highest accuracy or the lowest MAPE among the two adequate models. Therefore, SARIMA (0,1,1) (1,0,0)₁₂ is the best-fitted model.

4.3 Ordinal Logistic Regression Analysis

The OLR model was employed to examine the effects of destination, gender, and number of departures on migration skill levels. A total of 1,440,109 valid cases were analyzed with no missing data. Female migrants accounted for 30.6% of the total sample.

4.3.1 Model Fit and Goodness of Fit

The model significantly improved over the intercept-only model ($\chi^2 = 997,499.15$, $p < .001$), confirming that the predictors explained variations in skill levels. Although the Pearson and Deviance chi-square tests were significant ($p < .001$), this is expected in

large datasets. Nagelkerke $R^2 = 0.523$ indicated a strong model, explaining 52% of the variance in skill level, while McFadden $R^2 = 0.222$ reflected good predictive accuracy as shown in Table 3.

Table 3: Model Summary

Statistic	Value
χ^2 (Model Fit)	997,499.15 ($p < 0.001$)
Nagelkerke R^2	0.523
McFadden R^2	0.222

4.3.2 Predictor Effects

Table 4 illustrates the odds ratio (OR) results that were obtained. Gender and country of destination emerged as significant predictors ($p < 0.05$), whereas the number of departures was not statistically significant ($p = 0.142$).

Country Effects: Migrants to Jordan (OR = 10.41), Japan (OR = 6.59), and Singapore (OR = 2.44) were more likely to hold higher skill levels compared to those migrating to the UAE (reference). In contrast, migrants to Cyprus (OR = 0.25), Lebanon (OR = 0.11), and South Korea (OR = 0.12) had significantly lower odds.

Gender Effects: Females (Gender = 0) had much lower odds (OR = 0.015) of belonging to higher skill categories compared to males (Gender = 1).

Departures: The number of departures had no significant effect on skill level once gender and destination were controlled for.

Table 4: Selected Odds Ratios for Key Predictors

Predictor	Odds Ratio (OR)	95% CI	Interpretation
Jordan	10.41	10.17-10.66	Higher skill level likelihood
Japan	6.59	6.23-6.96	Higher skill level likelihood
Singapore	2.44	2.36-2.52	Higher skill level likelihood
Lebanon	0.11	0.10-0.12	Lower skill level likelihood
Female (vs. Male)	0.015	0.015-0.015	Much lower skill level likelihood

4.3.3 Test of Parallel Lines

The Test of Parallel Lines was significant ($\chi^2 = 421,293$, $df = 75$, $p < .001$), suggesting a potential violation of the proportional odds assumption. However, given the large sample size, this effect is minimal and does not invalidate the model's interpretive strength.

5 CONCLUSION

Sri Lanka's labour migration is highly responsive to domestic and global shocks. Migration fell sharply during COVID-19 but surged in the 2022 economic crisis, driven by inflationary pressures (NCPI). Despite stable unemployment by 2024, foreign employment remained high, reflecting

reliance on migration for income security. SARIMA forecasting predicted future flows with over 90% accuracy, confirming its suitability for seasonal migration data. OLR analysis showed skill levels are strongly shaped by gender and destination, with high-skilled migrants mainly going to Jordan, Japan, and Singapore, while women and migrants to Lebanon and South Korea remain concentrated in low-skilled categories. Departures volume was not a significant factor. To improve OLR accuracy, future research should consider partial proportional odds or multinomial logistic regression to address non-parallel effects and sparse categories, expand predictor variables such as education and work experience, and increase sample granularity to reduce empty cells.

These findings underscore the importance of gender-sensitive, skill-oriented, and crisis-resilient migration policies, while highlighting the need to strengthen domestic employment opportunities to reduce long-term reliance on foreign labour markets and mitigate potential brain drain.

REFERENCES

- Avato, J. (2010). *Migration pressures and immigration policies: New evidence on the selection of migrants*. World Bank Policy Research Working Paper No. 5356. The World Bank.
- Dharmadasa, R. A. P. I. S., & Herath, H. M. T. R. (2020). A gravity model analysis of international migration from Sri Lanka to Southeast Asian and European countries. *Sri Lanka Journal of Economic Research*, 81), 113–136.
- Dissanayake, L., & Samarakoon, I. (2021). The pattern of temporary international migration in Sri Lanka. *Colombo Journal of Multi-Disciplinary Research*, 6(1), 97–110.
- Ette, A., & Hagen, E. (2024). *Skill-based migration intentions and labour market integration: An ordinal logistic regression approach*. *Migration Studies*, 12(1), 89–110.
- Gawryluk, A., Komor, A., Kulisz, M., Zarębski, P., & Katarzyński, D. (2024). Artificial neural networks as a method for forecasting migration balance (*A case study of the city of Lublin in Poland*). *Sustainability*, 16(24), 11249.
- International Labour Organisation. (2015, August 5). *ILO and migrant workers*.
- Kaluarachchi, S., & Jayathilaka, R. (2024). Unveiling Sri Lanka's brain drain and labour market pressure: *A study of macroeconomic factors on migration*. *PLOS ONE*, 19(3), e0300343.
- Kupriyanov, A. (2024). Sri Lanka crisis: The long shadow of neocolonialism. *World Economy and International Relations*.
- Liu, X., Zhang, Y., & Wang, L. (2017). *Modelling migration intention using ordinal logistic regression: Evidence from rural China*. *Population, Space and Place*, 23(8), e2064.
- Saena, P., Lee, H., & Choi, J. (2020). *Hybrid SARIMA-ANN approach for forecasting international migration trends under economic shocks*. *Applied Soft Computing*, 94, 106468.



Integrating RFM Analysis and K-Means Clustering for Customer Segmentation in an Apparel Equipment Provider in Sri Lanka

De Mel WRU¹, Aponsu GMLM¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
rasangiudayanademel@gmail.com

ABSTRACT

This study aims to analyze customer segmentation and purchasing behaviour within the consumables and spare parts division of a leading apparel equipment provider in Sri Lanka, with a focus on understanding loyalty dynamics, identifying high-value segments and evaluate behavioral shifts over time. Using five years of transactional data (2020–2024), the study applied Recency–Frequency–Monetary (RFM) analysis and K-means clustering to classify customers, supported by transition matrix modelling to assess year-to-year changes in engagement and loyalty. This integrated approach enabled both behavioural description and temporal evaluation. Results show that “Champions” and “Loyal Customers” consistently contributed the highest value, forming the core revenue segment, while “At-Risk” and “Lost/Others” groups displayed declining engagement, especially during the 2022 disruption. Loyalty patterns stabilized in the following years, and combining RFM with clustering produced more nuanced and reliable segmentation than rule-based methods alone. The findings support improved customer profiling, more accurate loyalty tracking, and stronger behavioural segmentation for strategic decision-making in industrial B2B environments. The integration of RFM, clustering, and transition analysis enhances marketers’ ability to identify high-value customers and detect early signs of attrition. This study contributes rare empirical evidence on data-driven customer segmentation within Sri Lanka’s apparel machinery industry by integrating behavioural analytics, machine-learning clustering, and temporal transition analysis.

KEYWORDS: Apparel Industry, Customer Behaviour, Customer Segmentation, K-Means Clustering, RFM Analysis

1 INTRODUCTION

The Sri Lankan apparel industry represents one of the country’s most vital export sectors, relying heavily on continuous machine efficiency and the availability of consumables and spare parts. ABC Company (Pvt) Ltd serves as a regional distributor for globally recognized brands in apparel machinery and Computer Aided Design (CAD)/Computer Aided Manufacturing (CAM) solutions. Despite maintaining a comprehensive sales database, the organization historically depended on aggregated sales figures without analyzing customer-level behaviour.

Therefore, this research demonstrates how quantitative techniques can be applied to real-world industrial sales data to reveal customer behavior patterns and provides actionable insights into customer segmentation and retention strategies, enabling more informed and proactive decision-making.

2 LITERATURE REVIEW

Customer segmentation serves as a core principle in marketing analytics, allowing firms to categorize customers into homogeneous groups to enhance targeting and retention strategies (Wedel & Kamakura, 2012). The Recency,

Frequency, Monetary (RFM) Model proposed by Hughes (1994), remains one of the overal frameworks for assessing purchasing behaviour, measuring how recently customers have bought, and how frequently they purchase and their overall spending volume. The RFM approach, grounded in database marketing theory (Tsiptsis & Chorianopoulos, 2009), is valued for its simplicity and quantitative precision, making it suitable for industrial applications with continuous transactional data. Nevertheless, conventional rule-based segmentation often fails to uncover hidden behavioural patterns in large datasets. To overcome these limitations, researchers have integrated clustering algorithms such as K-means with RFM analysis, creating hybrid approaches that yield deeper insights into customer value and behaviour (Birant & Kut, 2007; Han, Kamber, & Pei, 2011). Such methods enhance accuracy and interpretability by identifying natural behavioural groupings (Rineiro, Pinto, & Silva, 2020). Empirical studies on apparel machinery and spare parts sectors, especially in emerging economies like Sri Lanka, are scarce. This study, therefore, addresses a critical research gap by applying integrated RFM-clustering to analyze customer behavior, providing actionable insights for strategic retention, and value optimization.

3 METHODOLOGY

This study adopted a quantitative research design to examine customer purchasing behavior using five years of transaction data (2020-2024) from an apparel equipment provider in Sri Lanka. The analysis was conducted in three stages namely RFM Analysis, K-means

Clustering and Transition matrix analysis for temporal behaviour tracking. The data set included variables such as customer ID, purchase date, quantity, unit price, and total value. Three key behavioural dimensions were computed: Recency (R) - days since last purchase, Frequency (F) - number of transactions, and Monetary (M) - total expenditure. Each customer received an RFM score ranging from 1 to 5, where higher scores indicated stronger engagement. The equations and rules that were used for RFM Scoring and Segmentation are mentioned below.

$$R_i = D_{analysis} - D_{last\ purchase,i} \quad 3.1$$

$$R\_Score_i = QuantileRank(R_i, ascending) \quad 3.2$$

$$F_i = \quad 3.3$$

$$Count\ of\ purchases\ by\ customer\ i\ in\ the\ a \quad 3.4$$

$$F\_Score_i = \quad 3.4$$

$$QuantileRank(F_i, descending)$$

$$M_i = \sum_{j=1}^{n_i} Purchase\ Amount_{ij} \quad 3.5$$

$$M\_Score_i \quad 3.6$$

$$= QuantileRank(M_i, descending)$$

$$RFM\_Value_i = \quad 3.7$$

$$(R\_Score_i, F\_Score_i, M\_Score_i)$$

1. Champion $\rightarrow R\_Score \geq 4$,
 $F\_Score \geq 4$, $M\_Score \geq 4$
2. Loyal Customers $\rightarrow F\_Score \geq 4$ &
 $M\_Score \geq 3$
3. Big-Spenders (Infrequent) $\rightarrow$
 $M\_Score \geq 4$ & $F\_Score \leq 2$
4. New Customers/Recent $\rightarrow R\_Score$
 $= 5$ & $F\_Score \leq 1$
5. At-Risk (was good, now inactive) $\rightarrow$
 $R\_Score \leq 2$ & ($F\_Score \geq 3$ OR
 $M\_Score \geq 3$)
6. Lost/Others $\rightarrow$ else

The segmentation rules were defined based on the 1–5 RFM quintile scores, where 5 represents the top 20% of customers and 1 represents the lowest 20%. Thresholds such as ≥ 4 or ≤ 2 were therefore chosen to distinguish clearly between strong and weak customer behaviors, with higher scores identifying high-value groups (e.g., Champions, Loyal) and lower scores indicating declining or inactive customers (e.g., At-Risk, Lost). The rule for New Customers ($R = 5$ and $F \leq 1$) isolates customers who are very recent but have minimal purchase history. These thresholds ensure

that the segmentation reflects the natural distribution of customer performance within the dataset. K-means clustering was applied to standardized RFM scores, with the optimal number of clusters selected using the Elbow and Silhouette methods. Subsequently, Transition Matrix Analysis was conducted to examine year-to-year customer movements, identifying behavioral shifts and loyalty retention trends in RFM-based segmentation.

4 RESULTS AND DISCUSSION

4.1 RFM Analysis

Table 1: Summary Table of RFM Analysis and Customers Segmentation of Consumables & Spare Pats Customers for 2020-2024

Year	Segment	Count	Average Recency (Days)	Median Recency (Days)	Average Frequency	Median Frequency	Average Monetary	Median Monetary
2020	Loyal Customers	11	52.45	51	8.82	8	35523.843	20677.500
2020	At-Risk	6	98.50	97	4.17	4	16258.615	13612.375
2020	Champions	11	10.36	9	12.82	12	33372.494	33625.250
2020	Lost / Others	28	101.36	68	3.21	1.5	5262.038	3158.125
2020	Big Spenders	3	54.00	65	3.33	3	23466.287	22198.250
2020	New Customers	1	1.00	1	1.00	1	283.500	283.500
2021	Champions	14	8.21	9	14.93	14	49634.559	43683.355
2021	Loyal Customers	11	27.55	24	13.27	13	29142.196	21899.500
2021	Lost / Others	27	71.41	51	3.78	3	12935.044	5919.970
2021	Big Spenders	4	44.00	47.5	5.00	5	53506.063	41186.250
2021	At-Risk	8	76.75	67	5.50	5	11428.283	11224.625
2021	New Customers	1	2.00	2	1.00	1	494.250	494.250
2022	Champions	6	279.50	279.5	5.83	5	21945.658	16800.063
2022	Loyal Customers	15	300.60	298	4.07	3	10336.456	9401.020
2022	Lost / Others	24	301.63	290	1.42	1	2087.763	1478.625
2022	At-Risk	8	317.25	312.5	1.63	2	4462.562	2649.058
2022	Big Spenders	4	310.00	309	1.00	1	10856.988	8700.350
2022	New Customers	1	283.00	283	1.00	1	525.000	525.000
2023	Champions	15	12.60	12	16.47	13	62743.899	53032.530

2023	Lost / Others	40	89.10	39	3.33	2	13897.437	6046.110
2023	Big Spenders	3	75.67	47	2.67	3	40072.737	37568.711
2023	Loyal Customers	17	32.35	37	10.65	11	29279.227	20220.030
2023	At-Risk	7	88.00	75	3.57	4	12707.318	14287.250
2023	New Customers	1	9.00	9	1.00	1	458.250	458.250
2024	Lost / Others	35	66.00	47	4.83	4	17447.108	8908.180
2024	Loyal Customers	14	25.43	19.5	13.07	12	40574.643	25205.380
2024	Big Spenders	6	57.83	55.5	3.67	4	56616.487	49484.115
2024	Champions	15	9.80	8	16.40	14	54011.717	47479.000
2024	At-Risk	12	90.17	71	6.42	5.5	15447.154	13975.115

Table 1 highlights that Champion and Loyal customers consistently demonstrated high frequency, and monetary contributions, reinforcing their role as the company's most valuable segments. Big Spenders maintained significant spending levels despite lower transaction frequency, reflecting project-based purchasing patterns typical of industrial clients. In contrast, At-Risk and Lost/Others segments exhibit longer recency and lower monetary values, indicating potential disengagement and the need for retention strategies. The steady inflow of new customers each year signifies sustained market reach and future growth potential.

4.2 K-Means Clustering

4.2.1 Year-By-Year Clustering

Over the five years, customer clustering evolved from a fragmented structure in 2020 to greater stability by 2024. Figure 1 shows the type of cluster plots obtained as an example. As for the other 4 years, also plots obtained. Following a decline in 2022, *Champions* and *Mid-value* segments re-emerged as key contributors, emphasizing the need

for sustained loyalty programs and targeted retention strategies.

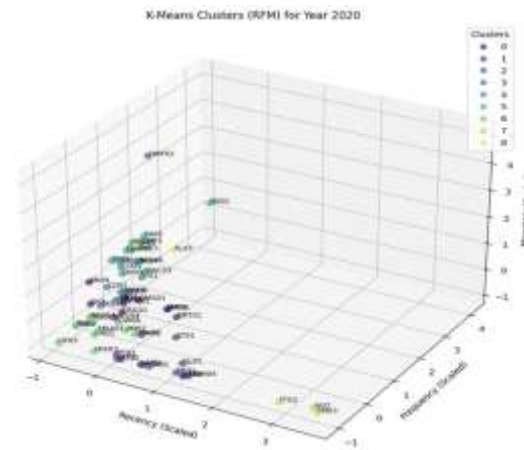


Figure 1: K-Means Cluster plot for the Year 2020

Table 2: Year-By-Year Clustering Summary

Year	Optimal No. of Clusters	Silhouette Score	Key Cluster Types / Characteristics
2020	9	Moderate separation	Highly fragmented structure with low-value, mid-tier, and elite groups (Clusters 3 & 7: €58,000–€121,000 avg. spend).

2021	2	0.65 (Good)	Highly fragmented structure with low-value, mid-tier, and elite groups (Clusters 3 & 7: €58,000–€121,000 avg. spend).	2022 → 2023	75% of <i>At-Risk</i> and all <i>New Customers</i> lost; <i>Loyal</i> partially stable (36% retained, 36% upgraded).
2022	2	0.54 (Fair)	Cluster 0: Very high recency (≈305 days), low frequency (≈1.7), minimal spend (€4,419); Cluster 1: Moderate frequency (≈6), higher spend (€18,977).	2023 → 2024	<i>Champions</i> retention improved to 33%; 35% of <i>Loyal</i> upgraded; 100% of <i>New Customers</i> lost again.
2023	3	0.61 (Good)	Cluster 0: <i>Mid-value</i> (Recency ≈35 days, Freq. ≈6.3, €19,711); Cluster 1: <i>At-risk</i> (Recency ≈207 days, Freq. ≈1.5, €5,676); Cluster 2: <i>Elite Champions</i> (Recency ≈20 days, Freq. ≈17, €78,029).		
2024	2	0.59 (Mode rate-Good)	Cluster 0: <i>Mid-value</i> (Recency ≈104 days, Freq. ≈3.3, €11,546); Cluster 1: <i>Champions</i> (Recency ≈21 days, Freq. ≈12, €43,735).		

4.3 Analysis of Transition Matrix

Table 3: Summary of RFM-based Customer Segment Transition Patterns (2020–2024)

Period	Key Transition Insights
2020 → 2021	64% of <i>Champions</i> retained; 40% of <i>Loyal</i> upgraded to <i>Champions</i> ; 100% of <i>New Customers</i> became <i>At-Risk</i> .
2021 → 2022	Only 23% of <i>Champions</i> retained; 31% downgraded to <i>Loyal</i> ; <i>At-Risk</i> split evenly between <i>Lost/Others</i> and <i>Loyal</i> .

The transition matrix analysis revealed consistent loyalty among *Champions* and gradual improvement in *Loyal Customer* retention, though overall volatility remained high. A recurring trend was the persistent churn of *At-Risk* and *New Customers*, with the *Lost/Others* segment emerging as a dominant absorbing category across years. While 2024 reflected a modest recovery in customer stability, the findings emphasize the need for stronger onboarding processes and sustained relationship management to enhance long-term retention.

5 CONCLUSION

This study employed RFM analysis, and K-means clustering to investigate customer segmentation and behavioral patterns in consumables and spare parts sales from 2020 to 2024 of ABCD Company. The decline in customer activity during 2022 reflects the combined impact of economic uncertainty and operational constraints, whereas the subsequent recovery highlights the resilience of core customer groups. The RFM analysis and K-means integrated framework provided more nuanced insights than using either approach alone, capturing both behavioural differentiation and underlying structure in purchasing

patterns. Transition matrices further clarified the dynamic nature of loyalty by revealing mobility across segments. Stable retention among Champions and Loyal Customers indicates strong long-term relationships, while the persistent presence of At-Risk and Lost/Others segments suggest vulnerabilities in customer engagement processes. These findings support more accurate customer profiling, enable early detection of customer disengagement, and highlight the firm's need for continuous monitoring of behavioral shifts beyond transactional volume.

However, the behavioural patterns of customers are clearly identified, behavioral motivations cannot be fully explained. Also, retention strategy design, reactivation program formulation, and inventory optimization could not be fully developed because the dataset lacked data and variables related to customer satisfaction, service response, product dependency, and customer tenure. These aspects should therefore be addressed in future work.

REFERENCES

- Birant, D., & Kut, A. (2007). ST-DBSCAN: An algorithm for discovering clusters in spatial-temporal data. *Data & Knowledge Engineering*, 60(1), 208–221.
- Gupta, P., Mehta, S., & Arora, R. (2020). Predictive segmentation for spare parts inventory optimization. *Journal of Industrial Marketing*, 36(2), 144–160.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
- Hughes, A. M. (1994). *Strategic Database Marketing*. McGraw-Hill.
- Kim, J., & Lee, D. (2022). *Hybrid RFM and predictive modeling for B2B customer retention*. *Decision Support Systems*, 156, 113–145.
- Ribeiro, J., Pinto, C., & Silva, F. (2020). Customer clustering and lifetime value prediction in B2B markets. *Decision Support Systems*, 130(1), 113–126.
- Tsipis, K., & Chorianopoulos, A. (2009). *Data Mining Techniques in CRM: Inside Customer Segmentation*. Wiley.
- Wedel, M., & Kamakura, W. A. (2012). *Market Segmentation: Conceptual and Methodological Foundations*. Springer.
- Zhao, L., & Wang, P. (2023). *Customer churn prediction using ensemble learning in industrial sales analytics*. *Expert Systems with Applications*, 225, 120093.



Forecasting Quality Outcomes in Apparel Manufacturing Using Time Series Models: A Case Study Based on Final Inspection Data

Sathsara WHT¹, Herath CHMAU¹, Dahanayaka SD¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
thamindusath123@gmail.com

ABSTRACT

This study develops a predictive time series framework to model and forecast weekly defect rates in an apparel production environment from 2019 to 2025. The goal is to identify trends, capture volatility behavior, and provide actionable forecasts for proactive quality management in the apparel industry. Stationarity is assessed with the Augmented Dickey-Fuller (ADF) test and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, while model selection using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) outlined ARIMA (4,1,2) as the best performing mean model. Residual diagnostics showed conditional heteroskedasticity, indicating some variation in weekly defect rates was time-related, with weekly variability explained by fluctuations within the production process itself. Therefore, GARCH (1,1) processes were included, which yielded an ARIMA - GARCH model. The research focuses on predicting future defect trends in order to support proactive, forward-looking quality interventions, which could help manufacturers prepare for periods of high risk instead of relying exclusively on past history. Forecasting performance metrics (RMSE = 0.405, MAE = 0.256, MAPE = 20.03%) indicate an excellent degree of predictive ability. Overall, volatility modelling noticeably improves accuracy over more traditional linear models. The results provide a practical, data-driven tool for improving quality control, resource planning, and defect-prevention strategies in apparel manufacturing environments.

KEYWORDS: Apparel Manufacturing, ARIMA, Defect Rate, Forecasting, GARCH

1. INTRODUCTION

The apparel industry which greatly invests in Sri Lanka's economy is widely known for its good quality and ethically manufactured products. The organization selected for this study is a major apparel manufacturer that operates several plants, and uses sophisticated technology, and sustainable practices to ensure the quality of its products. This report will focus upon its Quality Assurance (QA) Department and the final inspection when the manufacturing process assesses specification with the customer. Weekly defect rates from 2019-2025 were analyzed to determine the trend in quality performance over time. In garment manufacturing, variations in the level of defects will impact cost, efficiency, and

buyer satisfaction. Although the cause of defects can be evaluated, numerous industry analyses are retrospective and do not predict how the defect may or may not change with future production. Predicting is particularly important when a factory produces many garments that contain different defect risks. Predicting these various defects allows for resource planning and quality controls to be directed accordingly.

Accordingly, the proposed research employs ARIMA and GARCH time series models to predict weekly defect rates and capture both trends and volatility of the defects. The purpose of the proposed research is to assist in transitioning from reactive quality control to proactive, data-driven decision

making that supports defect prevention and quality is consistent across different manufacturing conditions.

2. LITERATURE REVIEW

According to Hyndman and Athanasopoulos (2021), time series forecasting is considered highly relevant to manufacturing, as they support early detection of deviations in processes and improvements in production efficiency.

Bollerslev (1986) introduced the general autoregressive conditional heteroskedastic (GARCH) model, a model that permits conditional variance to change over time. This model captures measures of high and low volatility in a method coined "volatility clustering." As a hybrid model, ARIMA–GARCH model accounts for both mean and variance dynamics, yielding more reliable forecasts in uncertain industrial settings.

Fatima et al (2024) conduct a review of time series forecasting methods for manufacturing systems. The researchers concluded that forecasting models of the ARIMA–GARCH variety outperform traditional models when data exhibit either structural breaks or nonlinear variance pertaining to volatility. Jaipuria and Mahapatra (2021) also apply the ARIMA–GARCH framework to forecasting supply-chain demand, and they achieve better accuracy by accounting for sudden surges in uncommon demand.

Lekamge and Ekanayake (2021) examined the fluctuation of defect rates in apparel production. They reported that variability generates irregular spikes over time due to fluctuations in operator efficiency, raw material, and equipment issues.

Hossain (2015) has stressed that nonlinear extensions of GARCH, such as EGARCH and TGARCH, are useful to capture asymmetric volatility responses in relation to sudden increases in defect rates.

By using an ARIMA–GARCH framework to predict weekly defects, this study fills a gap in the existing literature where techniques taking volatility into account are not employed. By modeling both trends and time-varying variance, the model supports proactive quality decisions, better inspection plans, and resource allocation decisions in the apparel manufacturing context.

3. METHODOLOGY

The research involved the use of 325 weekly defect-rate observations obtained from , records of final inspection in 2019 to 2025 at Sri Lankan apparel manufacturing facilities. Data was retrieved from the Quality Assurance arm of the organization, which is accountable for documenting inspection results from all major facilities. The defect rate was calculated as $(\text{total defects} \div \text{total samples checked}) \times 100$. The data set had no missing data, and preliminary checks detailed that there were no discovered statistical outliers. All analyses were carried out in R.

Modeling was developed in accordance to the Box–Jenkins methodology and with stationarity checked using the ADF and KPSS tests, indicated non-stationarity that needed first differencing. The ACF, PACF, and STL decomposition terms, were used to select AR and MA terms, while the selected ARIMA model was also selected using Auto ARIMA with minimum AIC and BIC. The residual diagnostics were

then performed using the Ljung–Box test, residual plots and the ARCH LM test. The findings indicated that from the residuals there was evidence of volatility clustering, and therefore included a GARCH component.

4. RESULTS AND DISCUSSION

Table 1 provides descriptive statistics for the weekly defect rates for years 2019-2025. The mean defect rate was 1.53%. The standard deviation of 0.88% provides evidence of week-to-week fluctuations. The maximum observed defect rate was 6.71%, which suggests there were instances driven by unforeseen production delays issues that led to increased defect rates.

As demonstrated in Fig. 1, defect rates fluctuated from week-to-week, indicating variability, albeit without any identifiable long-term trend. Defect rates remain generally stable with sporadic spikes in defect levels that are possibly due to disruptions in production, material issues or issues related to operators. These effects seem short-lived, as they do not reflect volatility over time.

Table 1: Descriptive Statistics of Weekly Defect Rate

Mean (%)	SD (%)	Max (%)	Min (%)
1.53	0.88	6.71	0.42

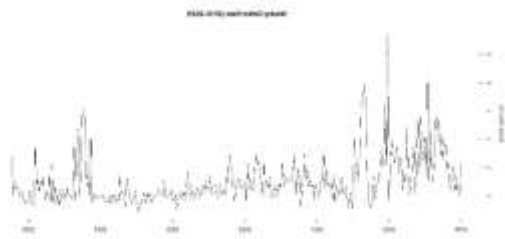


Figure 1: Time Series Plot of Weekly Defect Rate

Table 2: Results of ADF Test for Original and Differenced Series

	Original	Differenced
ADF P-value	0.078	0.001
KPSS test P-value	1.863	0.024

According to Table 2, the original weekly defect rate series is not stationary, as both the ADF and KPSS P-values indicate ($P\text{-value} > 0.05$).

However, when the series is first differenced, the ADF P-value decreases to 0.001 and the KPSS test P-value decreases to 0.024. Therefore both the ADF and KPSS test p-values indicate that the series is stationary ($P\text{-value} < 0.05$).

Fig. 2 revealed that the defect rate series contained a mild trend and small cyclical components but no strong seasonality. This justifies the application of ARIMA without seasonal terms.

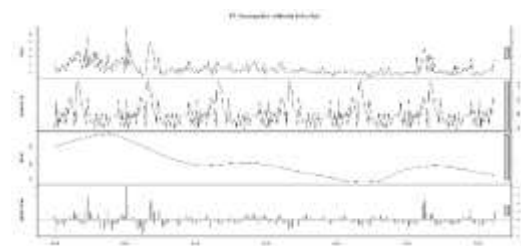


Figure 2: STL Decomposition of Weekly Defect Rate

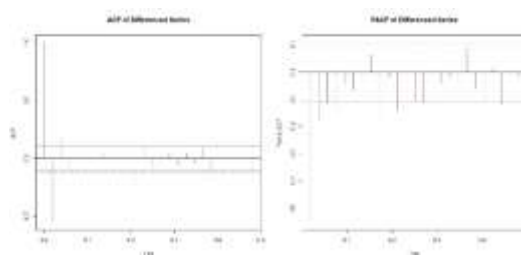


Figure 3: Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) for Differenced Series

Fig. 3 suggested the presence of an AR(4) structure followed by an MA(2) structure, which also supported the ARIMA(4,1,2) model, further validated through the Auto ARIMA procedure, which provided the model with the minimum AIC and BIC values. The residual diagnostics indicated that the residuals were uncorrelated (Ljung–Box $p = 0.9839$), indicated the residuals were not normally distributed (Shapiro–Wilk $p = 3.14e-15$), and suggested potential volatility effects (ARCH LM $p = 0.0217$), implying a GARCH extension was warranted.

Table 3 compares GARCH model specifications using AIC, BIC, and HQC values. The GARCH(1,1) model has the lowest values across all criteria, making it the best-fitting model for defect rate volatility.

Table 4 demonstrates that all estimated coefficients obtained from the ARIMA(4,1,2)–GARCH(1,1) model were significant, which implies that the mean and variance of this series were well modeled. The model appears to fit quite well and reflects the temporal structure and time-varying volatility of the defect rate data.

Fig. 4 shows the residuals closely follows the straight reference line, which indicates approximate normality.

Table 3: Model Comparison Based on Information Criteria

Model	AIC	BIC	HQC
GARCH (1,1)	1.8504	1.9698	1.8485
GARCH (1,2)	1.8563	1.9759	1.8534
GARCH (2,1)	1.8547	1.9742	1.8519
GARCH (2,2)	1.8598	1.9924	1.8549

Table 4: Parameter Estimation Results of ARIMA(4,1,2)–GARCH(1,1)

Parameter	Estimate	Pr(> t)
μ	1.4123	0.00000 ***
ar1	1.5378	0.00000 ***
ar2	-0.7015	0.000016 ***
ar4	0.2462	0.0012 **
ma1	-1.3510	0.000000 ***
ma2	0.7562	0.000000 ***
ω	0.0485	0.00029 ***
α_1	0.4868	0.000004 ***
β_1	0.5122	0.000000 ***

Fig. 5 shows a bell-shaped distribution which indicates that the standardized residuals are symmetrically distributed.

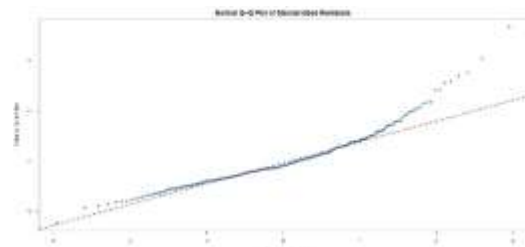


Figure 4: Normal Q–Q Plot

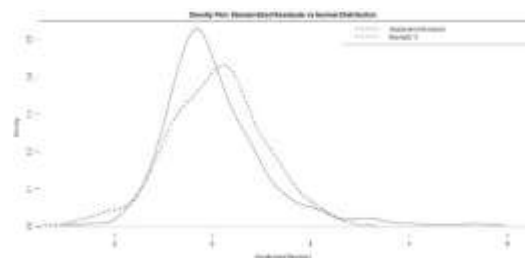


Figure 5: Density Plot of Standardized Residuals

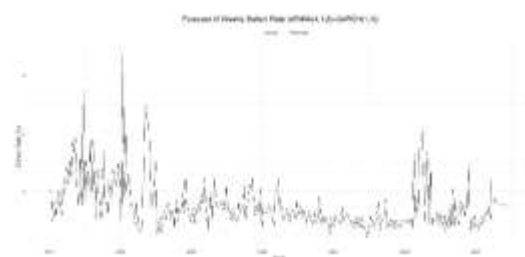


Figure 6: Forecast of Weekly Defect Rate (ARIMA(4,1,2)–GARCH(1,1))

Table 5: Forecasting Performance Metrics

Metric	Result
RMSE	0.405
MAE	0.256
MAPE (%)	20.03

Fig. 6 indicates gradual stabilization of defect rates, from 1.57% in week one to 1.50% in week twelve, reflecting improved consistency while remaining stable over the period.

According to Table 5, the accuracy of the forecasts demonstrates good predictive performance with a MAPE of 20.03% and an RMSE of 0.405. The ARIMA(4,1,2)–GARCH(1,1) model effectively captures both mean and volatility characteristics and provides strong normative guidance for quality assurance processes.

5. CONCLUSION

This study used ARIMA(4,1,2)–GARCH(1,1) model to forecast weekly garment defect rates. The ARIMA part estimated the mean movements, while the GARCH part estimated the volatility clustering common in quality data. The model was predictive and indicated its viability for short-term quality forecasting. The predictions provided QA managers the ability to identify high defect risk periods for inspection, time inspections, effective allocation of resources, and potentially minimize costs associated with rework or rejected products. Findings also bear implications for sound policy related to operational practice in apparel plants at the frequency or inspection regime level (e.g., increased inspection frequency for fabric with a high defect rate, inspection

focus on specific lines based on defect rates, etc.) and even number of inspectors based on anticipated quality arrangements or performance. Limitations include a single organization as the unit of analysis; the aggregation of weekly defect counts may mask intra-week variation in defects. Even with limitations, the ARIMA–GARCH has demonstrated its real-world applicability and applicability to the fast fashion experience, either through predictive quality deviation index or specific, shorter-term period forecasts, with the intention to improve proactive quality management, operational efficiency, and ultimately policy making with various production contexts.

REFERENCES

- Bollerslev, T. (1986). *Generalized autoregressive conditional heteroskedasticity*. *Journal of Econometrics*, 31(3), 307–327.
- Fatima, S. S. W., Imran, M., Raza, A., & Khan, S. A. (2024). *A review of time-series forecasting algorithms for industrial manufacturing systems*. *Machines*, 12(6), 380.
- Hossain, A. (2015). *ARIMA with EGARCH family modeling*. David Publishing.
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.
- Jaipuria, S., & Mahapatra, S. S. (2021). *A hybrid forecasting technique to deal with heteroskedastic demand in a supply chain: An ARIMA–GARCH application*. *Operations and Supply Chain Management: An International Journal*, 14(2), 192–204.
- Lekamge, R., & Ekanayake, N. (2021). *Internal quality failures of apparel industry: A case from Sri Lanka*. *Open Journal of Business and Management*, 9(5), 2280–2293.



A Machine Learning Approach for Classifying Production Performance: A Case Study

Dias PPVA¹, Appuhamy PADAN¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
virajidias95@gmail.com

ABSTRACT

To prevent supply chain disruptions in the apparel industry, supporting manufacturing sectors must operate efficiently. However, various internal and external factors create production inefficiencies. Although external factors are uncontrollable, identifying key internal contributors is crucial in reducing inefficiencies. This study applies machine learning to identify the best classifier for daily production performance and its key contributing factors. The daily production performance consisted of two categories viz, underperformance (if the actual production falls short of planned) and normal performance and eight internal factors were collected over 44 months from the ABC company, a leading fabric and elastic manufacturer in Sri Lanka. Initially, explanatory analysis was conducted to observe the variation in the data. Then five machine learning models viz. Decision Tree, Random Forest, XGBoost, CatBoost, and Logistic Regression were trained using 80% of data and tested with remaining 20% data. Finally, the classification performances were compared based on Accuracy, Precision, Recall, F1-score, and Area Under the Curve. It was revealed that more than 50% of the days of the study period were underperformed. Moreover, there was a significant difference in breakdown machines and experienced workers between underperforming and normal days. The Binary logistic model emerged as the best classifier for production performance with the highest accuracy (86.9%), precision (85.6%), recall (89.9%), and F1-score (87.7%). The Model highlighted that the higher machine breakdowns and fabric production days increase underperformance, while greater experience and night shifts guarantee normal production. These data driven information aid management to identify the best practices to mitigate production underperformance.

KEYWORDS: Apparel Industry, Logistic Regression, Machine Learning, Production Performance.

1 INTRODUCTION

The apparel manufacturing industry plays a vital role in Sri Lanka's economy, being the highest export earner and providing over 300,000 direct and 600,000 indirect employment (SLAEA, 2025). Moreover, apparel industry is the Sri Lanka's single largest generator of foreign exchange, contributing around 6% to national GDP (Central Bank of Sri Lanka, 2024). The elastic and fabric manufacturing industry supports the local supply chain by reducing reliance on imports and improving global competitiveness. ABC (Pvt) Ltd is a

leading elastic and fabric manufacturer in Sri Lanka, supplying high quality products to local and global markets. Despite its strengths, the company faces recurring production inefficiencies caused by external factors such as economic instability, political influences, and global market conditions, as well as internal issues including machine breakdowns, quality defects, and limited skilled labor. These challenges have led to revenue losses, customer dissatisfaction, and reduced market competitiveness. While external factors cannot be controlled, internal factors can

be managed if the key drivers of inefficiency are accurately identified. Traditional experience-based decision-making is subjective and reactive, making it difficult to capture complex operational interdependencies. Therefore, data-driven methods are essential for diagnosing the root causes of underperformance. Machine Learning (ML) provides a powerful approach for uncovering hidden patterns in production data and predicting efficiency with high accuracy (Mokrova et al., 2018; Kumar & Raj, 2023). This study aims to apply ML techniques to determine the most suitable classification model for daily production performance (categorized as either normal or underperforming) based on internal operational factors and to identify the main contributors to production underperformance.

2 LITERATURE REVIEW

Predicting production performance has become increasingly important as manufacturers adopt data-driven methods to improve efficiency. Traditional tools like SPC, Regression, and FMEA have been widely used (Ahmad & Bhat, 2024), but their linear assumptions limit the detection of complex operational relationships (Montgomery, 2019). Recent research therefore, shifts toward machine learning techniques capable of handling large, multivariate datasets more effectively.

Globally, studies show ML models outperform traditional methods in detecting production issues. Mokrova et al. (2018) found Random Forests and Decision Trees superior to regression, while Kumar and Raj (2023) used ensemble models to identify productivity losses. Jain and Kumar (2023) applied

image-based ML to detect fabric defects efficiently. In Sri Lanka, Amarasinghe (2015) highlighted workforce-related factors, skill gaps, low motivation, and absenteeism as major contributors to inefficiencies. Notwithstanding these results, research synthesizing both production and workforce-related variables in the apparel sector with ML is limited in Sri Lanka. This study fills the existing research gap by applying interpretable machine learning models to classify daily production performance in an elastic and fabric manufacturing company. It establishes a replicable, data-driven framework that can be effectively applied in similar manufacturing companies to identify inefficiencies, and support data-informed decision-making.

3 METHODOLOGY

The study utilized a dataset of 1,339 daily observations collected over 44 months, from January 1, 2022, to August 31, 2025, and sourced from the company's internal management system. The target variable (Y) of the ML models was defined as follows.

Daily Production Performance (DPP) = Actual Output – Planned Output

$$Y = \begin{cases} 1, & \text{if } DPP < 0 \text{ (Underperformed)} \\ 0, & \text{if } DPP \geq 0 \text{ (Normal)} \end{cases}$$

The data analysis utilized eight internal factors of the production process as independent variables, namely No. of breakdown machines per day, Total downtime per day (in minutes), Product Type (Elastic or fabric), Total length of defective fabric/elastic (in meters), Shift (day or night), Experience (no of workers with more than three year of experience), No of male employees , and No of

covering employees (employees assigned to cover the duties of absentees) per day.

At the preprocessing step, the data set was screened for missing values and outliers and none were present. Categorical variables (Shift and Product Type) were dummy encoded using ‘Day shift’ and ‘Elastic product’ as reference categories. Dummy encoding (k–1) was chosen over one-hot encoding to avoid multicollinearity. Exploratory data analysis was first conducted to examine variable distributions, identify production underperformance patterns, and assess correlations with daily performance. Five classification models, Binary Logistic Regression (BLR), Random Forest, XGBoost, CatBoost, and Decision Trees, were then developed using Google Colab. The dataset was split 80/20 for training and testing, and grid search with cross-validation was used for training. Since the two categories of the target variable were almost balanced, the classification performance of each model was assessed based on matrices viz. Accuracy, Precision, Recall, F1-score, and Area Under the Curve (AUC). Finally, models were compared to select the most accurate and interpretable one and identify key factors affecting production underperformance.

4 RESULTS AND DISCUSSION

It was revealed that 52.2% of the working days during the study period were underperforming and 47.8% days were normal performing, showing inefficiencies in the production process, which is not an occasional issue but a regular and frequent experience. On average, three out of 30 machines were non-operational daily, with total downtime ranging from 0 to 500 minutes

(mean 200 minutes). The number of experienced employees varied between 5 and 20 per day. Underperforming days had higher downtime and defective output and fewer experienced workers compared to normal days, while absenteeism and the number of male workers showed no significant difference. Figure 1 shows the correlation between the numerical variables and it suggests that none of the listed variables have a strong linear relationship with any of the others. Table 1 illustrates the classification performance of all five models. Among all models the Binary LR possessed the highest accuracy (0.8694), F1-score (0.8772), precision (0.8562) and recall (0.8993), outperforming more sophisticated ensemble algorithms such as XGBoost (accuracy = 0.8657, F1 = 0.8732) and CatBoost (accuracy = 0.8619, F1 = 0.8693). The Random Forest model performed fairly well (accuracy = 0.8545), whereas the Decision Tree model showed the lowest classification performance with accuracy of 0.7910. The AUC of more than 0.91 for all the ensemble models attested to outstanding discriminative ability, with BLR yielding a solid AUC of 0.9192. Five-fold cross-validation (folds of 267–268 samples each) was applied, showing stable performance across folds with mean accuracy = 0.834 (SD = 0.020) and

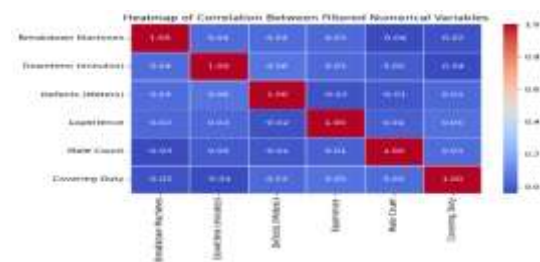


Figure 1: Correlation Between Numerical Variables

Table 1: Comparison of Models' Classification Performance

Model	Accuracy	Precision	Recall	F1-score	AUC
XGBoost	0.8657	0.8552	0.8921	0.8732	0.9319
Binary Logistic	0.8694	0.8562	0.8993	0.8872	0.9192
Random Forest	0.8545	0.8521	0.8705	0.8612	0.9304
CatBoost	0.8619	0.8542	0.8849	0.8693	0.9388
Decision Tree	0.7910	0.7823	0.8273	0.8042	0.7901

similar consistent precision, recall, F1-score, and ROC AUC values.

Although ensemble models like XGBoost and CatBoost achieved similar accuracy, ABC (Pvt.) Ltd prioritized an interpretable and easily deployable model. Binary Logistic Regression was selected as the best classifier due to its strong performance, clear coefficient-based insights, and high interpretability. Sensitivity analysis showed that removing any single variable kept the BLR accuracy above 84%, indicating strong robustness to omitted variable bias. The general form of the model used in this study is shown in Eq. (1).

$$\log\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \beta_0 + \sum_{i=1}^n \beta_i X_i \quad (1)$$

Where $P(Y=1)$ denotes the probability of daily production being underperformed, β_0 represents the intercept, and β_i are the coefficients corresponding to the independent variables X_i , where $i=1 \dots 8$.

Table 2 presents the coefficients, odds ratios, and p-values of the BLR. The number of employees with over three years of experience had the strongest negative effect (coefficient -0.4496 , OR = 0.64), indicating a 36% reduction in underperformance odds for

each additional experienced worker. Night shifts showed an 18% lower likelihood of underperformance compared to day shifts. Machine breakdowns had the strongest positive effect (coefficient 0.4439, OR = 1.56), meaning each additional breakdown increased underperformance odds by 56%. Fabric production also increased the likelihood of underperformance (coefficient 0.2745, OR = 1.32). Reference categories for the model were day shifts and elastic production. The remaining variables had minimal impact. P-values were used to assess statistical significance; Experience, Downtime, and Machine Breakdowns ($p < 0.001$) were highly significant predictors.

Table 2: The Coefficients of BLR model

Variable	Coeff(β)	OR(e^{β})	p-val
Intercept	2.1944	8.98	0.036
Breakdown_M	0.4439	1.56	<0.001
Product_Fab	0.2745	1.32	0.104
Downtime	0.0077	1.01	<0.001
Experience	-0.4496	0.64	<0.001
Shift (Night)	-0.1938	0.82	0.247
Covering D.	-0.0613	0.94	0.311
Male Count	-0.0039	0.996	0.694

The best Binary Logistic Regression model fitted to the daily production performance based on eight input variables was given in the following Eq. (2). Hosmer-Lemeshow test indicated imperfect fit (HL = 20.28, $p = 0.009$), and pseudo R^2 values suggested moderate explanatory power (McFadden = 0.433, Cox-Snell = 0.451, Nagelkerke = 0.601).

$$\log \left(\frac{P(\text{Underperformance}=0)}{P(\text{Normalperformance}=1)} \right) = 2.1944 + 0.4439(\text{Breakdown Machines}) + 0.0077(\text{Downtime}) + 0.0362(\text{Defects}) - 0.4496(\text{Experience}) - 0.0613(\text{Covering Duty}) - 0.0039(\text{Male Count}) - 0.1938(\text{Shift Night}) + 0.2745(\text{Product Type Fabric}) \quad (2)$$

The results align with previous studies (Amarasinghe, 2015; Sarkar et al., 2018), showing that interpretable machine learning can guide operational improvements. However, only 8 internal factors were considered, and data from a single facility limits generalizability, highlighting the need for external validation. Future work explore resampling techniques like SMOTE to address mild class imbalance and improve model robustness.

5 CONCLUSION

Efficient production is vital for competitiveness in fabric and elastic manufacturing. The study found that over half of ABC (Pvt.) Ltd.'s production days underperformed, largely due to machine breakdowns reaching up to 8 hours on some days and limited availability of experienced workers. Binary Logistic Regression performed best showing that breakdowns increase underperformance risk while experienced workers reduce it. Fabric production days

were more prone to underperformance, and night shifts performed better than day shifts. These findings offer clear, data-driven insights to improve workforce planning and operational stability.

REFERENCES

- Ahmad, N., & Bhat, M. A. (2024). Application of Statistical Process Control and FMEA for improving manufacturing performance: A review. *International Journal of Quality & Reliability Management*, 41(2), 215–229.
- Amarasinghe, R. (2015). Workforce-related factors influencing production performance in the Sri Lankan apparel industry. *Journal of Textile and Apparel Technology and Management*, 9(3), 45–53.
- Jain, P., & Kumar, R. (2023). Machine learning techniques for fabric defect detection and classification in the textile industry. *Journal of Manufacturing Systems*, 67, 421–432.
- Kumar, R., & Raj, A. (2023). Ensemble learning approaches for predicting manufacturing productivity losses: A comparative analysis. *Journal of Intelligent Manufacturing*, 34(5), 1893–1908.
- Mokrova, L., Petrov, I., & Novak, D. (2018). Predictive analytics for manufacturing process efficiency using machine learning methods. *Procedia Manufacturing*, 24, 55–62.
- Montgomery, D. C. (2019). Introduction to Statistical Quality Control (8th ed.). Hoboken, NJ: John Wiley & Sons.
- SLAEA. (2025). Sri Lanka Apparel Exporters Association Annual Report 2025. Sri Lanka Apparel Exporters Association.



Statistical Analysis of Data Quality Challenges Faced by Audit Professionals: A Case Study in the ABC Company

Diluxana K¹, Karunarathna WHDT¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
diluxana2000@gmail.com

ABSTRACT

This study investigates the data-quality challenges faced by audit professionals at ABC Company, a leading Sri Lankan firm specializing in audit and advisory services. The research was inspired by recurring issues observed during a six month period in the Governance, Risk, Compliance, Forensic, and Data Analytics (GRCFDA) Division, where the client provided datasets often contained missing entries, duplicate transactions, inconsistent formats, or corrupted files. The study aimed to identify, quantify, and analyze these data quality issues and evaluate their operational impact on audit workflows. A structured questionnaire was distributed among 52 professionals across four roles data analysts, internal auditors, external auditors, and IT auditors. Statistical techniques, including descriptive analysis, Analysis of Variance (ANOVA), and correlation tests were applied. Results revealed that missing records (70%), duplicate entries (55%), and inconsistent file formats (50%) were the most frequent problems, with corrupted files causing the longest delays. Correlation analysis showed a strong positive link between the number of data issues and additional audit hours ($r = 0.72$, $p < 0.001$). Reliability testing of the Likert-scale items yielded a Cronbach's alpha of 0.87, confirming strong internal consistency. Root cause analysis highlighted client-side weaknesses, inconsistent internal procedures, and limited automation. Because the sample was drawn solely from a single firm using non-random purposive sampling, the findings reflect the internal context of ABC Company and are not intended to generalize to the broader audit profession. The study recommends implementing standardized data templates and automated validation tools to improve audit quality, efficiency, and data reliability.

KEYWORDS: Auditing, Automation, Data Quality, Statistical Analysis, Workflow Efficiency

1 INTRODUCTION

The growing use of digital data has revolutionized auditing by enabling full population testing and advanced analytics. However, this transformation also introduces challenges when data provided by clients is incomplete or inconsistent. ABC Company, one of Sri Lanka's oldest audit firms, provides a broad range of professional services including audit, risk management, compliance, forensic accounting, and data analytics.

In recent years, professionals in the firm's GRCFDA division have

frequently encountered recurring data issues that delay audits and affect quality. These include missing entries, duplicated transactions, and corrupted or improperly formatted files. The motivation for this study emerged from direct observations of such challenges during a six-month industrial placement.

Given that the sample and observations are limited to a single organization, this study is framed as an organizational case study aimed at understanding internal operational challenges rather than making industry-wide generalizations.

2 LITERATURE REVIEW

The concept of data quality is widely discussed across disciplines. Wang and Strong (1996) define it as “fitness for use,” emphasizing dimensions such as accuracy, completeness, consistency, and timeliness. Later models (Pipino, Lee & Wang, 2002) provided quantitative frameworks for measuring these dimensions, showing that errors in datasets lead to inefficiency and poor decision-making.

In the auditing context, poor data quality directly increases audit risk. Studies by Redman (1998) and Kimmel et al. (2011) show that the cost of correcting data errors rises exponentially at later stages of the workflow, a finding highly relevant to audit engagements. Research in developing countries (Abiola & Oyewole, 2013; Duhan & Singh, 2021) confirms that heterogeneous accounting systems and weak internal controls are major causes of data defects in audit files.

Additionally, organizational culture plays a key role. Strong, Lee, and Wang (1997) emphasize that effective communication between data creators and users determines data reliability. In audit firms, this translates into close coordination between client staff and auditors. Sri Lankan studies (Perera & Chand, 2015; Weerathunga & Liyanage, 2018) highlight those variations in technology adoption and inconsistent ERP systems often contribute to data irregularities.

This review demonstrates that while the issue is global, limited evidence exists from the Sri Lankan audit sector, justifying the relevance of the present case study.

3 METHODOLOGY

This research employed a mixed-method approach combining both quantitative and qualitative analyses to examine data quality challenges faced by audit professionals at ABC Company. The study aimed to measure the frequency and operational impact of client data issues, assess role-based experiences, and evaluate potential solutions to improve data governance and audit efficiency.

3.1 Research Design

The research process was divided into two primary phases.

1. *Quantitative Evaluation* - A structured survey was distributed to audit professionals across four roles: data analysts, internal auditors, external auditors, and IT auditors. The quantitative component measured the occurrence and impact of common data-quality problems. Statistical tools, including descriptive analysis, one-way ANOVA, and Pearson correlation were used to analyze the relationships between data quality issues and workflow delays.

2. *Qualitative Evaluation* - Follow-up interviews with senior staff members were conducted to obtain deeper insights into the root causes of data-quality issues and to evaluate existing mitigation strategies. Qualitative feedback helped validate and interpret the quantitative results, particularly in identifying procedural and client-related challenges.

Because all participants were from a single firm and selected through non-random purposive sampling, results are not generalizable beyond ABC Company. Therefore, the inferential results (ANOVA, chi-square tests, and correlation) reflect patterns within ABC

Company and are not intended to represent causal relationships applicable to the wider auditing industry.

3.2 Tools and Environment

Data collection and analysis were conducted, using a combination of online and statistical tools:

- Google Forms – Used to design and distribute the structured questionnaire to participants.
- Microsoft Excel & IBM SPSS 26 – Used for data cleaning, coding, and performing statistical analyses such as ANOVA and correlation tests.
- Visual Tools – Bar charts, Pareto charts, and scatterplots were generated to display patterns and relationships among data-quality variables.

3.3 Questionnaire Design

The survey instrument was designed to collect both demographic and research-specific data. It included closed-ended and Likert-scale questions to capture measurable responses, along with a few open-ended items for qualitative insights.

The questionnaire consisted of four main sections:

- Demographics – Participant's role, years of experience, and division.
- Nature of Data-Quality Issues – Types and frequency of problems encountered.
- Operational Impact – Estimated extra audit hours and workflow delays.
- Mitigation Practices – Current strategies and tools used to manage data-quality issues.

A pilot test was carried out with five staff members to refine question clarity and structure before full deployment.

3.4 Quantitative Data Collection

Primary data were collected from 52 audit professionals between July and August 2025 using the structured Google Form. Respondents represented four major roles, such as data analysts, internal auditors, external auditors, and IT auditors ensuring diverse perspectives.

Key quantitative variables measured included:

- Frequency of Issues (%): Occurrence rates of missing, duplicate, inconsistent, or corrupted data.
- Additional Audit Hours (hrs): Average time lost due to poor client data.
- Severity Rating (1–5): Perceived level of disruption caused by each issue.
- Descriptive statistics, cross-tabulations, and inferential tests (ANOVA and correlation) were used to analyze results and identify significant patterns.

3.5 Qualitative Data Collection

The qualitative component involved semi-structured interviews with five senior audit staff to explore the underlying reasons for observed data-quality problems. The discussions focused on:

- Client-related Factors: Internal control weaknesses and inconsistent ERP exports.
- Internal Processes: Variations in engagement instructions and the lack of standardized templates.
- Technological Gaps: Low automation and limited use of validation scripts.
- Training and Awareness: Skill gaps among staff and client personnel.

Findings from these interviews were integrated with survey results to form a comprehensive understanding of both the technical and organizational aspects of

data-quality management within ABC Company.

4 RESULTS AND DISCUSSION

4.1 Descriptive Statistics

The first stage of analysis used frequency distributions, means and standard deviations to summarize the prevalence of each data-quality issue. Table 1 presents the distribution of respondents by role and years of experience, illustrating the composition of the sample used for analysis. This breakdown helps establish whether perspectives from different professional groups are adequately represented.

Table 1: Respondent profile by role and years of experience

Role	< 1 year	1–3 years	3–5 years	> 5 years
Data Analyst	04	02	01	0
Internal Auditor	05	07	02	0
External Auditor	12	13	02	01
IT Auditor	02	01	0	0
Total	22	23	05	01

4.2 Role-Based Differences

Cross-tabulation showed clear variations across professions. IT auditors frequently faced corrupt files (~70%), while data analysts reported the highest rate of missing records (~85%), as shown in Figure 1. External auditors mainly encountered inconsistent naming and mismatched templates, which slowed substantive testing.

Chi-square tests confirmed that these differences were statistically significant ($p < 0.05$) for corrupted files and file-format inconsistencies, supporting the first three hypotheses (H1–H3).



Figure 1: Frequency of major data-quality issues by professional role

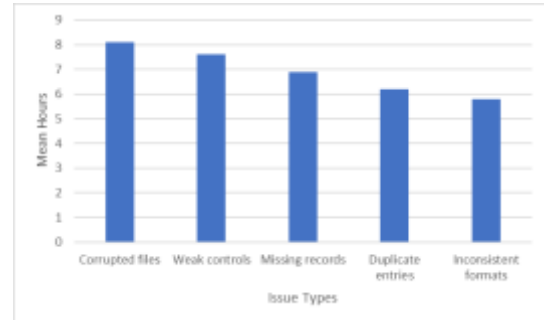


Figure 2: Mean additional hours by type of data-quality issue.

4.3 Impact on Workflow

Respondents experienced an average delay of 6.4 hours per engagement due to data-quality issues. ANOVA indicated significant variation in delays by issue type ($p < 0.01$). Tukey HSD tests showed that corrupted files caused significantly more additional hours than missing or duplicate records ($p < 0.05$). Weak client controls also led to greater delays than inconsistent formats. Figure 2 presents the mean additional hours by issue type.

4.4 Correlation Analysis

Pearson correlation analysis showed a strong positive association between the number of data-quality issues and total delay hours ($r = 0.72$, $p < 0.001$), supporting the hypothesis that poor client data quality disrupts audit workflows. Figure 3 illustrates this relationship using a scatterplot of severity score versus additional hours.

STATISTICAL ANALYSIS OF DATA QUALITY CHALLENGES FACED BY AUDIT PROFESSIONALS: A CASE STUDY IN THE ABC COMPANY

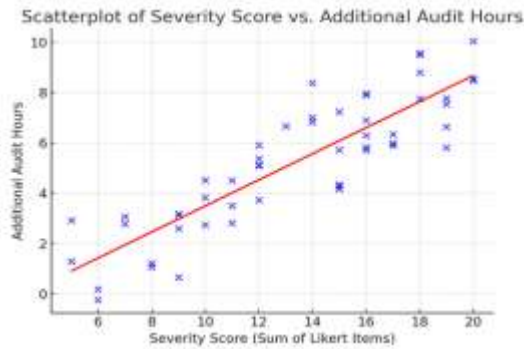


Figure 3: Scatterplot of severity score vs. additional audit hours

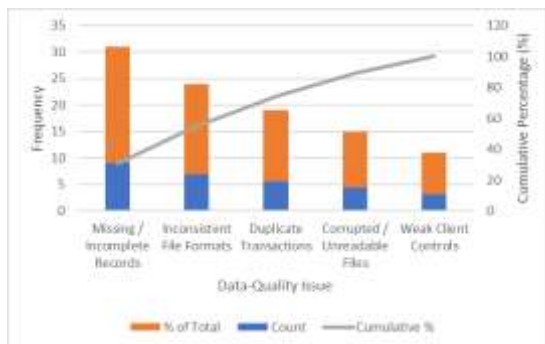


Figure 4: Pareto of Issue Frequencies and Cumulative percentage

These findings should be viewed in the context of ABC Company, as the single-firm sample limits broader generalization.

4.5 Pareto Analysis (80/20 Rule)

Figure 4 illustrates the cumulative contribution of each data-quality issue to overall disruptions. The top three issues, including missing records, duplicate entries, and inconsistent formats, comprise roughly 74% of all reported disruptions, consistent with the 80/20 principle.

4.6 Fishbone (Ishikawa) Diagram

Figure 5 illustrates the five major branches, such as Client Processes, Technology, Internal Instructions, Human Resources and External Constraints each feeding the “Effect” of poor data quality.

The statistical methods applied in this study are interrelated. Descriptive

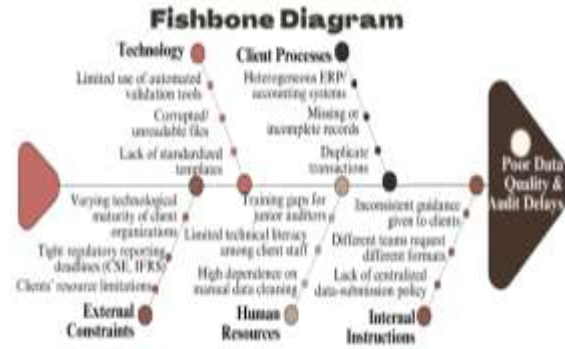


Figure 5: Fishbone (Ishikawa) Diagram

statistics quantified issue frequency, chi-square tests examined role-based differences, ANOVA evaluated variation in workflow delays, and Pearson correlation measured the relationship between total issues and audit hours.

5 CONCLUSION

This study confirms that data-quality challenges significantly affect audit performance. The issues most commonly encountered missing, duplicate, and inconsistent data result in additional audit hours and delayed reporting. Through systematic statistical analysis.

By adopting standardized submission templates and automated validation checks, ABC Company can improve both audit accuracy and operational efficiency. The research underscores the importance of proactive data governance and sets a foundation for future studies exploring automation and predictive analytics.

REFERENCES

- Abiola, J., & Oyewole, F. (2013). An empirical investigation of the impact of accounting records on the performance of small-scale enterprises. *International Journal of Academic Research in Business and Social Sciences*, 3(8), 168–188.
- Al-Zwyalif, I. M. (2015). The role of internal control in enhancing corporate governance: Evidence from Jordan.

International Journal of Business and Management, 10(7), 57–66.

- Appelbaum, D. (2017). Impact of business analytics and enterprise systems on managerial accounting. *International Journal of Accounting Information Systems*, 25, 29–44.
- Boakye, E. (2016). The role of internal control in the public sector: Evidence from Ghana. *International Journal of Financial Research*, 7(2), 38–47.



A Markov Chain-Based Approach for Modelling and Simulating Monsoonal Water Level Variations in Kalu River Basin

Fernando CNNM¹, Ekanayake EMP¹, Appuhamy PADAN¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
nishavinarthana@gmail.com

ABSTRACT

Water level forecasting in river systems of tropical island nations is a challenging task owing to seasonal hydrological variability and complex precipitation. The present study focused on the application of finite-state Markov chain models to characterise water level dynamics across Sri Lanka's four distinct monsoon seasons, with the aim of developing a computationally efficient probabilistic forecast tool for the Kalu River Basin in the Ratnapura District. The methodology involved a systematic analysis of hourly water level data to quantify seasonal transition probabilities, equilibrium state probabilities, expected length of stay in a state, and to validate model performance through stochastic simulation techniques. The findings revealed clear seasonal behaviour in water levels: the Southwest Monsoon (SWM) exhibited the highest vulnerability to flood events, while the First Inter-Monsoon (FIM) period showed no significant water level variations. The study further divulged that, in the long run, the Normal state consistently depicts the highest probability across all seasons, whereas the Alert state ranges from 0.003 to 0.054, and the Minor Flood state is limited to 0.011 during the SWM. Expected length of a stay highlighted persistence in the Normal state, complemented by shorter durations in the Alert state (7.33-19.07 hours). During SWM season, the expected duration of the Minor Flood level is relatively greater than the stay in the Alert level, which is 26.33 hours. These results support the fact that the Markov chain model developed in the study provides better insights into forecasting water level dynamics in Sri Lanka's monsoon seasons.

KEYWORDS: Hydrological variability, Markov chain, Monsoon seasons, Simulation, Transition probability

1 INTRODUCTION

Sri Lanka, a tropical island nation situated within the equatorial region between the Arabian Sea to the west and the Bay of Bengal to the east, is particularly vulnerable to variations in precipitation. Rainfall in Sri Lanka is predominantly influenced by the South Asian monsoon system, which consists of four distinct phases. The country experiences two principal monsoon seasons: the Southwest Monsoon (SWM) from May to September and the Northeast Monsoon (NEM) from December to February, along with two preceding inter-monsoon seasons: the First Inter-Monsoon (FIM) from March

to April and the Second Inter-Monsoon (SIM) from October to November. Based on the rainfall patterns associated with these monsoons, the island has long been considered to comprise two key climatic regions: the wet zone and the dry zone (Rajendram, 2021).

The Kalu River, which lies entirely within Sri Lanka's wet zone, originates from the central hills of Sri Lanka. Ratnapura is positioned at the centre of the Kalu River Basin and lies at the confluence of the Wey Ganga and Kalu River. The region has a diverse biophysical landscape, which includes mountainous areas with steep slopes. It receives notably high annual rainfall, 4,000–5,000mm (Ampitiyawatta & Guo,

2009). These factors account for severe flooding events in Ratnapura, with recent significant occurrences recorded in 2003, 2008, 2010, 2014, 2016, 2017, and 2024. The flood vulnerability in Ratapura is directly linked to the water level and rainfall of the Kalu River (Hettiwaththa & Abeygunawardana, 2018). In relevance to this catastrophic phenomenon, the present study aims to develop finite Markov chain models for the Kalu River water level variations in Ratnapura, based on monsoonal rainfall patterns and to simulate water level transitions to validate the model and generate steady-state probabilities for each monsoon season that support flood risk management and early warning systems.

2 LITERATURE REVIEW

The application of Markov chains in flood/ water level forecasting has appeared in some studies. Yapo et al. (1993) introduced a comprehensive streamflow forecasting methodology employing Markov chains combined with multi-objective optimisation principles. Building on this, San and Khin (2016) developed a three-state Markov model for river flood prediction in the Ayeyarwady River, Myanmar, while Tun and Sein (2019) introduced a more advanced Markov Chain model for flood forecasting at Mone Dam in Myanmar.

However, the use of the Markov chain approach for modelling Sri Lankan river systems, especially the Kalu River, remains unexplored. Recent studies on the Kalu River Basin have employed several approaches, like hydrological, hydrodynamic, statistical, and machine learning (ML). For instance, Nandalal

(2009) developed a hydrodynamic model (HEC-RAS) to determine water levels along the Kalu River. Hettiwaththa & Abeygunawardana (2018) proposed an ARIMAX (2,1,3)-GARCH (1,2) time series model to forecast river water level using rainfall and discharge as exogenous variables. More recently, Sudasingha and Koswatte (2024) compared ML models for Kalu River water-level forecasting. Additionally, a hybrid model combining Bi-LSTM with Time Series Regression was developed for flood forecasting (Saubhagya et al., 2025).

Despite these efforts, several issues arise in current methodologies for the Kalu River, such as high computational demands, large data requirements, and difficulties in interpretability that often exceed the technical capacity of local flood monitoring agencies. The Markov chain technique applied here addresses these limitations by developing season-specific, computationally efficient, yet probabilistically robust models suitable for local operational implementation.

3 METHODOLOGY

A Markov chain is a sequence of random variables $X_0, X_1, X_2, \dots$ taking the values in the state space (S) satisfying the Markov property. That is the conditional probability distribution of X_{n+1} for given values of $X_n, X_{n-1}, \dots, X_0$ depends only on the value of X_n , which is the most recent known value of the process and is denoted as in (1).

$$\begin{aligned} Pr(X_{n+1} = j_{n+1} | X_n = j_n, \dots, X_0 = j_0) \\ = Pr(X_{n+1} = j_{n+1} | X_n = j_n) \end{aligned} \quad (1)$$

; where $j_0, \dots, j_n, j_{n+1} \in S$

3.1 Markov chain modelling

The hourly water level data at the Ratnapura gauge station of the Kalu River Basin were obtained from the

Department of Irrigation, Sri Lanka, for the period spanning from 2020 to 2024. The missing values in them were imputed using the Kalman Smoothing method (Umar & Gray, 2023).

Let X_n denote the water level at the n^{th} hour. As n varies, X_n form a sequence of random variables that can be modelled as a finite Markov chain $\{X_n: n = 0, 1, 2, 3, \dots\}$, assuming that the water level of the next hour depends only on the current level. The process has a discrete state space $S = \{1, 2, 3, 4\}$ and a discrete parameter space $n = 0, 1, 2, 3, \dots$. Table 1 presents the classification of water level (states in S) relative to gauge zero and the mean sea level (MSL), based on operational thresholds established by the Department of Irrigation in Sri Lanka. The Chi-squared test of independence was applied to check the Markov property with the null hypothesis (H_0) that the water level at consecutive hours is independent, versus the alternative (H_1) that it is not independent at a 5% level of significance. To capture the effect of seasonal variations in water level behaviour, four distinct monsoon seasons were modelled separately.

3.2 One-step transition probability

Initially, data were grouped into training (80%) and testing (20%) sets to develop the Markov models and to validate the results, respectively. The one-step transition probability from state i to j is defined in (2),

$P_{ij} = P(X_{n+1} = j | X_n = i); i, j \in S$ (2)
Next, the number of observed transitions from state i to state j (n_{ij}) and the total number of transitions from state i (n_i) were recorded. The maximum likelihood estimator of P_{ij} is defined as,

Table 1: State Classification

State	Water Level (WL)	
	Gauge (in m)	MSL (in m)
Normal (1)	0.0–5.2	11.9–17.1
Alert (2)	5.2–7.5	17.1–19.4
Minor Flood (3)	7.5–9.5	19.4–21.4
Major Flood (4)	> 9.5	> 21.4

$$\hat{P}_{ij} = \frac{n_{ij}}{n_i} ; \text{ for } i, j \in S \quad (3)$$

Then the one-step transition probability matrix ($\mathbf{P}$) was constructed for each season separately, ensuring that it satisfies the properties $0 < P_{ij} < 1$ and $\sum_{j \in S} P_{ij} = 1$ for all $i \in S$.

3.3 Simulation and Validation

A Monte Carlo simulation analysis was then conducted to verify the validity of the Markov models for each season. The simulated dataset for each season was generated using the corresponding $\mathbf{P}$ and matched the size of the testing data. Using the same initial state as in the testing set and performing 10^4 Monte Carlo repetitions, the estimated transition probability matrices ($\hat{\mathbf{P}}$) were obtained. The validity was then quantified using the Mean Squared Error (MSE) as in (4) (Barsotti et al., 2014).

$$MSE = E \|\mathbf{P} - \hat{\mathbf{P}}\|^2 \quad (4)$$

3.4 Long-term behaviour

Equilibrium state probability distribution describes the long-term behaviour of water level states (Grag & Singh, 2010). For an irreducible, aperiodic Markov chain, the equilibrium probability distribution, π , can be obtained from (5).

$$\pi = \pi \mathbf{P} \quad (5)$$

Next, the number of hours (N) required for the system to reach

equilibrium was obtained by powering P until all rows became equal. Let D be a random variable denoting a sequence of consecutive hours spent at each water level state, preceded and followed by other states. Then, it follows a Geometric distribution and has an expected value given in (6) (Grag & Singh, 2010):

$$E(D_i) = \frac{1}{1-P_{ii}} \quad ; (i \in S) \quad (6)$$

4 RESULTS AND DISCUSSION

According to Table 2, which presents the descriptive statistics of water levels for each season, the FIM season showed minimal variation in water levels and remained in the normal state throughout the study period. As a result, the Markov chain modelling was not applicable for this season. Further, no major flood (4) events occurred during any of the three seasons, while the SIM and NEM seasons showed transitions only between the normal (1) and alert (2) states. All p-values of the Chi-square test statistics for the SWM, SIM, and NEM seasons were less than 0.05, resulting in the rejection of the null hypothesis and confirming that the Markov property holds for these three periods. Accordingly, water levels in SWM could be modelled as a three-state Markov chain, whereas water levels in SIM and NEM were modelled as a two-state Markov chain. The estimated transition probability matrices are presented below.

$$P_{(SWM)} = \begin{bmatrix} 0.9975 & 0.0025 & 0 \\ 0.0444 & 0.9475 & 0.0081 \\ 0 & 0.0380 & 0.9620 \end{bmatrix}$$

$$P_{(SIM)} = \begin{bmatrix} 0.9973 & 0.0027 \\ 0.0524 & 0.9476 \end{bmatrix}$$

$$P_{(NEM)} = \begin{bmatrix} 0.9997 & 0.0003 \\ 0.1364 & 0.8636 \end{bmatrix}$$

Table 2: Descriptive Statistics

WL (in m)	FIM	SWM	SIM	NEM
Min	0.00	0.13	0.46	0.13
Max	4.63	9.14	6.93	5.75
Mean	0.938	1.967	2.485	1.047
SD	0.699	1.594	1.365	0.807

The results obtained from the Monte Carlo simulation disclosed that the MSEs for SWM, SIM, and NEM were 0.1999, 0.0189, and 0.1959, respectively. These small MSE values suggest a good fit, confirming that the simulated sequences faithfully reproduce the Markovian transition structure derived from the training data.

According to Table 3, on average, the probability of the water level being in the normal state was considerably higher for all three seasons compared to other states, with the NEM season having the highest value. Moreover, there was a 1.1% risk of the water level transitioning to a minor flood during the SWM season, while that risk was zero for the other two seasons.

5 CONCLUSION

In this study, hourly water level data at the Kalu River Basin in Ratnapura over five years were analysed for the four monsoon periods. It was concluded that the highest water level variation occurred during the SWM season, while the lowest was observed during the FIM. Moreover, it was revealed that there was a greater likelihood of remaining in the same state in the next hour than transitioning to another state. Also, there was no sudden transition between the normal and minor flood levels during the SWM season. Furthermore, during SWM, the

Table 3: Steady State Probabilities, Expected Length, and No. of hours for Equilibrium

Season	π_1	π_2	π_3	Expected length of			N
				Normal	Alert	Minor Flood	
SWM	0.935	0.054	0.011	393.83	19.05	26.33	511
SIM	0.951	0.049	-	372.27	19.07	-	243
NEM	0.997	0.003	-	2883.33	7.33	-	93

probability of the water level being in the normal state was lower, while the likelihood of being in the alert or minor flood level was higher than in the other seasons. Once the water level reached the minor flood level during SWM, it typically remained there for about 26 hours, nearly a full day. In contrast, the number of hours spent in the alert state across all three seasons was less than 20. This indicates that the water level did not remain in the alert state for a full day but tended to transition to the normal state with high probability during the SWM, SIM and NEM seasons, and to minor floods with low probability during the SWM season. This offers reliable probabilistic insights critical for immediate flood response.

REFERENCES

- Ampitiyawatta, A., & Guo, S. (2009). Precipitation trends in the Kalu Ganga Basin in Sri Lanka. *The Journal of Agricultural Science*, 4.
- Barsotti, F., De Castro, Y., Espinasse, T., & Rochet, P. (2014). Estimating the transition matrix of a Markov chain observed at random times. *Statistics & Probability Letters*, 94, 98 - 105.
- Garg, V.K., & Singh, J.B. (2010). Three-State Markov Chain Approach on the Behaviour of Rainfall. *New York Science Journal*, Vol. 3, No. 12.
- Hettiwaththa, H., & Abeygunawardana, A. (2018). Measuring flood risk in Ratnapura town area in Sri Lanka. *2nd International Conference on Climate Change*, 2, pp. 31–41.
- Nandalal, K. (2009). Use of a hydrodynamic model to forecast Floods of Kalu River in Sri Lanka. (P. S. Simonovic, Ed.) *Journal of Flood Risk Management*, 151-158.
- Rajendram, K. (2021). Rainfall variability and drought in the dry and wet zones of Sri Lanka. *World Scientific News*, 160, 172-189.
- San, T.H.H., & Khin, M.M. (2016). River flood prediction using Markov model. *In Genetic and Evolutionary Computing, Advances in Intelligent Systems and Computing*.
- Saubhagya, S., Tilakaratne, C., Lakraj, P., & Mammadov, M. (2025). A Fusion of Deep Learning and Time Series Regression for Flood Forecasting: An Application to the Ratnapura Area Based on the Kalu River Basin in Sri Lanka. *Forecasting*, Vol. 7
- Sudasingha, H., & Koswatte, S. (2024). Realtime Flood Forecasting: River flow analysis using Machine Learning Techniques. *Asian Conference on Remote Sensing (ACRS 2024)*.
- Tun, P.P.P., & Sein, M.M. (2019). Flood Prediction System by using Markov Chain. *9th International Workshop on Computer Science and Engineering*
- Umar, N., & Gray, A. (2023). Comparing Single and Multiple Imputation Approaches for Missing Values in Univariate and Multivariate Water Level Data. *Water*, 15(8), 1519.
- Yapo, P., Sorooshian, S., & Gupta, V. (1993). A Markov Chain Flow Model for Flood Forecasting. *Water Resources Research*, Vol. 29, No. 7.



Evaluation and Enhancement of Sales Forecast Accuracy: A Case Study in a Cement Company

Weerakoon ND¹, Dahanayaka SD¹, Herath CHMU¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
nilmiweerakoon99@gmail.com

ABSTRACT

Accurate sales forecasting is essential for effective production planning, inventory management, and operational efficiency in the cement industry, where demand often shows high volatility and seasonal patterns. This study evaluates the performance of Holt-Winters Exponential Smoothing (HW) and Seasonal ARIMA (SARIMA) models in forecasting weekly cement sales over a 96-week period. The last six months of data have been reserved for out-of-sample testing to assess forecast accuracy. Both models had been implemented using Python, with SARIMA parameters selected through an automated grid search based on the Akaike Information Criterion (AIC). Forecast performance has evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), complemented by visualizations of actual versus forecasted sales to examine the models' ability to capture trend and seasonality. The results indicate that while the HW model effectively captures trend and monthly seasonal fluctuations, the SARIMA (0,1,2)(0,1,1)₄ model consistently outperforms HW across all accuracy metrics. SARIMA successfully models autocorrelation and recurring seasonal patterns, producing lower forecast errors and closer alignment with observed data. Residual diagnostics confirmed the validity of the SARIMA model, showing uncorrelated residuals and stable variance. Based on these findings, SARIMA (0,1,2)(0,1,1)₄ is recommended as the most reliable method for weekly cement sales forecasting. Its implementation can support more accurate production planning, inventory control, and operational efficiency. The study highlights the importance of selecting models that account for both trend and seasonality to generate precise, actionable forecasts in dynamic industrial environments.

KEYWORDS: Cement industry, Forecast accuracy, SARIMA, Sales forecasting, Time series analysis

1 INTRODUCTION

Accurate forecasting is crucial in cement manufacturing due to its impact on production planning, dispatch scheduling, inventory management, and cost control. Although weekly forecasts had used, high errors still lead to stockouts, overstocking, and inefficiencies. The current forecasting system has not been statistically evaluated, highlighting a research gap. This study addresses the lack of a validated model for highly volatile weekly cement demand. Its objective is to assess the accuracy of existing

forecasts and compare them with systematic forecasting models to identify the most suitable short-term forecasting approach.

2 LITERATURE REVIEW

Forecasting plays a significant role in demand planning across industries, with time-series techniques which are being highly adapted due to their ability to model temporal dependencies. Box and Jenkins (1976) introduced the ARIMA framework, emphasizing parameter estimation and autocorrelation structure. Holt-Winters exponential

smoothing is known for its ability to model deterministic seasonal patterns, while SARIMA extends ARIMA to include seasonal differencing and seasonal moving average terms.

Recent literature in the cement industry highlights the importance of selecting forecasting models that account for market volatility. Studies such as Patel et al. (2019) and Padhan (2012) show that SARIMA and exponential smoothing often outperform simpler models when seasonality is strong. However, research also emphasizes the need to incorporate external factors such as economic indicators, competition, and infrastructure development to improve predictive accuracy.

While several studies evaluate forecasting techniques, few explicitly compare internal company forecasts with statistically optimized models. This creates a clear knowledge gap, which the present study aims to address by critically evaluating multiple forecasting approaches using real industrial data.

3 METHODOLOGY

Forecasting had performed using Holt–Winters and SARIMA models. Weekly cement sales data for 96 weeks had been collected, preprocessed, and visually assessed for trend and seasonality. Stationarity had tested using the Augmented Dickey–Fuller (ADF) test. SARIMA model identification had conducted using AIC minimization and residual diagnostics. Forecast accuracy had evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). Additionally, actual versus forecasted sales have plotted to visually examine the models' ability to

capture patterns and fluctuations. The results provide insights into which method produces the most reliable and actionable forecasts for short-term planning.

4 RESULTS AND DISCUSSION

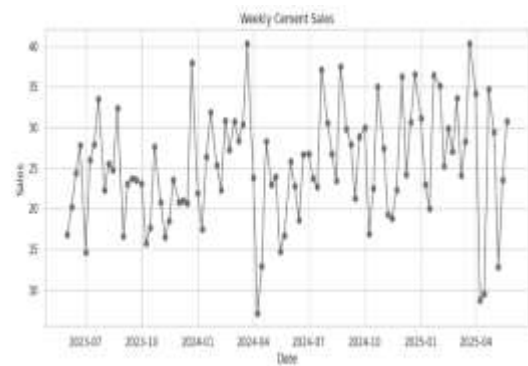


Figure 1: Time Series Plot of Weekly Cement Sales (KMT)

4.1 Data Collection

Weekly cement sales data for 96 weeks (24 months) were obtained from the company, including actual sales (KMT) and company forecasts (KMT). Missing or inconsistent data have corrected, and the dataset has stored chronologically.

4.2 Descriptive Analysis

The Weekly sales show an upward trend with recurring monthly seasonality in fig 1. Summary statistics have shown in n Table 1.

4.3 Trend and Seasonality

Additive decomposition with 4-week seasonality separated the series into trend, seasonal, and residual components. Both trend and monthly seasonality had confirmed in fig 2.

Table 1: Descriptive Statistics for Weekly Sales (KMT)

Statistic	Value (KMT)
Mean	25.198542
Standard Deviation	6.976959

Minimum	7.140000
Maximum	40.370000

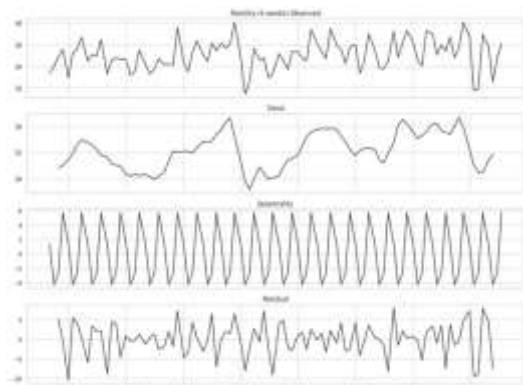


Figure 2:Additive Decomposition (Trend, Seasonal, Residual)

4.4 Stationarity Testing

Table 2 shows the ADF test results, where the original series has a high p-value (0.155) showing non-stationarity, while the seasonally differenced series has a p-value of 0.0002, confirming stationarity. Fig 3 visually supports this result by showing that the differenced series fluctuates around a stable mean with constant variance, demonstrating that seasonal differencing successfully transformed the data into a stationary series.

Table 2: Stationarity Testing Results

Test Step	ADF Statistic	p-value	Decision
Original Series	-2.353	0.155	Non-stationary
After Seasonal Differencing (lag = 4)	-4.452	0.0002	Stationary

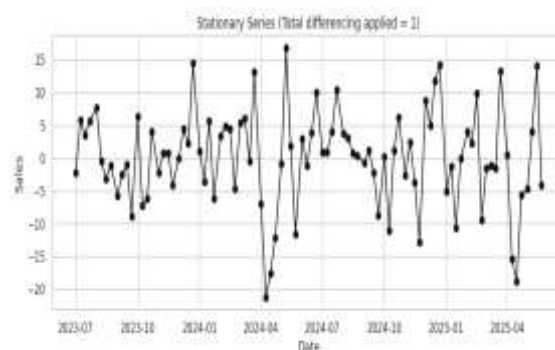


Figure 3:Stationary Series After Differencing



Figure 4:HW Model Forecast vs Actual Sales (KMT)

4.5 Modeling and Forecasting

Holt-Winters (HW) and Seasonal ARIMA (SARIMA) models were applied.

SARIMA parameters were selected using a grid search procedure and evaluated using AIC minimization, as summarized in Table 3.

Table 3: Top SARIMA candidate Models

Model	AIC	p-value
SARIMA(0, 1, 2) (0, 1, 1) ₄	546.52	0.6878
SARIMA(1, 1, 2) (0, 1, 1) ₄	548.49	0.6336
SARIMA(0, 1, 2) (1, 1, 1) ₄	548.52	0.6883
SARIMA(1, 0, 2) (0, 1, 1) ₄	549.14	0.0732

Best model is SARIMA (0, 1, 2)(0, 1, 1)₄

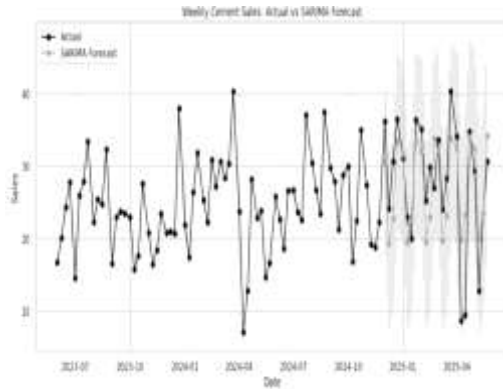


Figure 5: SARIMA Forecast vs Actual Sales (KMT)

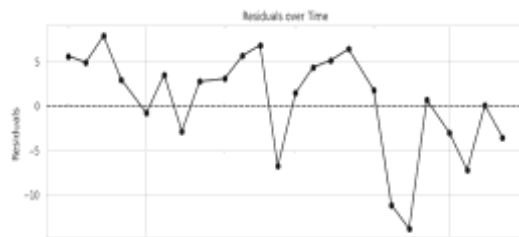


Figure 6: Residual Over Time

Fig 6 shows the residuals plotted over time, where the values fluctuate randomly around zero with no visible trend or pattern. This indicates that the model errors are stable and do not show systematic structure.

Fig 7 presents the distribution of residuals, which appears approximately normal. This suggests that the model errors are random and follow a reasonable error distribution without major skewness or outliers.

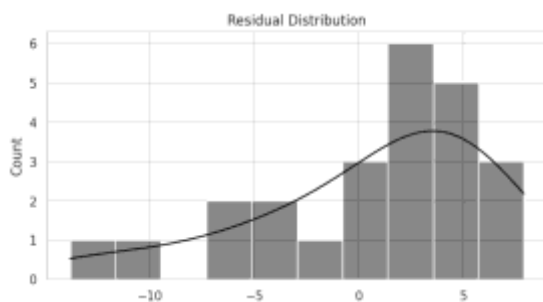


Figure 7: Residual Distribution

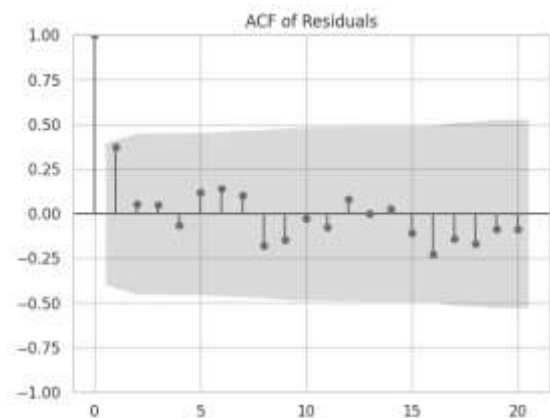


Figure 8:ACF of Residual

Fig 8 shows the autocorrelation function (ACF) of the residuals, where all spikes lie within the confidence limits. This confirms that the residuals have no significant autocorrelation, meaning the SARIMA model has captured all meaningful patterns in the data.

4.6 Forecast Accuracy Comparison

Table 4 presents the forecast accuracy metrics for Holt–Winters, SARIMA, and the company’s existing forecast. SARIMA shows the lowest MAE (4.67 KMT), RMSE (5.66 KMT), and MAPE (25.07%), showing that it provides the most exact predictions. Holt–Winters records the highest error values, while the company’s forecast performs moderately but still worse than SARIMA. This comparison confirms that SARIMA is the most suitable model for weekly cement sales forecasting.

Table 4: Forecast Accuracy Metrics

Model	MAE (KMT)	RMSE (KMT)	MAPE (%)
Holt-Winters	7.74	8.40	31.84
SARIMA	4.67	5.66	25.07
Company Forecast	5.73	8.34	26.67

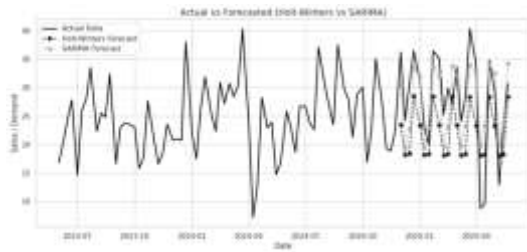


Figure 9:Forecast Comparison – Holt-Winters vs SARIMA

Although SARIMA's MAPE improvement (26.67% to 25.07%) is small, it is still meaningful in a highly volatile weekly sales environment. Even minor error reductions help reduce planning deviations and support more stable production and dispatch decisions. However, since MAPE values above 20% indicate moderate rather than high accuracy, SARIMA should be viewed as the best-performing model among those evaluated, not a highly accurate model in absolute terms.

Fig. 9 shows that SARIMA forecasts follow the actual sales more closely than Holt–Winters, especially during sudden increases or drops in demand. Holt–Winters tends to lag these sharp movements, creating larger gaps between predicted and actual values. This visual pattern supports the lower error values obtained for SARIMA and highlights its ability to capture short-term dynamics and seasonal effects more effectively.

Fig. 10 Fig. 10 further strengthens this conclusion, as SARIMA residuals are smaller and centered tightly around zero, indicating minimal bias and a better model fit. In contrast, Holt–Winters residuals are more scattered, showing that the model leaves more unexplained variation, confirming that SARIMA models both trend and seasonality more accurately and is therefore the more reliable forecasting method for this dataset.

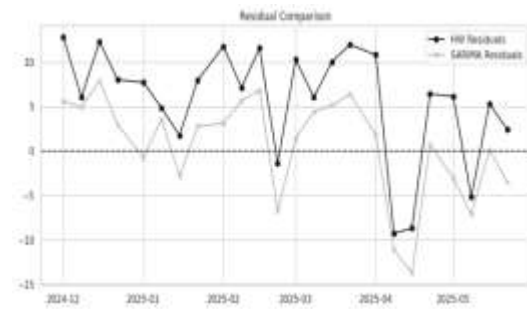


Figure 10: Residual Comparison Between HW and SARIMA Models

5 CONCLUSION

The study proves that SARIMA outperforms both Holt–Winters and the existing company forecast for weekly cement sales. Although accuracy improvements are modest, SARIMA provides better representation of trend and seasonality. Future work should incorporate external variables such as macroeconomic indicators, competitor activity, and industry demand drivers. Hybrid approaches integrating machine learning \with classical time-series models may further enhance forecast performance.

REFERENCES

- Box, G. E. P., & Jenkins, G. M. (1976). Time series analysis: Forecasting and control. Holden-Day.
- Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and practice (2nd ed.). OTexts.
- Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). Forecasting: Methods and applications (3rd ed.). John Wiley & Sons.
- Patel, R., Desai, A., & Patel, B. (2019). Forecasting cement demand using time-series models. *International Journal of Industrial Engineering*, 45(2), 67–75.
- Padhan, P. (2012). Modelling and forecasting of cement demand in India. *Construction Economics Review*, 8(1), 25–34.



Multi-Class Classification of Financial Market Growth and Interest-Level Trends Using CNN and LSTM

Upamali GI¹, Napagoda NADN¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
lmashaupamali093@gmail.com

ABSTRACT

This paper uses deep learning techniques capable of capturing the structural and temporal relationships in socio-economic, technological, and sentiment-based data to categorise global trend performance into Low, Medium, and High. Considering the Global Trend Analysis 2024 dataset, this research leverages appropriate preprocessing and normalisation, coupled with feature engineering, and a model development strategy involving Convolutional Neural Networks and Long Short-Term Memory networks to assess their effectiveness in modelling trend dynamics. Based on the findings, the CNN model returns an accuracy of 81%, while the LSTM model performs even better with 89% by learning sequential dependencies and long-range interactions between features. Such findings indicate that LSTM yields superior classification performance, particularly for medium-performing trends where temporal patterns are more influential. The results confirm that the deep learning models, particularly the LSTM, can be more practically useful for strategic decision-making, market monitoring, and forecasting activities under rapidly changing global conditions and offer a powerful means of interpreting complex trend behaviors. Furthermore, the comparative outcomes highlight the importance of selecting model architectures aligned with data characteristics and temporal complexity. This study also underlines the contribution of engineered interaction features to improved predictive accuracy. Overall, the study indicates that using deep learning architectures with temporal awareness, combined with well-engineered features, leads to better robustness and interpretability of global trend classification within complex, data-driven settings.

KEYWORDS: CNN, Deep Learning, Global Trend Analysis, LSTM

1 INTRODUCTION

The high pace of global technological change and changing trends in the market demand effective means of trend performance classification in order to effectively account for changing market trends. Effective categorization of trends into performance levels, Low, Medium, and High, is necessary for companies and stakeholders in making strategic decisions, best utilize resources, and achieve maximum competitive edge. The dataset includes a rich set of socio-

economic, sentiment, and growth indicators that one may use for this purpose. Regular statistical models are prone to ignoring complex relationships and temporal dependencies within the data. This work addresses this gap with two deep learning models: CNN and LSTM network for multi-class trend performance classification. The CNN excel at learning localized feature patterns, whereas the LSTM excel at learning sequential and temporal dependencies in data. Comparing the models on engineered features like

weekly_interest, region, sentiment, related_brands, and growth_rate, this research aims to show the relative strengths and practical applicability of deep learning in market trend classification. The method highlights the importance of higher neural network architecture in mitigating complexity in trend performance classification in very dynamic market environments.

2 LITERATURE REVIEW

This section shows an array of traditional and machine learning techniques for market trend classification, showcasing the broader methodology landscape outside CNN and LSTM models.

Kumar et al. (2021) compared the application of traditional econometric and machine learning models in the prediction of stock market prices. They indicated that ARIMA and Holt's LES were effective in short-term predictions but possessed less capability to explore nonlinear and unstable market structures common in financial data.

Gulmez (2022) compared the SVR and LSTM models' performance. The study found that SVR could do well in the high-dimensional data, but had more sequential dependence on LSTM. Nevertheless, the study did not account for preprocessing and interaction effects of features that would further improve the performance of the model.

Xu et al. (2024) utilized logistic regression and k-Nearest Neighbours (k-NN) for the prediction of market trends. Logistic regression offered interpretability but failed to address nonlinear relationships, whereas k-NN was effective in nonlinearity but was

computationally expensive and sensitive to outliers when applied to large data.

Campisi et al. (2024) compared SVR and decision tree models and found that nonlinear approaches were superior to linear regression in predicting market directions. Their study focused only on cross-sectional data and neglected time-dependent patterns crucial for financial forecasting.

Wang et al. (2024) examined the effect of feature engineering in deep learning-based classification. By adding interaction terms such as sentiment $\times$ weekly interest to LSTM models, they became more accurate.

Our research is an extension of this work by employing CNN and LSTM models separately, supplemented with hand-designed features. Our approach is focused on improving classification accuracy and robustness through extensive data preprocessing and model fine-tuning to deliver pragmatic and high-accuracy market trend classification tools that facilitate better strategic decision-making.

3 METHODOLOGY

A clearly established process of market trend performance classification using both CNN and LSTM models has been adopted in the research. Numerical feature normalization and one-hot encoding of categorical features were applied. Feature engineering operations, such as extracting temporal features from the date column and creating interaction terms, were conducted.

The CNN model used two Conv1D layers with 64 and 32 filters to extract local feature patterns. This setup helps capture relationships among sentiment, growth rate, and weekly interest.

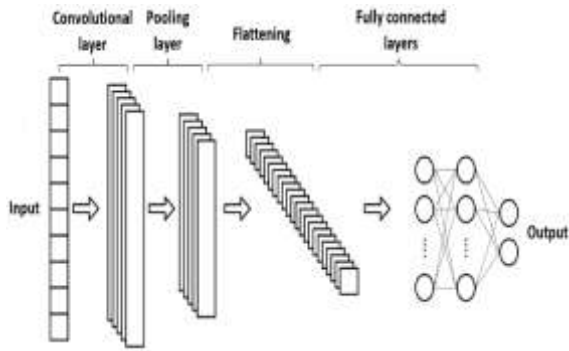


Figure 1 : CNN model architecture

The LSTM model included two layers with 128 and 64 units and a 0.2–0.3 dropout rate to learn sequential financial patterns. A softmax layer outputs the three trend classes.

Both models were optimized with categorical cross-entropy loss and the Adam optimizer. Both models were trained by dividing the dataset into a test set, a validation set, and a training set. Precision, accuracy, recall, F1-score, and confusion matrix performance measures were used to measure predictive accuracy.

Overfitting was prevented by early stopping and cross-validation. The solution utilizes an end-to-end model using the best aspects of CNN and LSTM networks coupled with engineering feature richness and aggressive testing for optimal-in-class classification accuracy and stability.

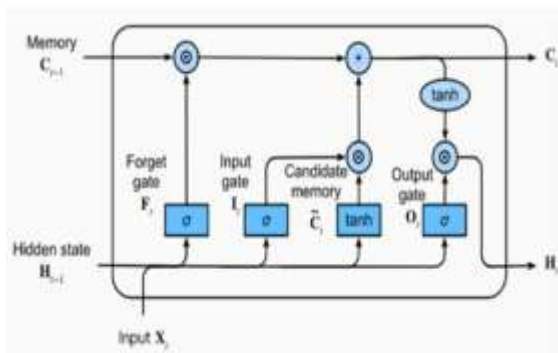


Figure 2: LSTM model architecture

4 RESULTS AND DISCUSSION

Table 1: Classification Report on Test Set.

	prec ision	rec all	F1score	sup port
0	0.78	0.86	0.82	657
1	0.68	0.60	0.64	569
2	0.85	0.80	0.82	774
accuracy			0.81	2000
macro avg	0.77	0.75	0.76	2000
weighted avg	0.80	0.81	0.80	2000

The CNN model achieved an overall accuracy of 0.81. Precision and recall were highest for class 0 and class 2, with lower values for class 1, indicating higher confusability with mid-level classes.

The confusion matrix is highly diagonal dominant, and most of the misclassifications occur between neighbouring classes.

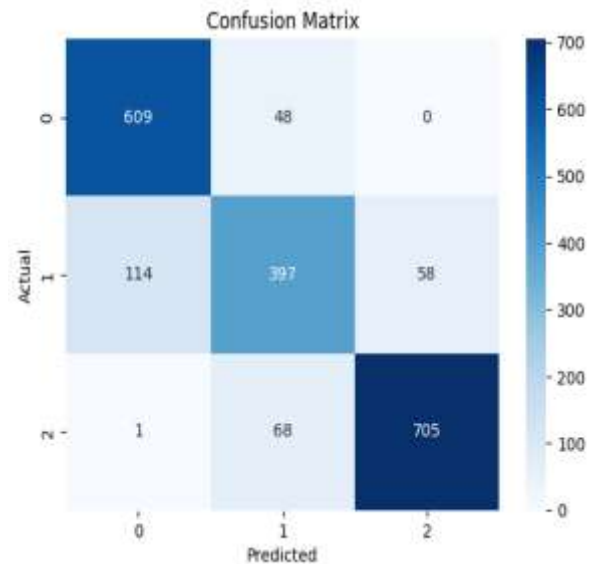


Figure 3: Confusion Matrix for CNN Model.

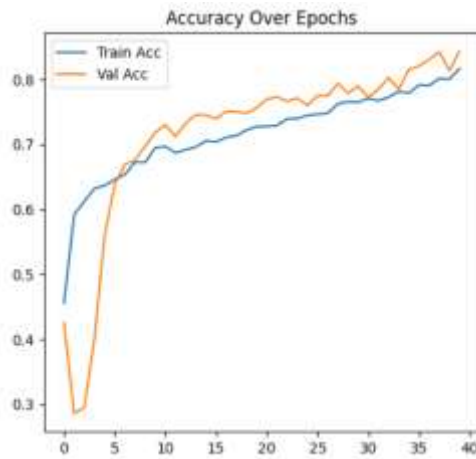


Figure 4: Training and Validation Accuracy over Epochs.

Accuracy over epochs curves demonstrated effective training without significant overfitting.

These results demonstrate that CNNs can be used to disentangle clear trends but might be improved using techniques such as data balancing or feature engineering to increase the classification of problematic cases.

Overall accuracy of the LSTM model for class prediction of market trend was 0.89. Class 0 and class 2 both maintained extremely high precision as well as recall, and hence high F1-scores. Macro-average and weighted-average F1-scores were 0.87 and higher, indicating that all the classes were performing equally well.

Table 2: Classification Report on Test Set.

	prec ision	rec all	F1score	sup port
0	0.89	0.93	0.91	657
1	0.85	0.82	0.83	569
2	0.92	0.90	0.91	774
accuracy			0.89	2000
macro avg	0.89	0.88	0.88	2000
weighted avg	0.89	0.89	0.89	2000

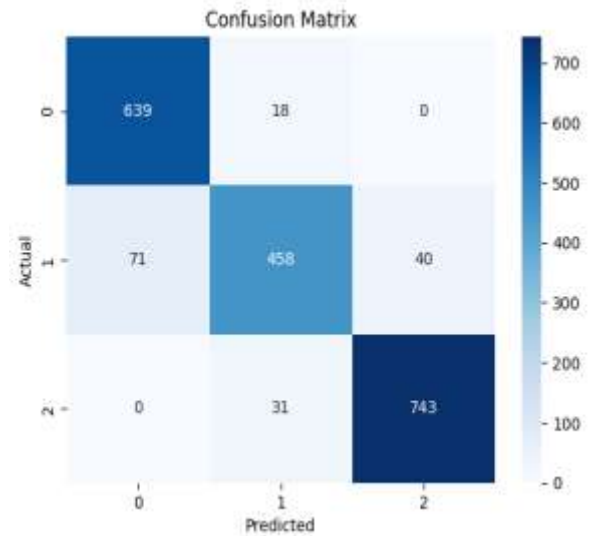


Figure 5: Confusion Matrix for LSTM Model.

The confusion matrix affirms that far more than the majority of predictions were correct, and the overwhelming majority of confusion was between adjacent trend classes.

The training curve for the LSTM model shows that both training and validation accuracy increase quickly and remain consistently high throughout the epochs, reflecting effective learning and strong generalization.

Compared to the CNN model, LSTM performed higher accuracy and recall, particularly in the distinction between the intermediate and high trend classes.

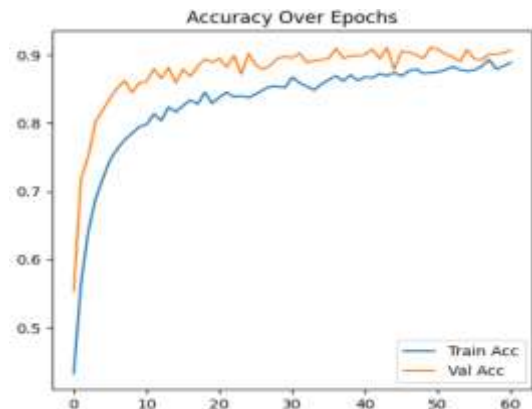


Figure 6: Training and Validation Accuracy over Epochs.

These findings show that LSTM is extremely suitable for market trend

classification with good and stable results for all classes.

5 CONCLUSION

This research focused on classifying market trend performance levels (Low, Medium, High) using deep learning methods, Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). Both models were experimented on the same data of socio-economic, sentiment, and growth-related features to determine their efficiency in pattern identification and temporal learning. The CNN model was 81% accurate, able to detect seemingly degrees of trends by learning spatial and structural pattern features. The LSTM model was at 89% accuracy, outpacing CNN, with its greater ability to model sequential dependencies and long-range relationships in the data.

The performance metrics and visualization analysis confirmed that the LSTM model was stronger and more consistent, with better recall and F1-scores across all classes. Its training process showed steady improvement in accuracy while minimizing overfitting, as reflected by stable validation performance. The CNN model, although less accurate, also produced informative feature extractions and was a good static data pattern recognition baseline.

Both models were very sensitive to interaction features like sentiment $\times$ weekly_interest, which links public opinion with market attention. This helps distinguish strong trends from temporary hype. CNN captures these interactions locally across features, while LSTM tracks their evolution over time, improving the detection of medium and emerging trends.

Overall, this study demonstrates the ability of deep learning models to predict trend performance classification. It is noted in the study that LSTM networks are optimum, while CNNs remain effective for quicker computation and spatial feature extraction.

REFERENCES

- Campisi, D., Gagliardi, D., & Gitto, S. (2024). Comparative analysis of machine learning and traditional models for stock market forecasting. *Journal of Computational Finance and Economics*, 18(2), 101–115.
- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654–669.
- Gulmez, A. (2022). Stock market forecasting using support vector regression and long short-term memory networks: A comparative analysis. *International Journal of Financial Studies*, 10(3), 45–60.
- Kumar, R., Singh, A., & Patel, S. (2021). Forecasting stock market prices using traditional econometric and machine learning techniques. *International Journal of Economics and Financial Research*, 9(4), 87–95.
- Nelson, D. M., Pereira, A. C., & de Oliveira, R. A. (2017). Stock market's price movement prediction with LSTM neural networks. *2017 International Joint Conference on Neural Networks (IJCNN)*, 1419–1426.
- Wang, Y., Zhao, L., & Chen, H. (2024). Feature engineering impact on deep learning-based market trend classification. *Expert Systems with Applications*, 238, 121743.
- Xu, L., Zhang, J., & Li, W. (2024). Market trend prediction using logistic regression and k-nearest neighbours. *Applied Artificial Intelligence*, 38(1), 56–72.



Performance Evaluation of CNN and RNN Architectures in Predicting the Stock Market Price Index

Kaushalya HWS¹, Napagoda NADN¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
shavani12.12@gmail.com

ABSTRACT

The stock market prediction plays a vital role in global finance, impacting international trade, investment decisions, and economic stability. Despite its stability and nonlinear characteristics, achieving accurate prediction remains a challenge. This study analyzes historical stock price data to identify complex temporal and spatial patterns using deep learning approaches. A dataset containing 4,795 daily records for the Dow Jones U.S. Completion Total Stock Market Index (DWCPF Index) is used, and the scale and complexity of this data make it ideal for neural networks. The main objective is to compare and evaluate the predictive capability of two neural network models: Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) for financial stock price data in the stock market. It involves weighted averaging, normalization, rolling window analysis, and model architectures. While the RNN model learns time dependencies, the CNN model aims to learn spatial patterns from sequential market data. Experimental results demonstrate that both models learn well, though the CNN model performed better than the RNN with a high R-Squared (R^2) of 0.93. The CNN model was able to capture complicated spatio-temporal relationships effectively and had strong generalization capability for future DWCPF Index prediction. These results demonstrate how CNN can possibly be used for predicting DWCPF data in the stock market with risk management and algorithmic trading.

KEYWORDS: Convolutional Neural Network, Deep learning, DWCPF Index, Recurrent Neural Networks, Stock Market Prediction

1 INTRODUCTION

The stock market is a dynamic, complex system subject to a wide range of political, psychological, and economic variables. Predicting stock price movements is one of the most challenging problems in financial analysis due to the nonlinear and volatile nature of financial time series data. Traditional methods like statistical models usually have gotten a lot of attention in evaluating the stock prices. Consequently, increasing interest can be observed in making use of data mining and machine learning models for the forecast of changes in the stock price. Deep learning models are now effective instruments for uncovering hidden

patterns in sizable financial datasets. The ability to recognize complex patterns and relations makes neural networks an appropriate tool for financial forecasting in recent years.

Recently, Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs) have been extensively applied in financial forecasting. Sequential information and temporal dependencies are captured by RNNs, whereas local patterns and spatial features in time series data are extracted by CNNs. Despite their successes, the selection of the most suitable model for financial prediction remains an active area of research.

The primary objective of this study is to develop and evaluate neural network models for predicting the Dow Jones U.S. Completion Total Stock Market Index (^DWCPF) using historical daily data comprising 4,795 observations. Although neural networks have been widely applied to financial forecasting, no studies have focused on comparing CNN and RNN models independently using a common dataset, highlighting a clear gap in directly comparing these two architectures on the same stock price dataset. Insights gained from this research are expected to contribute to the understanding of deep learning applications in financial forecasting and provide guidance for researchers in the selection of appropriate neural network models.

2 LITERATURE REVIEW

In this section, a detailed overview of the key findings of previous research studies is presented. Recent studies founded on such models are summarised, and their intent, approach, results, and limitations for stock price forecasting are examined.

Bhardwaj (2021) used one-dimensional time-series data from the National Stock Exchange Fifty Index (NIFTY 50 index) to study the use of CNNs to forecast the direction of stock prices. It was found that CNNs are good at recognizing the upward trend in the market, but only a one-dimensional CNN was employed, which is less effective at capturing temporal dependencies.

Zhu (2020) used 2009-2020 daily data and an RNN model to forecast the stock prices of Apple. The model successfully identified short-term stock trends with reasonable loss values and

more than 95% accuracy. RNNs work well for short-term stock forecasting but need further fine-tuning for longer prediction horizons, concludes the study.

Bao et al. (2025) carried out a systematic survey of data-driven stock forecasting neural network models during the period of 2015-2023. The research took into consideration conventional and advanced neural models such as RNNs, CNNs, Transformers, Graph Neural Networks (GNNs), Generative Adversarial Networks (GANs), and Large Language Models (LLMs). Lastly, the authors enumerated the main research issues, such as model interpretability, overfitting, and efficiency in computation.

Barua et al. (2024) in comparative studies compared deep models for predicting Indian market stock prices. RNN, CNN, Gated Recurrent Units (GRU), Long Short Term Memory (LSTM), and Attention-based LSTM models were utilized in this research for forecasting the values of big firms. It was found that the CNN model provided an effective compromise, and the study was largely focused on point forecasting without consideration of uncertainty estimation or interval forecasting.

Mao (2023) investigated the use of CNN in stock price prediction and compared its performance with LSTM and Conv1D-LSTM models on 1,304 days of China stock data. CNN was found to have the lowest Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) values, suggesting greater prediction accuracy. The study evaluated only RMSE and MAE, ignoring R^2 and overfitting risk.

A review of the literature indicates that there are no studies related specifically to the standalone use of CNN and RNN models together. It concludes that there is a gap in directly comparing these two architectures on this dataset. The traditional statistical methods are suitable under some assumption-based functions, or can perform even worse in very volatile and large-scale financial contexts. However, the neural networks are able to learn these patterns automatically from the big data. Comparative analysis of RNN and CNN models provides decision-making information to investors, economists, and policy-makers to select appropriate tools to forecast stock prices.

3 METHODOLOGY

In this study, DWCPF data from Yahoo Finance is utilized, spanning the temporal range from August 24, 2006, to September 28, 2025. The daily stock market values are represented by Open, High, Low, and Close columns. *Balanced_Weighted* is employed as the target variable for prediction, and using the balanced weighted average approach, as in (1), to combine Open, High, Low, and Close (OHLC) values into a single weighted measure.

$$\text{Balanced\_Weighted} = \frac{(\text{Open} + \text{High} + \text{Low} + \text{Close})}{4} \quad (1)$$

Initially, a systematic preprocessing part is done to ensure the dataset's reliability and quality. It is initiated with checking missing values, duplicate records, zero values, and statistical outliers. To maintain chronological order, the Date column is assigned as the index, and the Date column is set as the

index and data is normalized using Min-Max scaling.

Then, the architecture of an RNN consists of recurrent layers followed by fully connected layers, and the architecture of a CNN consists of multiple layers: convolutional layers, pooling layers, and fully connected layers. As shown in Table 1, the two models are proposed.

Table 1: Architectures of RNN and CNN

Characteristic	CNN	RNN
Input window	30 days	30 days
Layers	Conv2D (2), MaxPool (2), Dense(2), Dropout, Flatten	Simple RNN(2), Dense(2), Dropout
Activation functions	ReLU (hidden), Linear (output)	Tanh (RNN), ReLU (hidden Dense), Linear (output)
Regularization	Dropout (0.4)	Dropout (0.2 & 0.3)

Then, for generating multi-step forecast data, a rolling window technique is applied. Each output represents the future 10 business days, and each input comprises the previously chosen window size. The dataset is split into training, validation, and test sets of 70%, 15%, and 15% in order to preserve temporal consistency and avoid data leakage.

Model training is conducted with early stopping that is based on validation loss optimization. Further, Model performance is evaluated with MAE,

MSE, and R-Squared (R^2). Finally, both models forecast the next ten business days, with results and loss curves are plotted to assess convergence and stability.

4 RESULTS AND DISCUSSION

The Preprocessing step is done, and we made sure that there were no missing values, zeros, outliers, or duplicate records. It reflects that the data set is complete, consistent, and anomaly-free. While performing rolling window analysis, a 30-days window size has been selected as the optimal window size, as shown in Table 2.

According to Table 3, depending on the performance of the two proposed neural network models, the RNN model has an MAE of 0.0444, an MSE of 0.0031, and an R^2 of 0.76. The CNN model, on the other hand, has an MAE of 0.0242, an MSE of 0.0010, and an R^2 of 0.93, signifying great predictive precision and excellent generalization capability.

Table 2: Results of Rolling Window Analysis

Window Size (Days)	MSE	MAE	R^2
30	0.0009	0.0242	0.9281
60	0.0013	0.0286	0.9017
90	0.0017	0.0319	0.8834
120	0.0022	0.0351	0.8595

Table 3: Summary of the RNN & CNN models

Model	MSE	MAE	R^2
RNN	0.0444	0.0031	0.76
CNN	0.0242	0.0010	0.93

In Fig. 1, the values predicted by the RNN are more volatile, which demonstrates sensitivity to noise and overfitting to close temporal patterns. In Fig. 2, the steady performance indicates the CNN's ability to understand the complex spatial-temporal relationship built into sequences of stock prices. The prediction over the next ten business days is steady and uniform, which demonstrates CNN's ability to make accurate predictions for short-term events according to Table 4, Fig. 3 and 4.

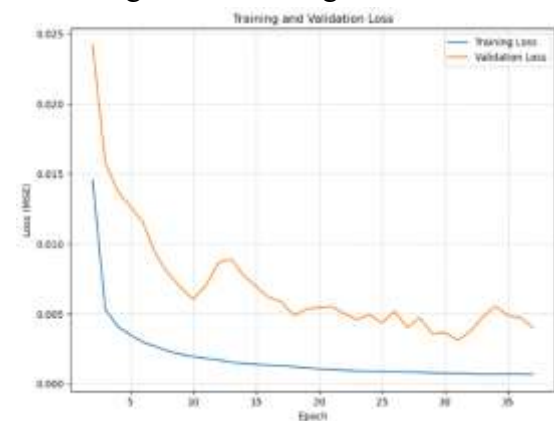


Figure 1: Training History for RNN model

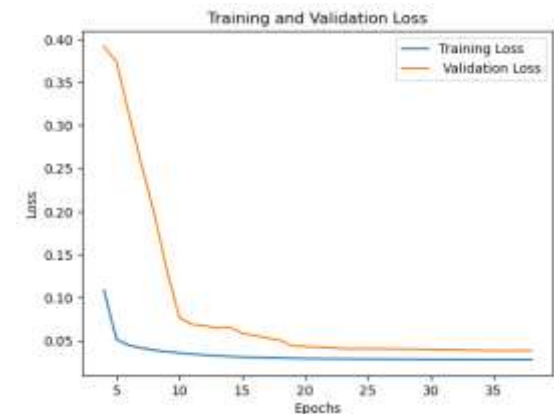


Figure 2: Training History for CNN model

Table 4: Predicted values of RNN & CNN models

Date (2025)	Actual	RNN	CNN
29-Sep	2502.63	2242.30	2428.26
30-Sep	2490.64	2213.87	2432.05

PERFORMANCE EVALUATION OF NEURAL NETWORK ARCHITECTURES IN PREDICTING STOCK PRICES

01-Oct	2493.78	2263.59	2435.77
02-Oct	2510.11	2245.57	2439.06
03-Oct	2527.84	2248.18	2429.22
06-Oct	2539.58	2314.89	2436.26
07-Oct	2519.52	2259.16	2436.89
08-Oct	2524.80	2190.64	2446.88
09-Oct	2526.22	2173.68	2445.20
10-Oct	2485.49	2277.33	2443.15



Figure 3: Actual vs Predicted Values Over Time for the RNN model



Figure 4: Actual vs Predicted Values Over Time for CNN model

Comparatively, the CNN model outperforms the RNN in all the performance measures. The above superior performance of the CNN is owing to the CNN's ability to learn multi-dimensional features and handle stock price volatility in data better. Besides, the higher R^2 value (0.93) indicates CNN's

high explanatory power and generalization predictability. The above results validate that the convolutional architecture is more suitable for stock rate prediction tasks compared to the traditional recurrent design.

4 CONCLUSION

RNN and CNN performance for stock market forecasting here is compared here utilizing the DWCPF index dataset with 4,795 observations of daily data. Both models are found to learn patterns from the data relevant to the task. However, CNN is found to be a more accurate predictor with lower values of MAE (0.0242) and MSE (0.0010) and a higher value of R^2 equal to 0.93 over that of RNN equal to 0.76. The capacity of the CNN to learn spatial and temporal patterns from DWCPF stock price data is shown to result in more stable and accurate short-term forecasts. Relative to the RNN, it is found that the RNN captures temporal relationships but exhibits higher volatility, indicative of higher sensitivity to noise and possible overfitting.

Thus, it is inferred from these findings that CNN models are better equipped for short-term stock price forecasting, particularly for highly nonlinear and unstable DWCPF stock price data. The need for deep learning in stock prices is emphasized because it enables better data-driven and more efficient forecasting. Future work should explore hybrid models involving the integration of RNNs or CNNs with deeper more deep learning techniques for more accurate stock price forecasting of the future in the long run.

REFERENCES

- Bao, W., Cao, Y., Yang, Y., Che, H., Huang, J., & Wen, S. (2025). Data-driven stock forecasting models based on neural networks: A review. *Information Fusion*, 113, 102616.
- Barua, M., Kumar, T., Raj, K., & Roy, A. M. (2024). Comparative analysis of deep learning models for stock price prediction in the Indian market. *FinTech*, 3(4), 551-568.
- Bhardwaj, K. (2021). Convolutional neural network (cnn/ConvNet) in stock price movement prediction. *arXiv preprint arXiv:2106.01920*.
- *Dow Jones U.S. Completion Total Stock Market Index (^DWCPF) Historical Data - Yahoo Finance*. (2025). Yahoo Finance.
- Mao, L. (2023). Analysis of Forecasting Stock Prices Using a CNN Model. *Academic Journal of Computing & Information Science*, 6(10), 79-84.
- Zhu, Y. (2020). Stock price prediction using the RNN model. *Journal of Physics: Conference Series*, 1650, 032103.



Time series analysis and forecasting of Sri Lanka's electricity usage: Statistical vs Machine learning models

Shehari KMDN¹, Francisco GS¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
nimshakulathunga@gmail.com

ABSTRACT

Accurate forecasting of electricity consumption is essential for effective operational planning, infrastructure development, and policy formulation in Sri Lanka's power sector. This study conducts a comprehensive comparative analysis of statistical and machine learning approaches for forecasting sector-wise monthly electricity consumption across 15 consumer categories from 1989 to 2024. The methodology integrates a multivariate Vector Autoregression (VAR) model with three machine learning (ML) techniques: Random Forest, XGBoost, and Long Short-Term Memory (LSTM) neural networks to capture both linear interdependencies and nonlinear temporal patterns in the data. Extensive preprocessing, including stationarity testing, differencing, lag selection, and feature engineering, was performed to enhance model performance. Model performance was evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the Coefficient of Determination (R^2). Results indicate that while the VAR model performs reasonably well for sectors exhibiting linear relationships, machine learning models, particularly the LSTM network, demonstrated superior predictive accuracy. The LSTM achieved the lowest RMSE (9.249) and MAE (5.185), along with the highest R^2 (0.992), highlighting its effectiveness in modelling complex, long-term dependencies in multivariate electricity consumption. Forecasts generated for 2025–2026 reveal sector-specific consumption trends with implications for grid planning, demand-side management, and national energy policy. The study underscores the value of deep learning techniques for high-precision electricity forecasting and provides actionable insights for utility providers and policymakers in Sri Lanka's evolving energy landscape.

KEYWORDS: Electricity Forecasting, Machine Learning Models, Multivariate Time Series, Sri Lanka, Statistical Model

1. INTRODUCTION

Sri Lanka has achieved nearly 100% electrification, with electricity sales reaching approximately 15,191 GWh in 2024. The Ceylon Electricity Board (CEB) and Lanka Electricity Company (LECO) are the primary providers, sourcing energy from hydro, thermal, and non-conventional renewable sources. Accurate forecasting of electricity consumption is paramount for revenue prediction, infrastructure investment, policy formulation, and ensuring grid stability. The 2022 power crisis, driven by economic constraints,

underscored the critical need for reliable demand forecasts to prevent such disruptions.

Traditional statistical models like ARIMA and VAR have been widely used for time series forecasting. However, they often struggle to capture the nonlinear and complex dynamics inherent in multi-sectoral electricity data. Modern ML models, such as XGBoost and LSTM networks, offer a promising alternative due to their ability to model complex patterns without strict linear assumptions.

1.1 Objectives

The primary objective is to identify the most accurate forecasting model for Sri Lanka's electricity consumption by comparing the performance of the VAR model with three machine learning models, thereby supporting data-driven decision-making in the energy sector. The secondary objective is to generate sector-wise electricity consumption forecasts for the next 24 months using the best model.

2. LITERATURE REVIEW

2.1 Forecasting in the Sri Lankan Context

Research on electricity forecasting in Sri Lanka has primarily relied on statistical models. A key study by Karunarathna and Vithyasangaran (2017) employed a VAR model to forecast sectoral demand, effectively capturing the interdependencies between three sectors and demonstrating the value of multivariate analysis for national planning. Fernando et al. (2017) applied the Box-Jenkins methodology, identifying optimal ARIMA models for long term electricity production and consumption. Similarly, Pathberiya and Dias (2013) compared statistical techniques, finding that ARIMA outperformed classical decomposition and exponential smoothing for short-term forecasting of Sri Lankan electricity sales. Collectively, these studies provide a strong baseline but are predominantly confined to linear statistical models, leaving the potential of machine learning largely unexplored in the local context.

2.2 Global Trends and Machine Learning Advances

Alim et al. (2020) demonstrated that XGBoost significantly outperformed ARIMA in forecasting a nonlinear, seasonal disease series, highlighting the advantage of tree-based models for complex temporal data. In energy-specific applications, Mishra et al. (2024) found that while ARIMA and XGBoost were competent, deep learning models, particularly LSTM and CNN-LSTM hybrids, achieved superior accuracy in forecasting volatile electricity prices due to their ability to model long-term dependencies and nonlinearities.

The efficacy of neural networks is further supported by Hussein and Awad (2024), who developed a hybrid NARX model for electricity consumption. Tessonni and Amoretti (2022) showed that statistical models like VARMA can remain highly competitive with RNNs in multivariate forecasting tasks. This study aims to bridge this gap by conducting a rigorous empirical comparison between a VAR model and multiple ML models to determine the most accurate forecasting framework for Sri Lanka's sector-wise electricity consumption, thereby contributing a data-driven model selection guideline for the country's energy sector.

3. METHODOLOGY

3.1 Data Collection and Preprocessing

The dataset comprises 432 monthly observations (January 1989 to December 2024) of electricity sales in GWh across 15 sectors. Data was sourced from the Ceylon Electricity Board's statistical books and the Ceylon Electricity Board website. The data was preprocessed by formatting the dates into a consistent structure and handling missing values using Excel-based techniques such as

filtering and replacing empty entries with zero to ensure accuracy and readiness for analysis. For non-stationary series, first-order differencing was applied, confirmed by the Augmented Dickey-Fuller (ADF) test.

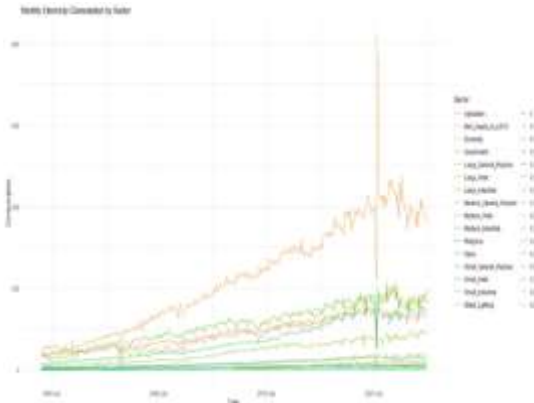


Figure 1: Overall time series plot

Figure 1 shows the overall time series plot according to each sector. Each colored line represents a distinct sector, revealing diverse consumption behaviors over time. The Domestic sector shows the most pronounced growth, while Bulk Supply to LECO, Industrial, and General-purpose sectors also exhibit steady upward trends. In contrast, sectors like Agriculture, Religious, and Small Hotels remain relatively flat, indicating low and stable demand. Notable spikes and drops, particularly around 2020, suggest external disruptions such as policy shifts or economic events.

3.2 Forecasting Models

Four forecasting models were selected based on their proven ability to handle multivariate and nonlinear electricity consumption data. The Vector Autoregression (VAR) model was chosen because it is specifically designed to capture linear interdependencies among multiple time series, making it well-suited for sector-wise electricity data where cross-sector influences are expected. The machine learning models

Random Forest (RF), XGBoost, and Long Short-Term Memory (LSTM) networks were selected to evaluate nonlinear forecasting capabilities. Random Forest provides robust predictions by averaging multiple decorrelated decision trees, thereby reducing variance and improving generalization. XGBoost was included due to its gradient-boosting architecture, which iteratively minimizes errors using optimized tree structures and is well known for superior performance in structured tabular time series. LSTM networks were selected because they incorporate memory gates that allow the model to retain long-term temporal dependencies, making them particularly suitable for sequential electricity data with seasonal and nonlinear patterns.

3.3 Model-Building Process and Evaluation

Lagged variables were generated ($t-1$, $t-2$, ..., $t-12$) for all models to capture short-term and seasonal temporal dynamics. The VAR model was constructed by first testing for stationarity using the Augmented Dickey-Fuller (ADF) test, followed by first-order differencing for non-stationary series. The optimal lag length was selected using the Akaike Information Criterion (AIC).

For the Random Forest model, the number of trees, maximum depth, and minimum samples per split were tuned using grid search on the training set. XGBoost was trained using a tuned learning rate, maximum tree depth and subsampling rate to balance bias and variance, with early stopping applied to prevent overfitting. The LSTM model required reshaping the input data into a three-dimensional structure (samples,

timesteps, features). A sequential architecture was implemented with an LSTM layer and dropout regularization. Training was conducted over multiple epochs, and early stopping was applied when validation loss failed to improve.

All models were trained using an 80/20 chronological split. Performance was compared using RMSE, MAE, MAPE, and R^2 to ensure consistent benchmarking across linear and nonlinear model classes.

3.4 Software

Analysis was conducted using R for the VAR model and Google Colab (Python) for the machine learning models, while Microsoft Excel was used to build and preprocess the dataset before model development.

4. RESULT AND DISCUSSION

4.1 Descriptive Analysis and Stationarity

Descriptive statistics revealed considerable variation in electricity consumption across sectors. The Domestic sector showed the highest volatility (Standard Deviation = 128.23 GWh), while sectors such as Religious and Agriculture remained relatively stable. Time series plots and Augmented Dickey–Fuller (ADF) tests indicated that most series were non-stationary. First-order differencing was applied to achieve stationarity before modelling.

4.2 Model Performance Comparison

Table 1 summarizes the forecasting performance of the VAR and the three ML models. Among all models, LSTM recorded the lowest RMSE and MAE, along with the highest R^2 value (0.992), demonstrating strong predictive accuracy. Its relatively high MAPE is largely due to sensitivity to

sectors with very low consumption values.

Table 1: Overall model performance comparison

Model	RMS E	MA E	MAPE (%)	R^2
VAR	16.35 6	7.65 9	116.06 6	0.97 4
Random Forest	11.04 0	6.37 0	15.939 8	0.98 8
XGBoo st	14.78 8	7.60 8	15.164 9	0.97 9
LSTM	9.249 5	5.18 5	84.866 2	0.99 2

A VAR model has a significantly higher Mean Absolute Percentage Error (MAPE) than the XGBoost model, even if other metrics are comparable, due to fundamental differences in how the two model types handle data characteristics. The core reasons are the VAR model's limitations with non-linear relationships and its sensitivity to low or zero values in the data.

4.3 Sector-Wise Forecasts and Insights

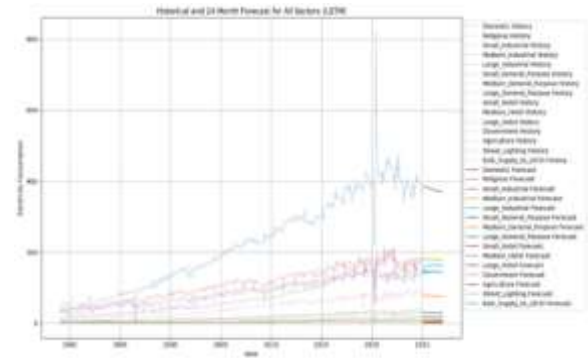


Figure 2: Sector-wise forecasted plot

The LSTM's dominance in RMSE, MAE, and R^2 provides a more robust validation of its accuracy, establishing it as the optimal model for this task. In Figure 2, the results indicate strong growth in sectors such as Large Industrial, Small General Purpose, and Street Lighting, suggesting continued expansion in these areas. Government and Hotel categories exhibit stable or moderately increasing consumption,

reflecting consistent operational needs. In contrast, the Domestic and Medium Industrial sectors display a slight downward trend, which may be attributed to market saturation, energy efficiency improvements, or behavioral changes. The Agriculture and Small Industrial sectors are expected to remain relatively flat, showing minimal variation over the forecast period.

5. CONCLUSION

This study successfully compared statistical and ML models for forecasting Sri Lanka's sector-wise electricity consumption. The results clearly demonstrate that the LSTM model provides the most accurate forecasts, exploiting its ability to capture complex, nonlinear temporal dependencies across multiple sectors. While tree-based models also performed robustly and offer greater interpretability, the LSTM's superior R^2 and error metrics make it the best fit for this multivariate time series task.

Future studies should incorporate statistical significance tests such as the Diebold Mariano test to formally compare forecasting accuracy across models. Since the dataset is limited (432 observations), robustness checks including rolling origin time-series cross validation and alternative validation schemes should also be applied. The very high R^2 of the LSTM model suggests possible overfitting; therefore, future work should explore regularization techniques, simplified architectures, and cross validation based hyperparameter tuning. Expanding the evaluation framework with additional error measures and using richer datasets or exogenous variables may further enhance

the reliability and generalizability of the forecasting models.

REFERENCES

- Alim, M., Ye, G. H., Guan, P., Huang, D. S., Zhou, B. S., & Wu, W. (2020). Comparison of ARIMA model and XGBoost model for prediction of human brucellosis in mainland China: a time-series study. *BMJ open*, 10(12), e039676.
- Fernando, J. L. L. S., Gunawardana, J. R. N. A., Perera, K. A. I. T., Perera, M. L. D. M., Shashikala, M. A. G., Jayasundara, D. D. M., & Rathnayaka, R. M. (2017). *Time Series Modelling approach for forecasting Electricity Demand in Sri Lanka*.
- Hussein, A., & Awad, M. (2024). Time series forecasting of electricity consumption using a hybrid model of recurrent neural networks and genetic algorithms. *Measurement: Energy*, 2, 100004.
- Karunarathna, K. A. N. K., & Vithyasangan, K. (2017). Modelling Sector-wise Electricity Demand in Sri Lanka: An Application of the Vector Autoregressive Model. *Asian Res. J. Math*, 3(4), 1-10.
- Mishra, B. K., Preniqi, V., Thakker, D., & Feigl, E. (2024). Machine learning and deep learning prediction models for time-series: a comparative analytical study for the use case of the UK short-term electricity price prediction. *Discover Internet of Things*, 4(1), 24.
- Pathberiya, H. A., & Dias, P. (2013). *Modelling electricity sales in Sri Lanka and Colombo city using different statistical methods*.
- Tesson, V., & Amoretti, M. (2022). Advanced statistical and machine learning methods for multi-step multivariate time series forecasting in predictive maintenance. *Procedia Computer Science*, 200, 748-757.



Investigating the Root Causes of High Glove Rejection Rates in the Thumb Crotch Reinforcement (TCR) Machine

Shehari KMDN¹, Karunarathna SPDSSK²

¹Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
nimshakulathunga@gmail.com¹

²Department of Mathematics, Faculty of Science, University of Kelaniya
sanjeewak@kln.ac.lk²

ABSTRACT

This study investigates the underlying causes of product rejection in the Thumb Crotch Reinforcement (TCR) process within ABC glove manufacturing company, aiming to identify statistically significant defect types and propose feasible corrective actions. Using multiple linear regression analysis on daily production data from the packing and printing department of a glove manufacturing company, the research quantifies the impact of various defect categories on the overall reject rate. Diagnostic tests such as residual plots, normality checks, and multicollinearity assessments (VIF) were employed to validate model assumptions. Exploratory data analysis identified Style_01 as the major source of quality issues, with a peak rejection rate of 16%, indicating process instability. Correlation analysis revealed that Compound Splash and Bubbles defects were most strongly associated with rejection rates ($r \approx 0.57$). A multiple linear regression model, refined through square root transformation, explained 61.83% of the variation in rejection rates ($p < 2.2e-16$). The model highlighted Burn, Compound Splash, and other defects as the most significant contributors, suggesting that thermal control and material handling are critical areas for quality improvement. Despite minor assumption violations, the model provided actionable insights for quality improvement. The study recommends machinery calibration, compound standardization, and sensor-based process monitoring as key strategies to mitigate rejections and enhance production efficiency.

KEYWORDS: Defects, Multiple Linear Regression, Rejection Rate, TCR Process

1. INTRODUCTION

In glove manufacturing, maintaining a low product rejection rate is critical for cost control and operational efficiency. The Thumb Crotch Reinforcement (TCR) process at ABC Glove Manufacturing currently exhibits a 4% rejection rate, which is four times higher than the company's 1% target and double the typical industry benchmark. This high rejection level leads to significant material waste and financial loss. This study aims to identify the key factors influencing the TCR rejection rate using regression-based analysis. The ultimate goal is to propose corrective

actions that can optimize quality and minimize production waste.

1.1 Objectives

This study aims to systematically reduce the 4% rejection rate in the TCR process by first identifying the most problematic glove style through exploratory data analysis, and then employing multiple linear regression to statistically quantify the impact of key defect types such as Burn, Compound Splash, and Bubbles on the rejection rate. The final objective is to determine the most significant root causes and based on this quantitative evidence propose a prioritized set of actionable solutions, including machinery calibration and

process standardization, to effectively mitigate the high rejection rate.

2. LITERATURE REVIEW

Defect minimization remains a critical priority in modern manufacturing, as high defect rates directly influence production costs, resource utilization, and product quality. To systematically address these challenges, many industries rely on structured quality-improvement methodologies particularly Six Sigma, Lean principles, and regression-based analytical frameworks. These approaches provide evidence-based strategies for identifying sources of variation, optimizing processes, and reducing waste.

A significant body of research has demonstrated the effectiveness of integrating statistical tools with Six Sigma in manufacturing environments. Evangelyne et al. (2018) highlighted the strong synergy between the DMAIC cycle and statistical techniques such as ANOVA and regression analysis, showing how data-driven diagnostics can substantially reduce process variability. Similarly, Nurprihatin et al. (2023) applied linear regression within a Six Sigma framework to quantify the impact of human-related factors on product defects, illustrating how statistical modeling can support targeted and sustainable improvements in production quality. Beyond manufacturing, Dinulescu et al. (2018) demonstrated that Lean Six Sigma principles can enhance operational efficiency in public systems as well, underscoring the versatility and scalability of these methods.

Uyanik and Güler (2013) demonstrated how multiple linear regression (MLR) can identify influential

predictors in complex datasets, highlighting the need for validating core assumptions such as linearity, normality, and multicollinearity to ensure robust and reliable inference. Because manufacturing processes often involve correlated variables, multicollinearity becomes a common challenge. Perez-Melo and Kibria (2020) addressed this issue by demonstrating that ridge regression and other regularization techniques produce more stable coefficient estimates when predictors are highly correlated, an insight particularly relevant to modeling defect drivers in production environments.

In addition, Ghelani (2023) reviewed global manufacturing practices and reaffirmed that Six Sigma remains one of the most widely adopted and effective methodologies for achieving operational excellence through process standardization and variation reduction. Complementing this viewpoint, Vinod (2022) proposed moving beyond traditional reliance on p-values by incorporating effect-size measures such as standardized or scale-free coefficients, improving the practical interpretability of regression results in real-world production settings.

Overall, the literature strongly supports the integration of statistical modeling with structured quality-improvement methodologies as an effective approach for diagnosing root causes of defects. While existing literature has extensively documented the application of Six Sigma methodologies and regression analysis in manufacturing environments, our study addresses a critical research gap by specifically investigating defect drivers in the Thumb

Crotch Reinforcement (TCR) process within glove manufacturing.

3. METHODOLOGY

3.1 Data Collection

The dataset was sourced from the internal rejection logs of a single TCR machine over a nine-month period. It consisted of 566 daily observations across three glove styles. For each observation, the total production volume, total rejects, and counts for 12 distinct defect categories were recorded. The defect categories included: Burn, Uneven, Wrinkle, Misalignment, High Thickness, Knitting Defects, Penetration, Compound Splash, Bubbles, Crotch Hole, Dipping Damage, and Other (an aggregate of miscellaneous causes like washing burn and machine faults).

3.2 Data Analysis

The dependent variable was the daily rejection rate (%). The independent variables were the daily glove counts of the 12 defect types. The analysis was conducted in R and Python (Google Colab) and proceeded in three stages: (1) Exploratory Data Analysis (EDA): Summary statistics and correlation analysis were performed to understand data distribution and relationships. (2) Regression Modeling: A multiple linear regression model was fitted as the optimal analytical approach due to its compatibility with the continuous nature of the rejection rate variable (0-16%). Stepwise regression was used for feature selection to retain only statistically significant predictors ($p < 0.05$). (3) Model Diagnostics: The model was validated by checking for linearity, independence, homoscedasticity, and normality of residuals. At the violations

of homoscedasticity and normality, Log, Box-Cox, and Square Root transformations were applied to the dependent variable to improve model adequacy. Further, Diagnostic tests, including Variance Inflation Factor (VIF), Shapiro-Wilk normality test, and Durbin-Watson statistics, were applied to ensure model robustness. Also, the study followed a stepwise regression approach to identify significant predictors.

4. RESULTS AND DISCUSSION

4.1 Exploratory Data Analysis

Initial analysis revealed that Style_01 was the primary contributor to quality issues, exhibiting severe instability and a peak reject rate of 16% (Figure 1).

Descriptive statistics (Table 1) indicate a right-skewed distribution of the reject rate, indicating a generally stable process punctuated by severe quality breakdowns. All values are reported to two decimal places.

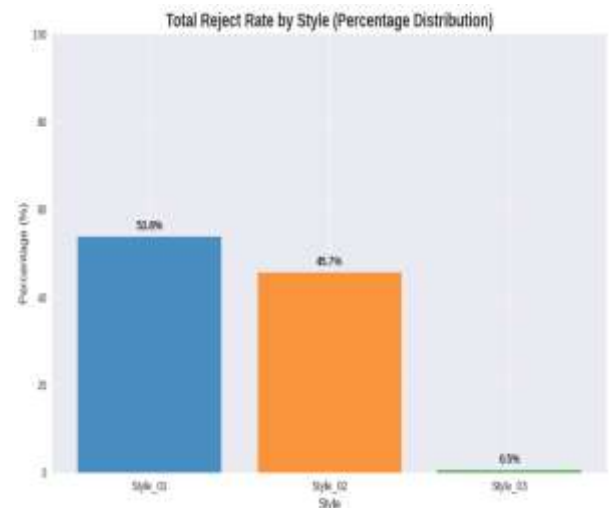


Figure 1: Style wise Reject rates

Table 1: Descriptive statistics table

Statistic	mean	std	min	max
Reject Rate	1.71	2.42	0.0	16.2
			1	0

Burn	2.56	23.36	0	493
Uneven	8	24.52	0	293
Wrinkle	18.1	56.52	0	891
	5			
Misalignme nt	3.23	16.39	0	227
High thickness	2.04	25.39	0	513
Knitting Defects	0.45	10.51	0	249
Penetration	4.41	34.86	0	458
Compound Splash	36.8	81.77	0	679
	7			
Reject Bubbles	45.0	122.6	0	1004
	7	8		
Crotch Hole	13.4	58.29	0	871
	9			
Dipping Damage	9.14	32.52	0	308
Other	21.2	65.95	0	561
	6			

The correlation heatmap and pair plots indicated that Compound Splash ($r=0.57$) and Bubbles ($r=0.57$) had the strongest positive correlation with the daily reject rate. The variables were assigned symbolic labels for analytical clarity. Accordingly, the values are denoted as $a_1, a_2, a_3, \dots, a_{12}$ where a_i refers to the i -th variable listed in Table 2.

Table 2: Correlation coefficients of different defect types with the rejection rate

Variable	Values	Label
Compound_Splash	0.571070	a_1
Bubbles	0.566236	a_2
Uneven	0.465828	a_3
Other	0.419782	a_4
Dipping Damage	0.391247	a_5
Wrinkle	0.389573	a_6
Crotch Hole	0.334163	a_7
Misalignment	0.325537	a_8
High thickness	0.227530	a_9

Penetration	0.222375	a_{10}
Burn	0.218552	a_{11}
Knitting	0.071193	a_{12}

4.2 Regression Analysis and Root Cause Identification

The initial MLR model was significant but violated assumptions of homoscedasticity and normality. After feature selection and testing various transformations, the square root-transformed model provided the best balance of explanatory power and diagnostic improvement. The final model was highly significant (F-statistic p-value $< 2.2e-16$) with an R^2 of 0.6183, meaning it explained 61.83% of the variance in the rejection rate. In the context of complex, multi-factorial manufacturing processes where unmeasured variables like raw material batch variations and human factors are always present, an R^2 value of 61.83% is considered moderately strong. This indicates that the model successfully captures the majority of the systematic variability in the rejection rate, providing a robust and actionable explanation of the primary defect drivers.

Table 3: Regression coefficient estimates for statistically significant defect types

Variable	Estimate	Std. Error	Pr(> t)
(Intercept)	0.7198	0.0229	$< 2e-16$
Burn	0.0050	0.0009	$9.58e-09$
Wrinkle	0.0021	0.0003	$7.32e-08$
Compound Splash	0.0030	0.0003	$< 2e-16$
Bubbles	0.0018	0.0002	$< 2e-16$
Crotch Hole	0.0016	0.0004	$4.75e-06$
Other	0.0034	0.0003	$< 2e-16$

The square root transformed model is then

Reject Rate

$$= 0.7198 + 0.005a_{11} + 0.0021a_6 + 0.0034a_4 + 0.003a_1 + 0.0018a_2 + 0.0016a_7$$

According to Table 3, all predictors in the final model were statistically significant ($p < 0.001$). All values are reported to four decimal places. Burn defects had the highest coefficient, identifying them as the most critical root cause, likely related to overheating or improper thermal settings. The "Other" defects category (encompassing loader faults, machine faults, etc.) and Compound Splash were the next most significant contributors, pointing to systemic issues in material handling and application consistency. VIF values for all predictors were below 1.32, confirming that multicollinearity was not a concern.

While the square root transformation improved linearity and normality, some heteroscedasticity persisted in the residual plots, a noted limitation. However, the model remains a powerful tool for root cause identification and prioritization.

5 CONCLUSION

This study successfully identified the root causes of high rejection rates in the TCR process using multiple linear regression. The analysis provides a clear, data-driven hierarchy of defect significance, with Burn, Compound Splash, and the "Other" category being the most impactful. Based on these findings, the following recommendations are proposed: (1) Immediate Action on Burn Defects: Implement enhanced temperature monitoring systems, conduct thermal profiling, and establish strict limits on material exposure to heat.

(2) Process Standardization: Address the "Other" defects by implementing quick-win procedural controls like standardized work instructions, daily equipment checks, and visual management to reduce human error and machine faults. (3) Compound Application Control: Target Compound Splash by upgrading to precision metering pumps, optimizing application parameters (viscosity, pressure), and installing sensor-based monitoring. The implemented model offers the company a framework for proactive quality control.

While multiple linear regression was appropriate for the available dataset and provided meaningful insights, alternative modeling approaches such as logistic or beta regression could not be applied due to the limited sample size. Future work will focus on expanding the dataset to enable more sophisticated regression analysis. A crucial next step is to incorporate interaction terms between defect types to analyze potential synergies.

REFERENCES

- Dinulescu R., Smeureanu I., Dobrin C. and Popa I. (2018), A statistical approach for improving the Romanian public healthcare system using the lean six sigma methodology, *Economic Computation and Economic Cybernetics Studies and Research*, No. 3, Vol. 52, pp. 5572, doi: 10.24818/18423264/52.3.18.04.
- Journal, I. J. E. R. T. (2018). IJERT-Realizing Improvements in Quality by Applying Six Sigma Methodology -A Foundry based Case Study. *International Journal of Engineering Research and Technology (IJERT)*.
- Ghelani, H. (2023). Six Sigma and Continuous Improvement Strategies: A Comparative Analysis in Global Manufacturing Industries. *Valley*

International Journal Digital Library, 954-972.

- Nurprihatin, F., Ayu, Y. N., Rembulan, G. D., Andry, J. F., & Lestari, T. E. (2023). Minimizing product defects based on labor performance using linear regression and six sigma approach. *Management and Production Engineering Review*, 88-98.
- Perez-Melo, S., & Kibria, B. G. (2020). On some test statistics for testing the regression coefficients in presence of multicollinearity: a simulation study. *Stats*, 3(1), 40-55.
- Uyanık, G. K., & Güler, N. (2013). A study on multiple linear regression analysis. *Procedia-Social and Behavioral Sciences*, 106, 234-240.
- Vinod, H. D. (2022). Kernel regression coefficients for practical significance. *Journal of Risk and Financial Management*, 15(1), 32.



A Comparative Study of Statistical and Machine Learning Techniques for Forecasting Leptospirosis Incidence in Kurunegala District

Karunasekara KMCG¹, Francisco GS¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
kgayani14@gmail.com

ABSTRACT

Leptospirosis is a zoonotic disease found worldwide, and the Kurunegala District of Sri Lanka has faced continuous outbreaks of leptospirosis, which is strongly affected by environmental and climatic factors, such as rainfall and flooding. This study was conducted to compare the predictive performance of statistical and machine learning (ML) techniques for forecasting weekly leptospirosis cases in Kurunegala District. The statistical model considered was Autoregressive Integrated Moving Average (ARIMA), and the ML techniques evaluated were Random Forest and XGBoost. Statistical models' performances were evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-Squared score (R^2), and ML models' performances were evaluated using RMSE and MAE. ARIMA (3,1,2) is the accurate statistical model, and among the ML models, XGBoost model outperformed the Random Forest model. Forecasting models provide valuable support for early warning systems and proactive public health planning. However, by the univariate nature of the analysis, its predictive performance remains constrained. By applying multivariate forecasting approaches including environmental, climatic, and socio-demographic factors, future research can address this limitation to increase the accuracy of the forecasting model.

KEYWORDS: ARIMA, Evaluation matrices, Leptospirosis, Random Forest, XGBoost

1 INTRODUCTION

Leptospirosis, a zoonotic disease, is a global public health concern, particularly in tropical and subtropical regions. The disease is caused by the *Leptospira* bacteria and is usually spread to humans through direct contact with the urine of infected animals or exposure to contaminated water and soil, especially in periods of heavy rainfall. This study aims to provide an informed understanding that can inform regional public health policies and disease prevention programs.

Previous studies have included multiple environmental and climatic variables to enhance prediction accuracy. This study focuses on the potential of univariate statistical and machine learning (ML) methods to effectively

model disease trends without relying on additional factors. Kurunegala district was chosen because a univariate forecasting model has not been developed specifically for the Kurunegala district. This comparison provides valuable insights into the respective strengths and limitations of each approach within the context of limited input data, thereby contributing to improve epidemiological modelling and supporting public health planning through evidence-based, targeted leptospirosis control policies.

1.1 Objective of the Study

The objective of this study is to develop forecasting models using both statistical and ML algorithms based on weekly historical leptospirosis incidences.

2 LITERATURE REVIEW

In past studies, few researchers have analyzed the leptospirosis outbreaks and the importance of modeling leptospirosis outbreaks. They have proposed different modeling methods, different sources for forecasting leptospirosis outbreaks.

Warnasekara et al. (2022) constructed Seasonal AutoRegressive Moving Average (SARIMA) and AutoRegressive Distributed Lag (ARDL) models to forecast leptospirosis incidence in Anuradhapura district, Sri Lanka. The more accurate univariate model was SARIMA(1,0,0)(0,1,1)₁₂, and the multivariate model was ARDL(1,3,2,1,0). With leptospirosis incidence, rainfall, rainy days, temperature, and the previous month's cases were positively associated, and relative humidity was negatively associated.

The spread of leptospirosis is significantly affected by environmental and occupational factors related to water, infrastructure, landscape, exposed animals, and occupation (Baharom et al., 2024).

A Back Propagation Neural Network (BPNN) model was developed by Fariq Rahmat et al. (2019) to predict leptospirosis cases in Seremban, Malaysia, using meteorological variables such as rainfall, temperature, and relative humidity. The results achieved 70% prediction accuracy. The most influential factor across all subregions of Seremban was rainfall.

A study about the relationship between meteorological factors and leptospirosis incidence in Hambantota District, Sri Lanka, was conducted by Ehelepola et al. (2021). Using time series and wavelet analysis, the study

concluded that changes in temperature, soil temperature, evaporation rate, and sunshine duration were affected to the leptospirosis incidence, and rainfall showed delayed effects in wavelet analysis.

Sawesi et al. (2025) systematically analyzed machine learning and deep learning approaches for leptospirosis prediction and diagnosis and concluded that models such as Support Vector Machines (SVMs), decision trees, and Convolutional Neural Networks (CNNs) demonstrated high predictive accuracy, with some achieving up to 98.02%, highlighting the promise of AI techniques in improving disease detection and forecasting.

3 METHODOLOGY

3.1 Data Description

This study was undertaken by using weekly leptospirosis incidence in the Kurunegala district, Sri Lanka. A retrospective analysis was conducted on the secondary data of weekly leptospirosis incidences that were collected from the Epidemiology Unit, Ministry of Health, Sri Lanka's official website, over the period of eighteen years from January 2007 to December 2024. The statistical analysis was done by using R 4.3.3 software, and the ML techniques were done by Jupyter Notebook using Python 3.11.

3.2 Data Analysis

3.2.1 Statistical model

The weekly leptospirosis incidence data set was loaded into R software, and the dataset was transformed into log scale. To improve the accuracy and reliability of the analysis, outliers were identified using the Inter Quartile (IQR)

method. The Multiple Imputation by Chained Equations method was applied to estimate missing values.

The Box-Jenkins method was employed to formulate a mean model. First, stationarity of the series was checked by using the Augmented Dickey Fuller (ADF) and the Kwiatkowski Phillips Schmidt Shin (KPSS) tests. If the series is not stationary, it should be done trend or seasonal differencing accordingly. The AutoCorrelation Function (ACF) and Partial AutoCorrelation Function (PACF) were used to identify the mean model. The best model for forecasting leptospirosis incidence was selected by comparing Akaike information criteria (AIC) and analyzing the model's parameters and residuals.

Following the assessment of parameter significance, it is imperative to evaluate the goodness of fit for the model. In order to check randomness, normality of residuals, and heteroscedasticity, the serial Correlation LM test, Jarque-Bera test, and ARCH LM test can be employed. The model's accuracy was evaluated using Root Mean Squared Error (RMSE), Mean Squared Error (MAE), and R-Squared (R^2).

3.2.2 Machine learning models

The research is initiated with data loading to Jupyter Notebook, and then data preprocessing, which involved the dataset was indexed by the time variable column to make it a proper time series analysis, and the dataset was randomly split into training and testing sets for modeling and to validate the model's performance using 80:20 ratio. To capture temporal dependencies in the dataset, features such as `patient_diff52`,

`is_outbreak`, and `is_high_patient` were generated.

Two ML models, Random Forest and XGBoost, are trained and fitted. XGBoost is chosen due to its efficiency, capacity to manage missing values, and robust performance, while Random Forest reduces overfitting and increases generality. RMSE, and MAE are used for model evaluation.

4 RESULTS AND DISCUSSION

Table 1 depicts descriptive statistics of the weekly leptospirosis incidence over the period 2007 January to 2024 December. Fig.1 shows time series plots of weekly leptospirosis incidence for the original data, and the Fig.2 displays the series after outlier removal and logarithmic transformation. The original series is characterized by generally low case counts punctuated by sudden, sharp spikes indicative of outbreak events. These outliers are clearly discernible in the original data. On the other hand, the

Table 1: Descriptive statistics of weekly leptospirosis incidence 2007-2024

Mean	SD	Min	Max
7.608974	17.06116	0	292



Figure 1: Time series plot of the original series

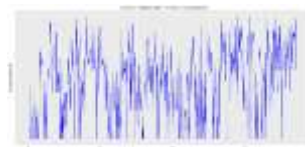


Figure 2: Time series plot of outlier removal and logarithmic transformation

Table 2: ADF test and KPSS test

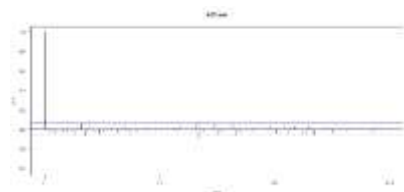
	ADF Test		KPSS Test	
	Orig inal	Differ enced	Orig inal	Differ enced
P- value	0.01	0.01	0.01	0.1

transformed series, following outlier removal and log transformation, exhibits a slight upward trend over the past three years, starting in early 2022.

Table 2 confirms that the original series is not stationary, as indicated by the ADF and KPSS tests. After the first order differencing, both ADF and KPSS tests reveal that differenced data are stationary.

4.1 Statistical Model

Fig.3 indicates that $q = 2$ and Fig.4 indicates that $p = 5$ are suggested by significant ACF and PACF, respectively. Based on the AIC values, the ARIMA(5,1,2) model is selected, and then by removing non-significant terms, the ARIMA(3,1,2) model is selected as the best model to predict the weekly leptospirosis incidence, where all the parameters are significant at their respective significance levels. The coefficient probabilities in the mean equation's parameters are all less than

**Figure 3:** ACF plot**Figure 4:** PACF plot**Table 3:** Parameter Estimation of ARIMA (3,1,2)

Parameter	Estimate	Pr(> z)
ar1	0.322036	< 0.05
ar2	0.149082	3.08e-05
ar3	0.077096	0.0259
ma1	-1.94999	< 0.05
ma2	0.949993	< 0.05

Table 4: Residuals Diagnostic Tests on ARIMA (3,1,2)

Test	p-value	Conclusion
Serial Correlation LM Test	0.3078	Residuals are random
Jarque-Bera Test	0.09304	Residuals are normal
ARCH LM Test	0.2125	Residuals homoscedastic

0.05, as shown in Table 3. As a result, every parameter in the fitted model is significant at the 5% level.

Table 4 shows the results of residuals diagnostic tests. Therefore, the ARIMA(3,1,2) model was determined to be the best and most desirable model by comparing the outcomes. The forecast performance of the ARIMA(3,1,2) model is presented in Table 5.

4.2 Machine Learning Models

The XGBoost model is fitted on the training set using the optimal parameters. The average prediction error between predicted and actual values is 5.94. The average magnitude of error between predicted and actual values is 3.45.

The Random Forest model is fitted

Table 5: Forecasting Performance of ARIMA(3,1,2) model

RMSE	MAE	R ²
0.6461736	0.5113964	0.34

on the training set using the optimal parameters, and the average prediction error between expected and actual values is 7.13. The average magnitude of error between predicted and actual values is 4.27. Figure 5 shows actual vs predicted values on the test set for both ML models. Despite the minor deviations, the predicted values generally fall within the range of the actual data, indicating that the model is well calibrated to the observed variability in leptospirosis incidences.

5 CONCLUSION

Prediction models are created using statistical and ML approaches. Comparison of the evaluation matrices of statistical and ML models are shown in Table 6.

In this study, ARIMA(3,1,2) was identified as the most appropriate model for the statistical analysis based on the significance of its AR and MA terms. The model's performance was assessed using evaluation matrices, RMSE, MAE, and the R^2 , which yielded values of 0.6461736, 0.5113964, and 0.34, respectively.

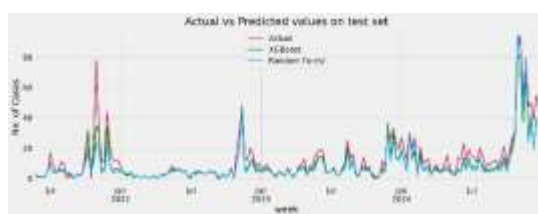


Figure 5: Actual vs predicted values on the test set for ML models

Table 6: Comparison of the evaluation matrices

	ARIMA	RF	XGBoost
RMSE	0.6461736	7.13	5.94
MAE	0.5113964	4.27	3.45

Among the ML models, the XGBoost model outperforms the Random Forest model. Therefore, for the ML models, the XGBoost model was chosen for forecasting the weekly leptospirosis incidences, and the forecast performance of the XGBoost model was evaluated, using RMSE, and MAE. They were found to be 5.94, and 3.45, respectively.

A comprehensive investigation of leptospirosis predictive approaches is essential to improve disease surveillance and reduce morbidity and mortality. Strengthening predictive capacity could ultimately support early interventions in high-risk areas and reduce the public health burden of leptospirosis.

REFERENCES

- Baharom, M., Ahmad, N., Hod, R., Ja'afar, M. H., Arsad, F. S., Tangang, F., ... & Osman, Y. (2024). Environmental and occupational factors associated with leptospirosis: a systematic review. *Heliyon*, 10(1).
- Ehelepola, N. D. B., Ariyaratne, K., & Dissanayake, D. S. (2021). The interrelationship between meteorological parameters and leptospirosis incidence in Hambantota district, Sri Lanka 2008–2017 and practical implications. *PLoS One*, 16(1), e0245366.
- Rahmat, F., Ishak, A. J., Zulkaflī, Z., Yahaya, H., & Masrani, A. (2019). Prediction model of leptospirosis occurrence for Seremban (Malaysia) using meteorological data. *International Journal of Integrated Engineering*, 11(4).
- Sawesi, S., Jadhav, A., & Rashrash, B. (2025). Machine Learning and Deep Learning Techniques for Prediction and Diagnosis of Leptospirosis: Systematic Literature Review. *JMIR Medical Informatics*, 13, e67859.
- Warnasekara, J., Agampodi, S., & NR, A. (2022). SARIMA and ARDL models for predicting leptospirosis in Anuradhapura district Sri Lanka. *PLoS one*, 17(10), e0275447.



Statistical Analysis of Production Loss Reasons in Lean Garment Manufacturing

Sandunika AMKS¹, Karunarathna SPDSSK²

¹*Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka*
sayurisanidunika@gmail.com¹

²*Department of Mathematics, Faculty of Science, University of Kelaniya*
sanjeewak@kln.ac.lk²

ABSTRACT

Identifying the key factors of production losses enables targeted improvement and greater operational efficiency in lean garment manufacturing (LGM). This study, presents a statistical investigation into the primary drivers of production losses within a LGM context, where persistent operational disruptions continue to challenge efficiency despite the adoption of Lean principles. The study employs a robust quantitative explanatory design that strategically integrates multiple linear regression for inferential analysis with a Random Forest machine learning algorithm to enhance predictive accuracy and identify feature importance. The data collection has been conducted through the systematic manual gathering of hourly data from multiple production modules, capturing key variables such as Standard Minute Value (SMV) and categorized loss reasons. Regression analysis quantified that each additional minute of Standard Minute Value (SMV) increases production loss by 0.328 Standard Allocated Hours (SAH), establishing garment complexity as a key factor. However, the machine learning model provided a more nuanced hierarchy, revealing that operational disruptions are the primary driver (48% importance), surpassing garment complexity, SMV (42%), with material input delays identified as the most severe specific disruption. The study conclusively demonstrates that advanced statistical modeling is an indispensable tool for converting operational data into strategic, actionable intelligence. Consequently, it is recommended that manufacturers prioritize eliminating supply chain bottlenecks and enhancing supply chain synchronization, while simultaneously implementing complexity-based production planning. This work substantively contributes to the practical evolution of Lean manufacturing by underscoring the critical integration of data analytics into decision-making to foster a culture of continuous, evidence-based improvement.

KEYWORDS: Garment Industry, Production Loss, Random Forest, Regression Analysis, Standard Allocated Hours

1 INTRODUCTION

The garment manufacturing industry operates under intense pressure from global competition, short lead times, and low profit margins. In this context, production losses significantly affect both cost efficiency and delivery performance. Lean manufacturing principles such as waste reduction and process flow improvement (Womack & Jones, 2003) are widely used; however, operational disruptions—including input delays, machine breakdowns, and line

imbalances—continue to hinder productivity. These losses are measured using Loss Standard Allocated Hours (Loss SAH), a metric capturing both unproduced units and garment complexity.

Existing managerial practices often rely on descriptive statistics, which identify issues but do not quantify the relative statistical impact of each loss factor. This study addresses this gap by applying advanced analytical techniques to determine the individual contribution of categorized loss reasons to Loss SAH,

and the relationship between garment complexity, measured via Standard Minute Value (SMV), and production loss.

2 LITERATURE REVIEW

The global garment industry faces strong competition and low profit margins, making efficiency essential. It is no secret that despite adopting Lean manufacturing principles to reduce waste, many production inefficiencies still remain.

Extensive research has successfully categorized the root causes of these losses, including machine breakdowns, material delays, and line balancing issues, with frameworks like Total Productive Maintenance (TPM) offering structured analysis (Nakajima, 1988; Fernando et al., 2020). Recent studies in the apparel sector further reinforce that implementing lean tools such as value-stream mapping, 5S, and line optimization significantly improves throughput, reduces work-in-progress, and lowers waste (Bizuneh & Omer, 2024; Kong et al., 2024). However, categorization alone is insufficient for strategic intervention. The financial impact of each disruption must be quantified to guide resource allocation effectively. The metric of Standard Allocated Hours (SAH) addresses this need by incorporating garment complexity measured by Standard Minute Value (SMV), thus providing a more accurate measure of lost productive capacity than simple unit counts. While studies, including those in the Sri Lankan apparel sector, identify prevalent issues through descriptive statistics, a significant methodological gap remains. The literature largely relies on frequency-

based analyses that describe *what* is happening but fail to model *how* different factors collectively and independently influence the magnitude of loss. This study addresses this gap by integrating Lean theory with advanced statistical modeling (Gunarathne & Kumarasiri, 2017; Aryarathne & Galahitiyawe, 2020).

3 RESEARCH QUESTIONS

The study focuses on six key production loss reason categories that commonly occur in garment manufacturing: TM Absent, Input Delay, Potential Issues, Production Issues, Machine Breakdowns, and Balancing Issues. Alongside these disruption categories, garment complexity measured using the Standard Minute Value (SMV) plays an important role in determining the magnitude of production loss. These variables together represent the primary contributors to Loss SAH, and analyzing them allows the study to quantify both the individual impact of each loss reason and the influence of garment complexity on production efficiency.

4 METHODOLOGY

4.1 Data Collection

Data were collected manually on an hourly basis over a three-month period from a single garment manufacturing factory consisting of multiple production modules. The final dataset comprised 517 complete daily observations, capturing SMV, Loss SAH, output quantities, and six categorized disruption reasons.

4.2 Analytical Framework

A multi-stage analytical framework was applied in this study, beginning with exploratory data analysis to understand

the structure and distribution of the dataset. This was followed by multiple linear regression to quantify the relationship between the predictors and Loss SAH. Finally, Random Forest modelling was employed to capture complex, non-linear patterns and assess the relative importance of the variables in predicting production loss.

4.3 Categorical Encoding

All six loss reason categories were converted to numerical inputs using one-hot encoding. As category frequencies were moderately imbalanced, class-weight adjustments were applied to ensure stable model training.

4.4 Model Selection

Following initial model development using multiple linear regression, assumption checks indicated multicollinearity among the categorical loss reason variables and limited explanatory contribution. Accordingly, backward stepwise selection was performed to ensure model parsimony. This resulted in a final simplified model in which SMV emerged as the sole statistically significant predictor of Loss SAH. Thus, although multiple linear regression was used initially, the final results are presented as a simple linear regression model, consistent with statistical diagnostics and model selection criteria.

4.5 Validation and Implementation

All analyses were conducted using R statistical software with comprehensive validation through assumption testing and cross-validation. This integrated approach balanced the interpretable parameter estimates from regression with enhanced predictive accuracy from

machine learning, ensuring both statistical rigor and practical relevance for manufacturing decision-making.

5 RESULTS AND DISCUSSION

5.1 Exploratory Analysis Findings

Initial analysis revealed substantial variation in production losses, with Loss SAH ranging from 0.81 to 34.88 hours and SMV values spanning 6.13 to 41.10 minutes. Distribution analysis showed Input Delay and potential Issue categories exhibited the highest average SAH losses per incident. A moderate positive correlation ($r = 0.399$) between SMV and Loss SAH confirmed garment complexity as a significant factor in production efficiency.

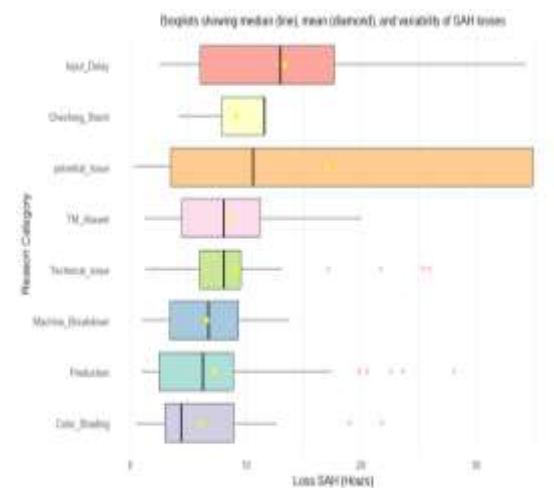


Figure 1: Loss Distribution by Reason Category

The box plot reveals Input_Delay as the most severe and variable disruption, necessitating its prioritization over more consistent, lower-impact issues like Color_Shading.

5.2 Regression Analysis Results

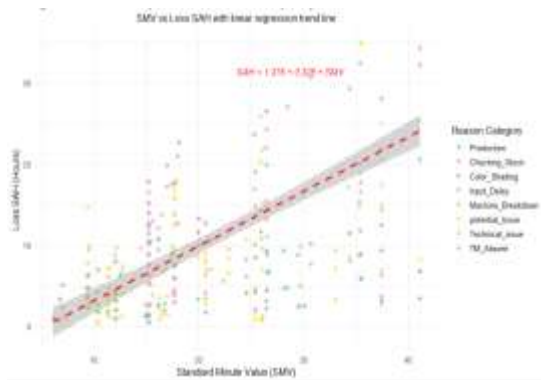


Figure 2: SMV vs Loss SAH Relationship

The scatter plot confirms a positive linear relationship between SMV and Loss SAH, yet the significant data scatter visually justifies the need for a more complex, non-linear model.

Multiple regression analysis produced a statistically significant model (F-statistic = 75.89, $p < 0.001$) explaining 21.5% of variance in Loss SAH. The final simplified model:

$\text{Loss SAH} = 1.376 + 0.328 \times \text{SMV} + \varepsilon$
 indicates each minute increase in garment complexity increases production loss by 0.328 SAH hours. Although the study initially applied a multiple linear regression model incorporating SMV and all six disruption categories, the diagnostic evaluation revealed substantial multicollinearity among the categorical predictors. Backward stepwise selection further indicated that only SMV remained statistically significant. Accordingly, the final simplified model is presented as a simple linear regression, consistent with the diagnostic outcomes.

5.3 Machine Learning Enhancement

Random Forest analysis provided complementary insights through non-parametric modeling, revealing that

operational disruption categories (48% importance) serve as the primary driver of production loss, slightly surpassing garment complexity measured by SMV (42% importance). Within these disruption categories, Input_Delay emerged as the most severe specific issue, followed by Potential_Issue, demonstrating the machine learning approach's capability to capture complex non-linear relationships that traditional regression modeling could not detect.

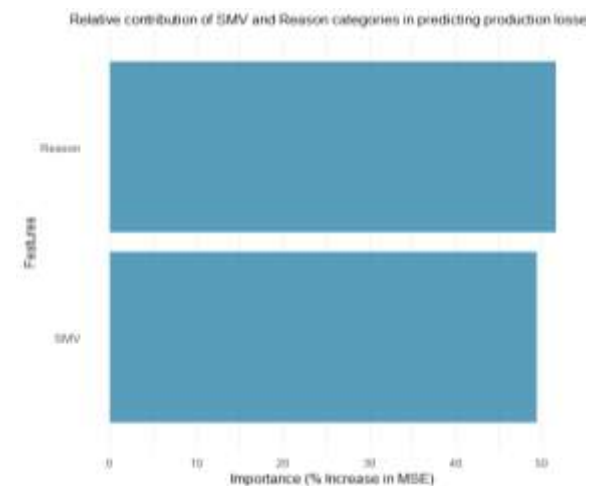


Figure 3: Feature Importance from Random Forest Model

The feature importance chart quantitatively establishes that operational disruption categories (48%) are a more critical predictor of production loss than garment complexity, SMV (42%).

Random Forest analysis demonstrated superior predictive capability, with feature importance revealing that operational disruption categories were the primary driver (48% importance), slightly surpassing garment complexity, SMV (42% importance). Within these disruptions, Input_Delay was identified as the most severe type. The exceptional performance of a specialized model using only these variables ($R^2 = 0.839$) confirmed that

SMV and reason categories collectively provide a sufficient foundation for accurate production loss forecasting.

5.4 Integrated Interpretation and Implications

Based on the complementary analyses, this study provides compelling evidence for targeted manufacturing optimization. The significant relationship between garment complexity (SMV) and production loss necessitates differentiated planning for high-complexity garments, while machine learning identifies Input Delay and Potential Issue as critical disruptions requiring prioritized intervention. The high predictive accuracy ($R^2 = 0.839$) of a simplified model confirms that monitoring SMV and reason categories enables proactive management and data-driven resource allocation. This integrated analytical approach delivers actionable insights for evidence-based decision-making, advocating for complexity-adjusted planning alongside focused disruption reduction to drive continuous improvement.

6 CONCLUSION

Based on the comprehensive analysis, this study concludes that operational disruptions, particularly material input delays, constitute the primary driver of production losses in lean garment manufacturing, closely followed by garment complexity. The demonstrated efficacy of integrating regression analysis with Random Forest modeling provides a robust framework for transforming production data into actionable intelligence. These findings underscore the necessity of implementing targeted disruption mitigation strategies

and complexity-adjusted production planning to enhance operational efficiency. This research establishes the vital role of advanced data analytics in advancing Lean manufacturing principles and fostering evidence-based continuous improvement in the apparel industry.

This study's finding that operational disruptions are the primary driver of production loss aligns with Fernando et al. (2020), who identified similar categories of inefficiencies in apparel assembly lines. However, our research extends this prior work by quantifying the precise predictive importance (48%) of these disruptions using Random Forest analysis, moving beyond descriptive categorization to provide a data-driven hierarchy for decision-making.

REFERENCES

- Akçagün, E., Dal, V., & Yilmaz, A. (2012). Using value stream mapping at apparel industry: A case study. In *International Textile, Clothing & Design Conference-Magic World of Textiles* (pp. 1-6).
- American Apparel Manufacturers Association. (1976). *Quality and Productivity: Cornerstones of Apparel Manufacturing*. Arlington, USA, 3.
- Aryarathne, S. P., & Galahitiyawe, N. W. K. (2020). Impact of Lean Manufacturing Practices on Operational Performance: A Study in Sri Lankan Apparel Sector. *Peradeniya Management Review*, 2(2).
- Delgado-Romero, L. A., Alvarado-Cespedes, A., & Meza-Ortiz, R. N. Improving Sewing Efficiency in Garment SMEs: A Lean-TPM Production Model Adapted to the Peruvian Context. *International Research Journal of Economics and Management Studies IRJEMS*, 4(5).
- Dissanayake, D. M. A. S., Perera, V. V., & Herath, D. (2024). An Empirical Study On Front-End Downtime Affecting Sustainability In Garment

Manufacturing Sector Of Sri Lanka:
Case Of Linea Aqua Avissawella
Factory. *Edith Cowan University, Sri
Lanka Ii*, 178.

- Gunarathne, G. C. I., & Kumarasiri, W. D. C. K. T. (2017). Impact of lean utilization on operational performance: A study of Sri Lankan textile and apparel industry.
- Hiba, J. C. (1998). Improving working conditions and productivity in the garment industry: An action manual
- Nallaperuma, W., Ekanayake, U., & Punchi-Manage, R. (2016). Identifying factors that affect the downtime of a production process. *Proceedings of the 15th Open University Research Sessions (OURS 2017)*.



Unit Profit Forecasting in Foreign Exchange Based on Buying and Selling Rate Margins of a Government Bank in Sri Lanka

Padmasiri MDP¹, Aponsu GMLM¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
dananjaniipyumi18@gmail.com

ABSTRACT

Foreign exchange is very critical to the profitability of the banking sector, as a buying and selling spread on rates is mainly how commercial banks remain solvent. This study is an empirical investigation into the forecasting of unit profits on foreign exchange operations within a state-owned commercial bank in Sri Lanka. The goals of this study are to model and forecast unit profit spread on the transactions resulting from buying and selling currency rates operating within the framework of computational intelligence. Machine learning and deep learning models were applied to forecast unit profit margins in foreign exchange. The RNN had 50 units, linear output, 2,751 parameters; the LSTM had 50 units, linear output, 10,851 parameters. Both were trained with Adam optimizer, batch size 64, 100 epochs, and RMSE loss. A dataset of 4,916 daily observations spanning the period from 2005 to 2025 was used for training and testing model performance. Model performance was assessed based on conventional statistical measures such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2). Random Forest model was the best performing in predicting the unit profit with an R^2 of 0.9930 among all the models, evaluated. Deep learning models including LSTM and RNN, showed promise in predicting sequences, but at a lower precision. The results show that ensemble ML models are effective for financial forecasting and can support profit management and banking decision-making.

KEYWORDS: Banking, Deep Learning, LSTM, Machine Learning, Random Forest

1 INTRODUCTION

The foreign exchange (forex) sector is one of the most active and profitable parts of modern banking. Unit Profit margin is calculated from the difference between buying and selling currency rates, commonly known as the spread. While these margins may seem small for individual transactions, their total value across large transaction volumes greatly adds to a bank's overall profits. However, the unpredictability of rate changes, influenced by global financial crises and changes in national policies, makes profit margin forecasting

challenging and uncertain. In Sri Lanka, where foreign remittances and trade play critical roles in the national economy, accurate profit forecasting is essential for maintaining liquidity and maximizing returns. The Bank manages extensive foreign exchange operations, making it an ideal case study for this analysis. Traditional statistical forecasting methods, such as moving averages or autoregressive models, often fail to capture the nonlinear and dynamic interactions within foreign exchange data.

Primarily, ML and DL have enabled new opportunities for financial

forecasting. These models may uncover complex patterns and relationships in large sets of historical data, providing more accurate predictions.

2 LITERATURE REVIEW

Globally, several studies have demonstrated financial forecasting by providing data-driven methods to model nonlinear market behaviors. Patel and Shah (2020) found that ensemble techniques like RF or Gradient Boosted Trees, combining multiple weak learners to learn the complex market spread pattern in forex markets, tend to improve accuracy.

Zhang and Wang (2020) showed that by combining different algorithms to form ensemble learning, the uncertainty of profit predictions reduces because a variety of classifiers with differences could eliminate the variance and bias. New hybrid models have merged ML and DL methods to utilize the strengths of both. Zhang, Chen, and Zhao (2024) introduced an ensemble-enhanced hybrid model that outperformed conventional methods in predicting financial trends. Kirkby and Andrean (2024) and Enkhbayar (2025) found that the feature selection as well as model tuning proved to improve significantly in terms of forecasting accuracy about bid-ask spread and algorithmic trading.

In Sri Lanka, Jayasinghe and Fernando (2019) researched the effect of exchange rate volatility on banks' profitability and emphasized the requirement for forecasting tools to mitigate foreign exchange risk. Perera and De Silva (2022), as well as De Silva and Hewage (2022), highlighted the growing application of AI-based

analytics in local banks where enhanced profitability coupled with efficiency in operations was witnessed. Together, these studies provide a solid basis for applying ML and DL models to predict profit margins in Sri Lankan banks.

3 METHODOLOGY

This study uses a quantitative research design based on predictive modeling. Data was gathered from the official annual reports of the Bank, covering daily buying and selling rates (USD/LKR) from 2005 to 2025. The methodology has three main phases: data preprocessing, model development, performance evaluation, and forecasting. To forecast unit profit margins, RNN and LSTM models were employed to capture temporal patterns, while Random Forest and XGBoost ensemble models were applied to handle nonlinear relationships and enhance predictive accuracy.

3.1 Data Preprocessing

Data preprocessing involved cleaning and transforming 4,916 daily profit margin observations from January 2005 to May 2025. Missing values were filled using a forward-fill method, and outliers were eliminated using the interquartile range (IQR) method. Feature engineering created lag variables, moving averages, and calendar-based features like month and day of the week to improve model sensitivity.

3.2 Model Development

This study uses two machine learning models (RF and XGBoost) and two deep learning models (RNN and

LSTM). The data was modeled into test and training (20:80) sets, normalized between 0 to 1 during deep learning using Min-Max scaling. RF and XGBoost learned from nonlinear patterns, while RNN and LSTM were used to capture temporal dependencies.

3.3 Evaluation and Forecasting

The performance of the model was evaluated based on RMSE, MAE, and R^2 . Loss plots from training and validation were observed for all learning models to monitor convergence and avoid overfitting. The model with the best performance was selected for the next 30- days unit profit margin prediction.

4 RESULTS AND DISCUSSION

Table 1 shows an average unit profit margin of 3.99 with positive skewness and high dispersion ($SD = 3.22$), indicating outliers from market volatility and the need for models that capture non-linear behavior.

Table 1: Descriptive Statistics

Statistic	Value
Mean	3.99
Median	3.26
Minimum	-0.90
Maximum	19.14
Standard Deviation	3.22
Skewness	1.66
Kurtosis	2.43
Jarque-Bera	3484.75
Jarque-Bera p alue	< 0.001

XGBoost Model

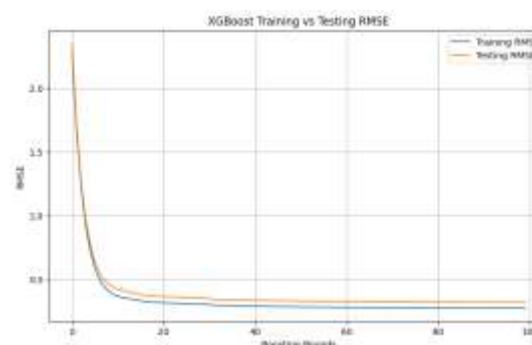


Figure 1: XGBoost Training and Validation Loss (RMSE)

Figure 1, Loss plot shows that the XGBoost model confirms the steady decline of training and validation RMSE, indicate that accurate learning and minimal over-fitting.

The XGBoost model accuracy was obtained by $RMSE = 0.1156$, $MAE = 0.0429$, and $R^2 = 0.9926$.

Overall, the model provides robust and reliable predictions for practical decision-making.

4.1 Random Forest Model

Unlike iterative models, Random Forest does not produce a training/validation loss curve. Therefore, out-of-bag (OOB) validation uses ~37% of excluded data to estimate accuracy, with OOB R^2 showing stable performance after 100-200 trees for reliable and efficient modelling.

OOB values of Table 2 show the Random Forest has stabilized, generalizes well ($OOB R^2 = 0.9929$), and adding more trees offers little performance gain but increases computation.

Table 2: OOB R^2 values

n_estimators	OOB R^2
10	0.9352
50	0.9897
100	0.9900
200	0.9898
300	0.9899
500	0.9900
700	0.9899
1000	0.9899

The highest predictive accuracy was obtained by the RF model with $RMSE = 0.1122$, $MAE = 0.0412$, and $R^2 = 0.9930$. This R^2 value confirms that the OOB internal validation generalizes well ($OOB R^2 \sim R^2$).

4.2 LSTM Model

Figure 2 shows a marked reduction in loss for both the training and validation sets, indicating that the LSTM successfully learned the temporal dependencies inherent in the unit profit margin data. While fluctuations in the validation loss are observable, they generally follow the downward trajectory of the training loss, suggesting effective learning without severe over fitting. The LSTM model had an $RMSE$ of 0.4282, $MAE=0.3267$ and an R^2 of 0.9658.

**Figure 2:** LSTM Training and Validation Loss (RMSE)

4.4 RNN Model

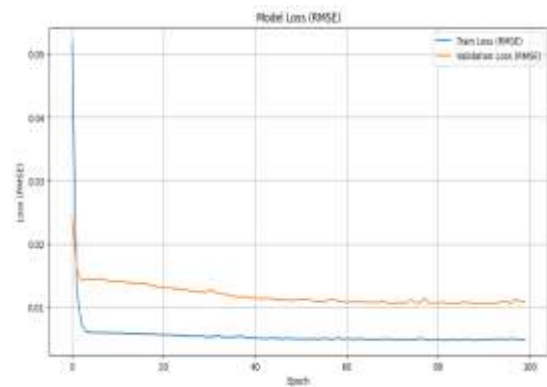
**Figure 3:** RNN Training and Validation Loss (RMSE)

Figure 3 shows the Loss plot for the RNN model confirms the steady decline of training and validation RMSE indicates that accurate learning without over-fitting.

The RNN model also achieved strong results of $R^2 = 0.9528$ with $RMSE$ 0.6653 and $MAE=0.4274$. Therefore, it has good performance in recognize trend traction, but is less precise during volatile periods. The relatively high $RMSE$ of the RNN model is expected because it is less capable of capturing long-term dependencies. Despite the slightly lower metrics, both deep learning models effectively modeled temporal dependencies and long-term fluctuations.

4.4 Model Comparison

The highest predictive accuracy was obtained by the Random Forest model. This ensemble model was able to capture nonlinear relationships in unit profit margins, demonstrating its power in financial prediction, outperforming other models.

Table3: Comparison of the Model Performance

Model	MAE	RMSE	R ²
Random Forest	0.0412	0.1122	0.9930
XGBoost	0.0429	0.1156	0.9926
LSTM	0.3267	0.4282	0.9658
RNN	0.4274	0.6653	0.9528

4.5 Forecasting

Overall, based on a holistic consideration of RMSE, MAE, and R², the Random Forest model was used for the next 30-day prediction of unit profit.

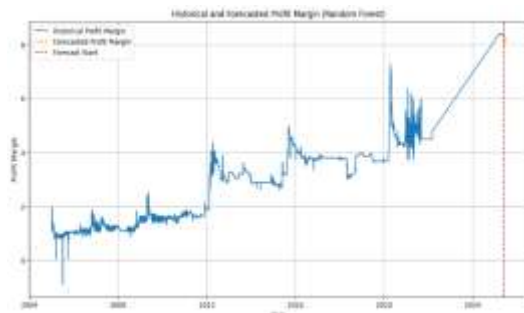


Figure 4: Historical and Forecasted Profit Margin using Random Forest Model

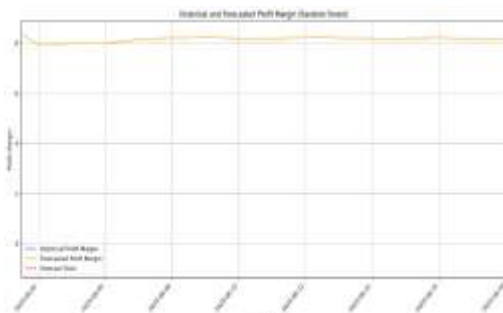


Figure 5: Forecasting profit margin values for Random Forest Model

The findings indicate that ensemble models, particularly Random Forest, outperform deep learning models in forecasting unit profit margins, aligning with Patel and Shah (2020) and Zhang and Wang (2020), who highlighted the superior ability of ensemble methods to capture nonlinear financial patterns and reduce variance. LSTM and RNN models captured temporal dependencies but were less precise, consistent with Kirkby and Andrean (2024). Compared to Sri Lankan studies (Jayasinghe & Fernando, 2019; Perera & De Silva, 2022), the results confirm that USD/LKR spreads exhibit high volatility, emphasizing the practical advantage of ensemble learning for accurate and stable operational-level forecasting in forex markets.

5 CONCLUSION

This study demonstrates the application of machine learning and deep learning models in forecasting unit profit margins in Forex operations. The Random Forest model achieved the highest predictive accuracy (R² = 0.9930, RMSE = 0.1122, MAE = 0.0412), followed closely by XGBoost, while LSTM and RNN captured sequential dependencies with lower precision. These results highlight the value of data-driven forecasting for banking operations, enabling proactive currency spread adjustments, risk management, and optimized liquidity planning.

Nevertheless, the study has notable limitations. The analysis focused on a single currency pair (USD/LKR) and a single state-owned bank, which may limit generalizability across other currencies or institutions. Additionally, the models relied exclusively on historical profit data without incorporating macroeconomic variables

or policy changes that influence exchange rate dynamics.

The deep learning architectures were relatively simple and not extensively optimized, potentially constraining their performance in capturing long-term trends or extreme market fluctuations. Despite these constraints, the findings provide strong evidence of the practical utility of ensemble machine learning models and indicate directions for future research incorporating broader datasets and advanced architectures.

financial trend forecasting. *Finance Research Letters*, 59, 104–122.

- Zitis, G. (2024). Complexity measures in deep learning models for predicting forex market volatility. *Journal of Financial Risk Analytics*, 14(2), 205–221.

REFERENCES

- Enkhbayar, B. (2025). Machine learning-based algorithmic trading system for forex pairs: Daily and intraday applications. *International Journal of Computational Finance*, 18(3), 120–138.
- Jayasinghe, P., & Fernando, R. (2019). Exchange rate volatility and bank profitability in Sri Lanka: An empirical analysis. *Sri Lankan Journal of Banking and Finance*, 3(1), 15–29.
- Kirkby, J., & Andrean, L. (2024). Predicting bid–ask spreads using supervised machine learning. *Journal of Quantitative Finance and Economics*, 22(4), 87–104.
- Patel, A., & Shah, K. (2020). Ensemble machine learning approaches for forecasting forex spreads. *Applied Artificial Intelligence in Finance*, 7(2), 55–71.
- Perera, D., & De Silva, K. (2022). AI-driven tools and decision-making in Sri Lankan banks. *Journal of Financial Technology and Innovation*, 9(3), 145–160.
- Zhang, L., & Wang, P. (2020). Ensemble learning for bank profit prediction. *Computational Economics Review*, 12(4), 201–219.
- Zhang, X., Chen, Y., & Zhao, L. (2024). Ensemble-enhanced hybrid model for



Analyzing the Factors Influencing the Water Quality Index of the Kelani River and Predicting Its Future Water Quality Trends

Padmasiri MDP¹, Aponsu GMLM¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
dananjaniapiyumi18@gmail.com

ABSTRACT

Water quality prediction is a very important key component of sustainable river basin management, public health protection, and environmental decision-making. Nevertheless, due to the multivariate and dynamic nature of water quality observations, reliable long-term prediction is hard to achieve. The current study examines extensive water quality monitoring data for the Seethawaka Ferry catchment area of the Kelani River, Sri Lanka, using both statistical and time series modeling methodologies. A structured dataset containing monthly data from 2009 to 2023 with eight basic physicochemical and microbiological parameters were preprocessed, cleaned, and worked upon to develop predictive models. The major focus is to model the Water Quality Index (WQI) with Multiple Linear Regression (MLR) and forecast future values through Auto-Regressive Integrated Moving Average (ARIMA) time series analysis. The regression model was exceptionally explanatory with an adjusted R^2 of 0.8685, and it reflected strong correlations between WQI and indicators such as Dissolved Oxygen, Biochemical Oxygen Demand, Temperature, and Fecal Coliform. For time series modeling, different ARIMA configurations were tried out, and ARIMA (1,0,1) was found to be the best-fitted model with a Mean Absolute Percentage Error (MAPE) of 8.50% and an accuracy of 91.50%. One of the significant contributions of this research is the development of a simplified forecasting model that avoids subjective WQI sub-index scores and weights and relies on raw monitoring data. This model can be applied to other Kelani River stations in future research and incorporating machine learning to automate decision support.

KEYWORDS: ARIMA, Environmental Monitoring, Kelani River, Multiple Linear Regression, Water Quality Index

1 INTRODUCTION

Water is a vital natural resource supporting human life, ecological integrity, public health, and economic growth. Particularly in the developing world, urbanization, industrialization, and population growth have led to extensive resource degradation and water contamination. Rivers are particularly susceptible to multi-point contamination since they are the main supplies of fresh water for agricultural and industrial uses. The Kelani River originates at Sri Pada and flows approximately 145 km to the North-Western coast, where it empties into the Indian Ocean at Modara in

Colombo. However, as it flows through the heavily urbanized and industrialized areas around Colombo, the river is highly polluted by additional untreated sewage, industrial waste and agricultural runoff generated in these areas. Currently, Kelani River is one of the most polluted rivers in Sri Lanka, and a study by the Central Environmental Authority (CEA) and National Water Supply and Drainage Board (NWSDB) captured the importance of ongoing monitoring and forecasting analysis. The focused area of this study is the Seethawaka Ferry area, a semi-urban section of the Kelani River influenced by agricultural and industrial activities. It reflects human and

ecological interactions and serves as a key point for monitoring water quality, identifying pollution trends, and supporting sustainable river management.

2 LITERATURE REVIEW

Globally, several studies have demonstrated the effectiveness of WQI in diverse contexts. Singh, Malik, and Sinha (2002) used the weighted arithmetic WQI to assess the Gomti River in India and concluded poor quality water due to untreated municipal and industrial effluent. Similarly, Boyacioglu (2007) and Bordalo, Teixeira, and Wiebe (2001) applied WQI in Turkish and Iberian rivers, respectively, for hotspot pollution detection and European water directive compliance.

In Sri Lankan contexts, most studies have prioritized descriptive analyses over predictive modeling. In their study on Colombo urban canals, Herath, Ratnayake, and Weerasinghe (2018) observed that these water bodies are highly polluted by sewage effluent and industrial effluents. Also, Edirisinghe, Gunawardena, and Ekanayake (2020) reported on the low to moderate water quality of the Mahaweli River basin due to agricultural runoff. Jayathilake et al. (2021) also highlighted industrial effluents and municipal waste as the key contributors to pollution of the Kelani River basin. Wijesekara et al. (2019), Gunawardena et al. (2013) and Weerasinghe et al. (2007) also found that Bolgoda Lake had a high pollution level due to microbial contamination and nutrient enrichment from seasonal variation, while Colombo Wetlands had a moderate pollution level due to Microbial–nutrient contamination

influenced by seasonal variation may also increase the system's risk of excessive algae growth.

3 METHODOLOGY

3.1 *Data Collection*

Multiple Linear Regression and time series modeling were used in this study to evaluate and predict the water quality of the Seethawaka Ferry region along the Kelani River.

The data from 2009 to 2023 on a monthly basis were collected from CEA, covering laboratory-tested parameters such as DO, BOD, Temperature, Nitrate, Turbidity, Total Phosphate, pH, and Fecal Coliform. Data preprocessing in R included validation, linear interpolation for missing values, and IQR-based outlier removal for reliability. WQI combines key water quality parameters into a single score, weighted by their importance, to reflect overall water quality. WQI was used as the dependent variable representing overall river water quality.

3.2 *Data Analysis*

A Multiple Linear Regression (MLR) model was developed to assess how eight water quality parameters affect WQI, with key predictors selected through correlation and significance testing. Model fit, significance test of the predictors, diagnostic assumptions, and multicollinearity were tested by diagnostic tests such as t-test, F-test, R^2 , diagnostic plots, and Variance Inflation Factor (VIF) values.

Time Series Analysis was carried out with the Box-Jenkins method, including stationary analysis (ADF test), difference identification, and model identification, that are conducted using

the Autocorrelation function (ACF) plot and Partial Autocorrelation function (PACF) plot. To identify the temporal relationship parameters and to forecast future trends of WQI, AR, MA and ARIMA(p,d,q) models were estimated. The randomness, parameter redundancy and the normality of the residuals were confirmed by Diagnostic tests prior to any final forecasting.

Collectively, this methodological approach combines multiple linear regression and time series analysis to provide durable data-driven insights on long-term water quality trends and drivers aiding sustainable Kelani River basin management.

4 RESULTS AND DISCUSSION

4.1 Multiple Regression Analysis

This section discusses the development and evaluation of a multiple linear regression model for WQI based on eight water quality parameters. Each parameter's individual contribution to water quality is explained by the model in an attempt to provide an indication profile for monitoring and pollution control of the Seethawaka Ferry location of the Kelani River.

The model was constructed using the R programming and the precise model coefficients, estimates, p-values, and standard error are presented in Table 1.

Table1: Summary of the Multiple Linear Regression Model

Predictor	Estimate	Std. Error	p-value
Intercept	63.763	4.233	< 2e-16
Dissolved			
Oxygen(DO)	4.260	0.194	< 2e-16
Biological	-4.087	0.225	< 2e-16

Oxygen Demand (BOD)			
Temperature	-0.178	0.060	0.0034
Nitrate	-1.452	0.308	< 2e-16
Turbidity	-1.044	0.102	< 2e-16
Total Phosphate	-1.822	0.753	0.0167
pH	1.358	0.479	0.0052
Fecal Coliform	-0.041	0.004	< 2e-16

According to the regression analysis, it was observed that DO is positively and significantly related to WQI meaning, high DO contributes good water quality. In contrast, BOD, Nitrate, Temperature, and Fecal coliform are found to have significant negative relationships, which signify their negative contribution towards water quality. pH has a positive effect, meaning that water conditions are improved as the acidity becomes more in balanced. Overall model is significant ($p = 2.2 \times 10^{-16}$) since p-value is far below the common significance level indicate that the set of eight water quality parameters collectively explains a significant amount of the variability in WQI. All the predictors are found to be statistically significant ($p < 0.05$), and among them, DO, BOD, and Fecal Coliform are considered to be the most influential predictors.

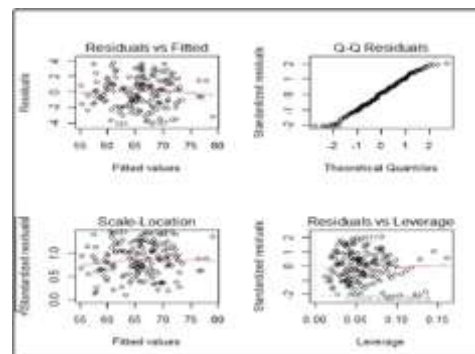


Figure 1: Diagnostic Plots

The overall model diagnostics confirm that the assumptions of regression are well satisfied. Each of the coefficients in the model represents the expected change in WQI when there is a unit increase in the respective variable, while others are kept constant. The model fit statistically well with an adjusted R^2 of 0.8685, which implies that approximately 86.85% of WQI variability is explained by the selected predictors.

The VIF values DO (1.015), BOD (1.043), Temperature (1.046), Nitrate (1.084), Turbidity (1.026), Total Phosphate (1.015), pH (1.034), and Fecal Coliform (1.040) are all below 5, indicating no multicollinearity and that each predictor contributes independently to WQI variations.

The MLR model equation is:

$$\text{WQI} = 63.76 + 4.26(\text{DO}) - 4.09(\text{BOD}) - 0.18(\text{Temperature}) - 1.45(\text{Nitrate}) - 1.04(\text{Turbidity}) - 1.82(\text{Total Phosphate}) + 1.36(\text{pH}) - 0.041(\text{Fecal Coliform}) \quad (1)$$

The model is significant and reliable, as DO and pH increase water quality while BOD, Nitrate, Turbidity, and Fecal Coliform decrease it, and can be used for the estimation of WQI and river management.

4.2 Time Series Analysis

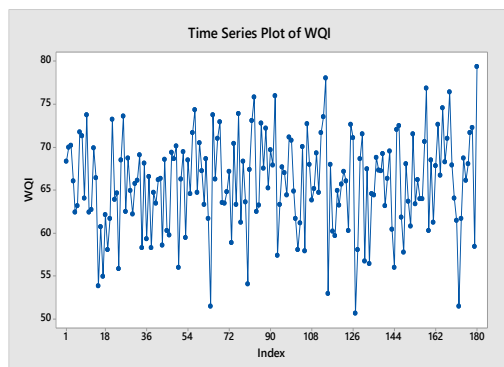


Figure 2: Time Series Plot of WQI

The seasonal and annual fluctuations seen in Figure 1 occurred due to the changes in river discharges, precipitation and human interference, including industrial effluent and urban runoff.

Before model development, the stationarity of the WQI series was examined using the Augmented Dickey-Fuller (ADF) test. The test yielded a Dickey-Fuller statistic of -5.7452, lag order 5, and p-value 0.01, indicating the series is stationary without a differencing. ACF and PACF were used to detect potential AR and MA terms. Various ARIMA(p,0,q) models were then evaluated through a grid search, systematically testing combinations of p and q to identify the most effective forecasting structure.

The ARIMA(1,0,1) model is selected as best model, all parameter p-values are below 0.05, indicating significance at the 5% level.

Table 3: Final Estimate of Parameters for ARIMA (1,0,1)

Type	Coef	SE Coef	T	P
AR 1	1.000	0.000	6296.71	0.00
	0	2		0
MA 1	0.981	0.000	2664.97	0.00
	5	4		0

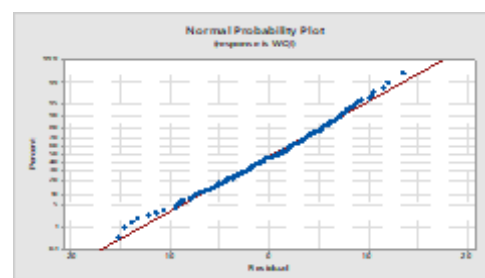


Figure 3: Normality of Residuals Plot -ARIMA (1,0,1)

Figure 3, the Residuals plot shows an approximate straight line, indicating that the residuals are approximately normally distributed. The Modified Box-Pierce (Ljung-Box) Chi-Square test for ARIMA(1,0,1) shows P-values of 0.998, 0.910, 0.952, and 0.916, all above 0.05, indicating that the residuals are random. In the correlation matrix, 1 (AR(1)) and 2 (MA(1)) have a correlation of 0.625, which is below 0.9, indicating no high correlation or parameter redundancy.

The performance of the models ARIMA(1,0,1), ARIMA(2,0,0) and ARIMA(3,0,0) is selected by goodness of fit criteria namely AIC, BIC, MAPE, and Accuracy.

The minimum AIC (1136.075) was observed in ARIMA(1,0,1), and the highest accuracy (91.50%) with MAPE at 8.50%. The parameters were statistically significant, and the residuals were randomly scattered without a serial correlation pattern, providing confidence in the model's capability to describe the WQI dynamic.

Table 4: Comparison of the Model Performance

Model	MAPE	Accuracy	AIC	BIC
ARIMA(1,0,1)	8.49771	91.5023	1136.075	1148.847
ARIMA(2,0,0)	15.0497	84.9503	1136.15	1148.922
ARIMA(3,0,0)	13.2209	86.7791	1138.07	1154.038

The ARIMA(1,0,1) model provided precise short-term WQI forecasts. The forecasted mean WQI was

approximately 66, indicating moderate water quality in Seethawaka Ferry area.

Forecasts were consistent, indicating consistent improvements and homogeneous trends during the recent years.

5 CONCLUSION

This study demonstrates that water quality at Seethawaka Ferry, Kelani River, can be effectively monitored and forecasted using MLR and time series models. The forecasted WQI trends indicate periods of potential water stress, which may pose risks to human health, aquatic ecosystems, and local communities relying on the river. The technique streamlines WQI monitoring by using raw data and allows for periodic updates.

The findings further highlight how pollution behavior in the Kelani River is shaped by surrounding land-use patterns, industrial activity, and rainfall dynamics. Urban and industrial zones increase contaminant loads through effluent discharge and runoff, while rainfall events wash pollutants into the river and influence temporal variations. These environmental linkages help explain fluctuations in water quality and emphasize the need for targeted management across different catchment areas.

Overall, the modeling approach can be expanded across the river to provide early warnings, manage pollution, and promote sustainable water management, supporting long-term freshwater conservation. These insights can inform proactive water management strategies and offer a robust, data-driven framework that strengthens Sri Lanka's

national efforts in freshwater governance and river policy planning.

REFERENCES

- Amarathunga, A. A. D., Perera, L. A. S. R., & Gunawardena, U. A. D. P. (2021). Assessment of microbial and nutrient pollution in Colombo wetland systems using water quality indicators. *Sri Lanka Journal of Aquatic Sciences*, 26(1), 45–58.
- Bordalo, A. A., Teixeira, R., & Wiebe, W. J. (2001). A water quality index applied to an international shared river basin: The case of the Douro River. *Environment International*, 26(7–8), 581–591.
- Boyacioglu, H. (2007). Development of a water quality index based on a European classification scheme. *Water SA*, 33(1), 101–106.
- Edirisinghe, E. A. P. K., Gunawardena, U. A. D. P., & Ekanayake, E. M. A. P. (2020). Assessment of Mahaweli River water quality using Water Quality Index (WQI). *Journal of Environmental Professionals Sri Lanka*, 9(1), 1–11.
- Herath, I., Ratnayake, R. R., & Weerasinghe, M. (2018). Monitoring and assessment of urban canal water quality in Colombo, Sri Lanka. *Environmental Monitoring and Assessment*, 190(5), 261.
- Jayathilake, W. M. S. K., Chandrasekara, W. U., & Wijesundara, W. M. S. M. (2021). Water quality status of the Kelani River and its tributaries in Sri Lanka: A case study from an urbanized region. *Sri Lanka Journal of Aquatic Sciences*, 26(2), 95–110.
- Singh, D., Malik, D. S., & Sinha, S. (2002). Water quality assessment of the Gomti River in Uttar Pradesh, India using the water quality index. *Nature Environment and Pollution Technology*, 1(4), 447–450.
- Wijesekara, H. H. D. M., Gunawardena, U. A. D. P., & Weerasinghe, M. (2019). Water quality status and pollution levels in the Bolgoda Lake system. *Sri Lanka Journal of Aquatic Sciences*, 24(1), 27–37.



Predicting the Job Completion Time in a Continuous – Shift Polythene Manufacturing Plant

Iwanthi MAY¹, Vithanage GVRK¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
iwanthiyashodha98@gmail.com

ABSTRACT

In a continuous-shift polythene manufacturing plant, efficient scheduling and correct prediction job completion time are the major challenges. Delays in any of the stages of printing, lamination, slitting, and pouching disrupt the flow of production, resulting in late delivery and eventually reduced customer satisfaction. Therefore, this research has developed a multiple linear regression-based predictive model to estimate the final job completion time by utilizing the historical production data at Clean Poly Packaging (Pvt) Ltd. Regression models fitted separately for printing (Flexo and Gravure of print materials), lamination, slitting, and pouching were integrated into a complete computational program developed in R that estimates the final output date after a customer order enters production. In contrast to prior studies, which often use queuing-theoretical or machine-learning models, this work identifies the practical factory-specific application of statistical modeling. First, the models are validated based on train-test evaluation involving RMSE, MAE, and residual diagnostics. The results reflect that material type, number of colors, machine type, and processing category have a crucial impact on production time. By employing the proposed model, the prediction accuracy has significantly improved, whilst average scheduling errors were reduced and more reliable production planning could be achieved. This research makes several contributions, as it provides a validated, data-driven decision-support tool for operational scheduling in polythene manufacturing and lays a foundation for future integration with real-time production data and machine learning mechanisms.

KEYWORDS: Job Completion Time, Polythene Manufacturing, Production Planning Regression Analysis, R Programming

1. INTRODUCTION

Polythene production is related to multiple independent processes like extrusion, printing, lamination, slitting, and pouching. Each job has to follow these steps in a sequence, which means any interruption in one process can have an impact on the entire production process. It is therefore essential to properly estimate the job completion times for effective resource allocation, improved scheduling, and timely delivery of finished products.

In 24-hour operation plants such as Clean Poly Packaging (Pvt) Ltd., production does not halt for long periods.

Nevertheless, product planners tend to use manual estimates or traditional judgments based on experience to estimate completion times. Such techniques are questionable due to their failure to consider differences in material types, machine settings, operator performance, and color processing times.

As a response to problems such as these, the study formulates a statistical model from using data to enhance the estimation of the completion times of the jobs. Using historical job card data and regression analysis, the model computes the impact of factors like material type, machine type, number of colors, and total

measurement on production time. The outcome is a computerized application that production planners to make the right decisions, reduce scheduling errors, and maximize the efficiency of operations in a continuous shift polythene plant setup.

2. LITERATURE REVIEW

Predictive modeling in manufacturing has attracted attention in operational research and industrial engineering:

Groover (2015) highlights the need for time prediction models in batch processing systems. Montgomery & Runger (2014) emphasize regression models for manufacturing time estimation. Singh et al. (2017) used multiple regression to forecast production time in textile manufacturing. Kumar & Rajan (2019) applied DOE and ANOVA to evaluate cycle time in automotive component manufacturing.

Li et al. (2020) found that print quality complexity significantly affects run times in packaging plants.

Despite global research on production prediction, very few studies address polythene manufacturing, particularly continuous-shift plants where: Machines run 24 hours, work orders pass through multiple dependent processes, material and color combinations cause large variations.

This research is the first to produce a plant-specific, regression-based job completion model for a Sri Lankan polythene manufacturer.

3. METHODOLOGY

This research adopts a quantitative and computational approach to develop a predictive models for the printing process

in a continuous-shift polythene manufacturing environment. The study was conducted at Clean Poly Packaging (Pvt) Ltd, using real production data obtained from the company's Production Department.

3.1. Data Collection

Data was collected from job cards, production logs, and daily reports. Each job included:

- Material type (e.g., PET, Nylon, BOPP)
- Number of print colors
- Machine type (Flexo or Gravure)
- Total meterage of the job
- Process starts and end times

Additional data on lamination, slitting, and pouching were extracted from historical records and verified with the production planning assistant to ensure accuracy. The dataset was cleaned, standardized, and organized for analysis using the R programming language.

3.2. Model Development

Multiple linear regression models were developed to predict the Total Time per 1000 meters for Flexo and Gravure machines separately. The dependent variable was Total Time per 1000m ($\text{Average Time} + \text{Machine Setup Time} + \text{Color Matching Time}$). Independent variables included number of colors, material type, and machine type. The predicted printing time was then combined with the estimated times from the Laminating, Slitting, and Pouching which are created equations to obtain the final total production time.

3.3. Implementation

An automated program was developed to perform calculations. The

program takes user input such as job start date, material, machine type, color count, and total meterage, and outputs the estimated job completion date. It computes total hours, converts them to days using a 24-hour continuous-shift model, and adds curing or setting up times where applicable.

4. RESULTS AND DISCUSSION

The data analysis process involved statistical and mathematical techniques to uncover relationships between job parameters (meterage, material, colors) and total job completion time. The study considered both Flexo and Gravure printing machines, focusing on five printing materials. The Flexo machine uses three printing materials (PET, NYLON, BOPP) and The Gravure machine five print materials (PET, NYLON, BOPP, MPET, LD).

4.1. Regression equation for PET material for Flexo and Gravure machines

For the PET material, significant multiple linear relationships were found between the number of colors (x), machine type and total time per 1000m components:

PET Material

According to figures, some of figures there are positive correlations between total time per 1000m and colors, machine type and meterage.

The summary is obtained the regression model including independent variables. The intercept and Meterage p-values of the model are greater than 0.05 therefore the model is not significant. Then removed intercept and used stepwise regression.

Model 1

Step: AIC= 1464.98 and final model is ,
 $Total\_Time\_Per\_1000m \sim Machine + Color - 1$ After following model is obtained.

$$TotalTime_{per1000m} = 35.170 * Machine + 10.051 * Color - 1 \quad (1)$$

Then the p-value of the model is $< 2.2e-16$, since the p-value of the overall model is less than 0.05. Therefore, the overall model is significant.

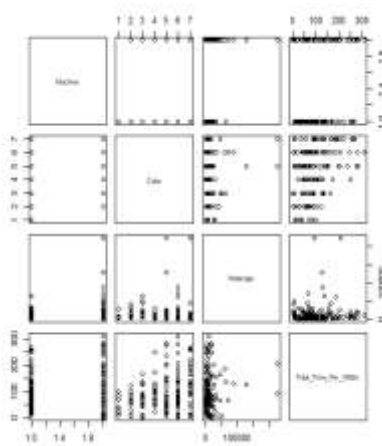


Figure 1: Scatter Plot

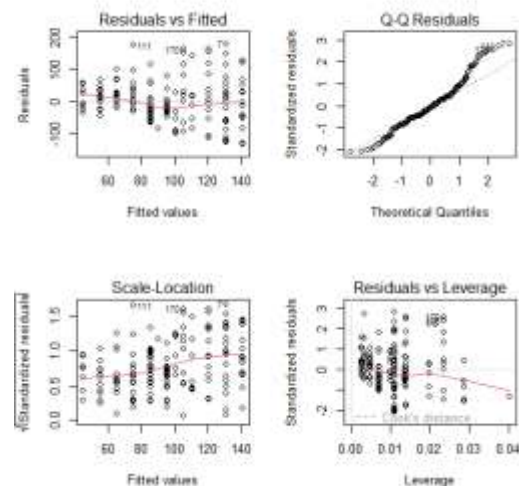


Figure 2: Checking Model Assumptions

The residual vs fitted plot shows mostly random scatter with slight curvature, indicating the relationship is largely linear. The Q-Q plot suggests residuals are approximately normally distributed. The scale-location plot shows increasing variance, indicating heteroscedasticity, while no influential

outliers are detected. VIF values of 6.81 for machine and color indicate multicollinearity, but the overall model remains statistically significant.

4.2. Regression equation for NYLON material for Flexo and Gravure machine

NYLON Material

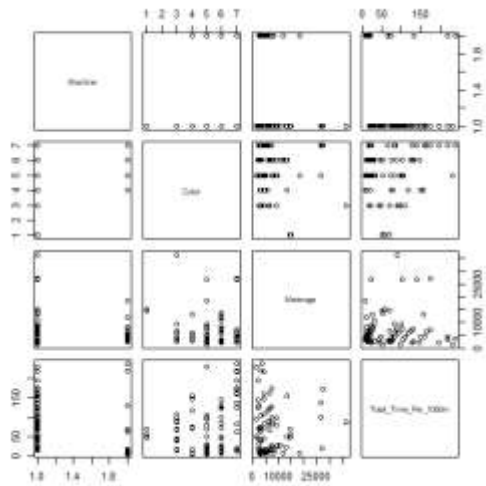


Figure 3: Scatter Plot

The meterage p-value of the model is greater than 0.05. After stepwise regression.

Model 2

Step: AIC= 685.7 and our final model is ,
 $Total\_Time\_Per\_1000m \sim Machine + Color$

After following model is obtained,

$$Total_{TimePer1000m} = 78.52 - 54.54 * Machine + 11.76 * Color \quad (2)$$

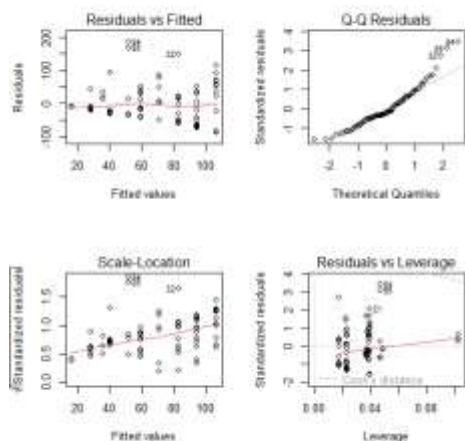


Figure 4: Checking Model Assumptions

The model p-value is $5.472e-05$, which is less than 0.05, so the overall model is significant. The residual vs fitted plot shows mostly scattered points with a few possible outliers, indicating an approximately linear relationship. The Q-Q plot suggests the residuals are normally distributed. The scale-location plot shows increasing variance, and the fourth plot indicates some influential observations. VIF values for machine and color are 1.03, showing no multicollinearity, and the model remains statistically significant.

4.3. Regression equation for MPET material for Gravure machine

MPET Material

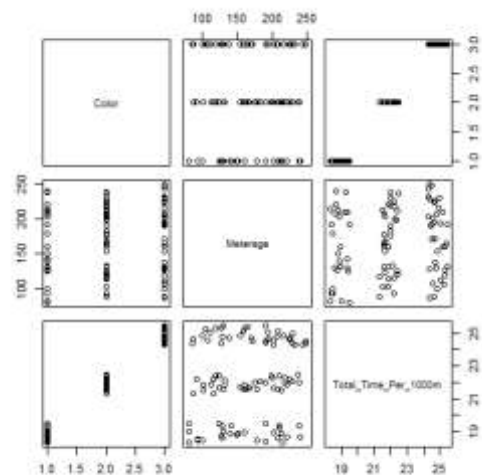


Figure 5: Scatter Plot

The meterage p-value of the model is greater than 0.05. After used stepwise regression.

Model 4

Step: AIC= -204.31 and our final model is,

$$Total\_Time\_Per\_1000m \sim Color$$

After following model is obtained,

$$Total_{TimePer1000m} = 15.952 + 2.954 * Color \quad (3)$$

The model p-value is $< 2.2e-16$ with an R-squared of 0.9777, indicating a highly significant model. Diagnostic

CONSTRUCTING A PROGRAM FOR ESTIMATING JOB COMPLETION TIME IN A CONTINUOUS – SHIFT POLYTHENE MANUFACTURING PLANT

plots show a roughly linear relationship, approximately normal residuals, mostly constant variance with minor heteroscedasticity, and no serious influential outliers.

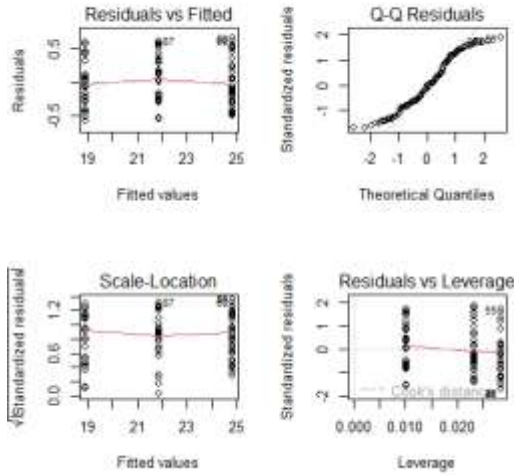


Figure 6: Checking Assumption

Step-by-Step Prediction Process

1. Printing Time Calculation (Flexo / Gravure):

Based on the chosen machine and material type, the model applies the corresponding regression equations for:

- Total Time per 1000 m
- Machine Type
- Color

Total printing time is then computed as:

$$T_{\text{Printing(min)}} = \frac{\text{Meterage}}{1000} \times \left(\frac{\text{Total Time}}{\text{Per 1000m}} \right) \quad (4)$$

2. Lamination Time (if applicable):

For jobs requiring lamination, the lamination time per 1000 m is applied, and a curing time (usually 24 hours) is added to allow the lamination adhesive to set properly. In here categorized lamination part. According to Table 1, the equation for lamination part.

$$T_{\text{Lamination(min)}} = (l_0 \times \quad (5)$$

$$\frac{\text{Meterage}}{1000} + 1440)$$

- $l_0 = \text{Average Time Per 1000m for Lami category}$
- 1440 minutes = 24-hour cure time

Table 1 – Lami Category

Print Material	Lami Material 1	Lami Material 2	Average Time Per 1000m
PET	LD	-	20.202
PET	CPP	-	12.553
PET	ALUFOIL	-	23.104
PET	MPET	-	15.736
NYLON	CLD	-	21.171
PET	PET	CPP	41.894
PET	PET	ALUFOIL	46.413

3. Slitting Time:

The model includes slitting time based on regression equations or standard operational averages.

Also categorized the slitting part.

Table 2 - Slitter Category

Order Type	Average Time Per 1000m
Reel	29.703
Pouch	17.718

Created the equation for slitter part.

$$T_{\text{Slitting (min)}} = s_0 \times \frac{\text{Meterage}}{1000} \quad (6)$$

- $s_0 = \text{Average time per 1000m for slitter category}$

4. Pouching Time (if applicable):

For pouch jobs, the model calculates pouching time based on the number of bags produced per 1000 bags, again using regression estimates or historical averages.

Table 3 - Pouch Category

Pouch Type	Average Time Per 1000 Bags
CS	15.649
3Z	16.818
GAZZERT	36.722
STANDUP	27.169
POUCH	
4 SIDE	85.062
SEAL	
OFF	53.828
CENTER	
SEAL	
STANDUP	41.587
WITH	
SHIPPER	

According to Table 3, constructed the equation for Pouch part.

Also categorized the pouch types.

$$T_{Pouching} \text{ (min?)} = p_0 \times \frac{\text{No of Bags}}{1000} \quad (7)$$

- $p_0 =$
Average time per 1000 bags for pouch type

5. Total Production Hours:

The total time is obtained by summing the times for all stages:

$$T_{Total} \text{ (hours)} = \frac{T_{Print} + T_{Lamination} + T_{Slitting} + T_{Pouching}}{60} \quad \dots (13)$$

6. Conversion to Calendar Days:

Since the plant operates on a continuous 24-hour shift system, total hours are converted into days by:

$$\text{Required Days} = \frac{T_{Total} \text{ (hours)}}{24} \quad (9)$$

7. Final Completion Date:

The required number of days is added to the job entry date to produce the estimated completion date:

$$\text{Completion Date} = \text{Job Entry Date} + \text{Required Days} \quad (10)$$

5. CONCLUSION

This study developed a regression-based model to estimate job completion time in a continuous-shift polythene manufacturing environment. Using real production data, the model accurately predicts job output dates by considering material type, machine type, and color count. Its predictions are in extremely proximity to actual production rates. It is an excellent decision-support system for schedulers which helps optimize scheduling, reduce downtime, and increase customer satisfaction. Future enhancement may involve incorporating real-time inputs and machine learning techniques to make predictive accuracy and responsiveness even more accurate and responsive.

REFERENCES

- Groover, M. (n.d.). *GlobAl editIOn Automation, Production Systems, and Computer-Integrated Manufacturing FoUrTheeditIOn*.
- Taifa, I., & Vhora, T. (2019). Cycle time reduction for productivity improvement in the manufacturing industry. *Journal of Industrial Engineering and Management Studies*, 6(2), 147–164.
- Yue, W., Tang, R., Guo, A., Tao, X., Hu, Y., Sheng, X., Qü, P., Wang, S., Li, J., & Li, F. (2023). Enhancing Starch-Based Packaging Materials: Optimization of Plasticizers and Process Parameters. *Materials*, 16(17), 5953–5953.
- Alexander, M. T., Montgomery, D. C., & Runger, G. (1995). Applied Statistics and Probability for Engineers. *Technometrics*, 37(4), 455.
- Pirola, F., Zambetti, M., & Cimini, C. (2021). Applying simulation for sustainable production scheduling: a case study in the textile industry. *IFAC-PapersOnLine*, 54(1), 373–378.



Data-Driven Cost of Spare Parts in Selected Production Machines in Korean SPA Packaging (Pvt) Ltd

Jayakody KAND¹, Vithanage GVRK¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
divyanirupadhi@gmail.com

ABSTRACT

Effective control of spare part costs is one of the critical determinants of equipment reliability, production continuity, and overall manufacturing efficiency. Spare parts, which contribute a significant proportion to maintenance expenditure, have usually been poorly tracked. This research investigates the cost structure and spending behavior of spare parts used on five major production machines at Korean SPA Packaging (Pvt) Ltd. Using descriptive and exploratory data analysis with Microsoft Excel, R, and Power BI, this study analyzed a dataset comprising 8,255 transactions extracted from the ERP system for the period 2023-2025. It was found that the total expenditure came to LKR 71.4 million and more than 63% of the amount was accounted for by two machines: Factory 06 Corrugator Plant No. 02 and Flexo Printer No. 06. A few items, such as the Ink Roller, Servo Driver, and Rubber Sleeve, accounted for a disproportionately large share in the overall cost. Cost distribution was skewed to the right, which indicates the impact of high-value items on the overall spare parts expenses. This will help in devising data-driven strategies like bulk procurement, diversification of supplies, and predictive maintenance to achieve cost transparency and improvement in evidence-based managerial decisions.

KEYWORDS: Cost Analysis, ERP Data, Power BI, Predictive Maintenance, Spare Parts Management

1 INTRODUCTION

The spare part is a significant factor to maintenance costs in manufacturing companies. When you don't understand your actual costs, or are reactive in your approach to maintenance, inefficiencies result augmenting unplanned downtime and procurement spending. Korean SPA Packaging (Pvt) Ltd., is one of the leading manufacturers of corrugated print.

Initial glimpsing of the company's ERP data revealed that specific machines and spares represented a majority of expenses. For instance, the machines Corrugator and Flexo in Factory 06 accounted together for over 60% of all spare parts spending. We present our work here addressing what cost patterns exist, where the opportunities for high-

value components are and how costs can be reduced through analysis-driven actions.

2 LITERATURE REVIEW

Studies suggest that cost due to spare-parts comprises, on average, 40-70% of total maintenance costs in industrial manufacturing (Gupta et al., 2019). Besides, poor maintenance management of spare parts leads to increased costs, downtime, and reduced machine efficiency (Huiskonen, 2001). Contemporary research focuses on the application of big data and Enterprise Resource Planning (ERP) system in facilitating cost transparency and rational decision-making (Perera & Silva, 2020).

Inventory management theories Inventory control models like the

Economic Order Quantity (EOQ) and ABC classification guide firms to determine optimal order frequency and concentrate on high value items (Silver et al. 1998). Likewise, the Life Cycle Cost (LCC) analysis combines investment cost and lifetime maintenance costs (Huang et al., 2017). The 80/20 rule, Pareto Principle is also applied to this issue because a limited number of items usually account for the large majority of costs (Kennedy et al. Predictive maintenance, supply chain risk management and ERP- integrated dashboards are identified as successful measures to optimize the cost effectiveness of spare parts.

3 METHODOLOGY

The study adopted a quantitative exploratory research design, using secondary data from the company's ERP system. Data covering the period from June 2023 to May 2025 were extracted for five major machines.

Table 1: Selected factories and machines analyzed

Factory	Machine
Factory 01	Corrugator Plant No:01
Factory 01	Flexo Printer No:05
Factory 06	Corrugator Plant No:02
Factory 06	Flexo Printer No:06
Digital Section	Digital Printer Line Corrugation No:01

A total of 8,255 purchase transactions were analyzed. Variables included machine name, part code, purchase date, quantity, unit price, and total cost. Data preprocessing was done in Microsoft Excel to remove duplicates

and standardize fields. Descriptive analysis and visualizations were performed using R (for statistical summaries and plots) and Power BI (for interactive dashboards).

4 RESULTS AND DISCUSSION

4.1 Overall Spare Parts Expenditures

The two-year ERP dataset had 8,255 procurement records from 202 to 2025. The descriptive statistics show a very skewed cost pattern — there were very few highly expensive transactions, but most had low-value items. The mean unit cost was LKR 4,625 and median was LKR 160, indicating that significant spends by a small number of items were driving overall cost.

The total expense amounted to LKR 71.4 million of which over 63% was attributed to Factory 06 machines.

Table 2: Machine-wise spare part cost distribution

Machine	Total Cost (LKR)	Share of Total
Factory 06 – Corrugator Plant No:02	23,042,260	32.3%
Factory 06 – Flexo Printer No:06	22,299,745	31.2%
Digital Printer Line Corrugation No:01	15,671,066	21.9%
Factory 01 – Flexo Printer No:05	9,232,752	12.9%
Factory 01 – Corrugator Plant No:01	1,166,437	1.6%

These findings indicate the Factory 06 equipment accounts for the majority of the spare part budget due to higher

usage, aging machinery, and costlier components.

4.2 Monthly Cost Trend Analysis

Trends by month revealed rampant volatility in spending. There were several sharp spikes in February 2024, October 2023, and March 2025 primarily due to bulk purchases of rollers and electronics.

This spasmodic disposition indicates reliance on reactive maintenance rather than scheduled preventive procurement.

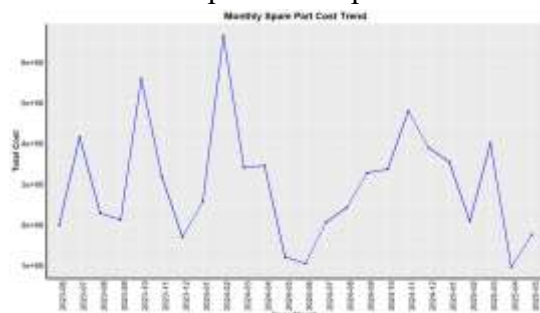


Figure 1 : Monthly spare part cost trend

Regular fluctuations highlight the need for budget forecasting and predictive maintenance scheduling to smooth out expenditure patterns and prevent financial shocks.

4.3 Department-Wise Cost Distribution

Analysis by department confirmed that the Engineering Department was responsible for nearly 91% of total expenditure, followed by production units with minimal consumption.

Table 3: Department-wise spare part expenditure

Department	Total Cost (LKR)
Engineering	65,281,007
Corrugated Factory VI	4,174,300
Digital Section	962,281
Corrugated Factory I	548,015
Human Resources & Misc.	349,653

4.4 Machine Wise Cost Distribution

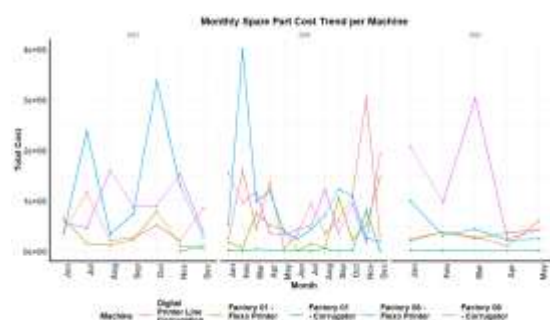


Figure 2: Monthly spare part cost trend per machine

Figure 2 presents the monthly spare part cost variation from 2023 to 2025 for some machines. The highest and least regular cost increases, reflecting major repair events, were found for the Flexo Printer (Factory 06), while corrugator machines demonstrated much more stable and relatively lower monthly costs. Overall, the cost patterns do not show any clear seasonal trend, which means that the main driver of maintenance needs is unexpected machine failures and operational load.

4.5 Pareto Analysis of Spare Parts

In order to determine the "vital few" items making up most of the costs, a Pareto analysis was conducted. Findings indicate that ten spare parts only make up close to 25% of the overall costs, led by consumables and high-tech electrical components.

Table 4: Top 10 costly spare parts by total expenditure

Spare part description	Total Cost (LKR)
Ink Roller (Y-BC-2061-28)	4,926,994
Servo Driver – 30KW	2,584,010
Rubber Sleeve (SPS-08-24)	2,252,793
Grease (Klubersynth HB 74-401)	2,102,511

Lead Edge Feeding Roller	1,810,837
Polyurethane Pad – Die Cutter Unit	1,682,045
Thinner	1,624,327
Anvil Pad – Die Cutter Unit	1,510,073
Slotting Blade	1,474,722
High-Speed Cotton Belt	1,325,111



Figure 3: Top 10 costly spare parts

The Ink Roller alone accounted for almost 7% of total costs. This dominance justifies negotiating long-term supplier contracts and evaluating the feasibility of local manufacturing for repetitive consumables.

4.6 Power BI Dashboard Integration

A Power BI dashboard was developed as a managerial decision-support tool. It enables dynamic filtering by year, machine, and department, and visualizes real-time procurement metrics.



Figure 4: Power BI Dashboard

This digital solution bridges data analytics and operational decision-making, allowing managers to quickly

identify cost anomalies and prioritize cost-saving actions.

5 CONCLUSION

This study demonstrates that spare part costs in Korean SPA Packaging (Pvt) Ltd., are highly concentrated among a few machines and components. Data-driven analysis revealed that procurement volatility, supplier dependency, and the absence of predictive maintenance contribute to inefficiencies. The study recommends the adoption of bulk procurement contracts, budget forecasting using R and predictive maintenance systems to enhance cost stability. Integration of Power BI dashboards into management operations is also advised to improve cost visibility and decision-making. These strategies will collectively improve operational efficiency and long-term financial sustainability within the manufacturing environment.

REFERENCES

- Caniato, F., & Ronchi, S. (2007). Supply chain management in the spare parts industry. *Journal of Purchasing and Supply Management*, 13(1), 34–45.
- Fernando, L., & Jayasekara, P. (2021). Predictive maintenance for cost optimization in industrial environments. *International Journal of Production Research*, 59(10), 2985–3002.
- Gupta, S., Patel, R., & Singh, A. (2019). Spare part cost control in manufacturing industries. *Journal of Industrial Engineering*, 45(3), 201–212.
- Huiskonen, J. (2001). Maintenance spare parts logistics: Special characteristics and strategic choices. *International Journal of Production Economics*, 71(1–3), 125–133.
- Huang, Y., Liu, X., & Zhang, D. (2017). Life cycle cost analysis and predictive maintenance in manufacturing systems. *Journal of Manufacturing Systems*, 45, 20–31.

- Kennedy, W., Patterson, J., & Fredendall, L. (2002). An overview of recent literature on spare parts management. *International Journal of Production Economics*, 76(2), 201–215.
- Perera, S., & Silva, R. (2020). ERP-driven analytics for spare part procurement in Sri Lankan industries. *Journal of Applied Management Research*, 12(4), 66–78.
- Silver, E. A., Pyke, D. F., & Peterson, R. (1998). *Inventory Management and Production Planning and Scheduling*. New York: Wiley.
- Wickramasinghe, D., & Fernando, R. (2020). Challenges in spare parts management in Sri Lankan manufacturing firms. *Asian Journal of Operations Research*, 9(2), 55–68.



Factors Associated with Self-reported Health in University Students: A Case Study

Alweera AMIH¹, Panahatipola PMOP¹

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
imashahanshani1999@gmail.com

ABSTRACT

University students often face various lifestyle, dietary, and psychological challenges that impact their academic performance and overall wellbeing. This investigation explored the primary elements linked to self-reported health among undergraduates at Wayamba University of Sri Lanka, filling a void in local studies that incorporate various behavioral and psychosocial factors. A cross-sectional survey involving 100 students was carried out utilizing a structured questionnaire designed to evaluate dietary habits, physical activity, sleep patterns, stress levels, and general health perception, which was measured using a 5-point Likert scale. The analysis of data involved the application of descriptive statistics, correlation analysis, Principal Component Analysis (PCA), and Exploratory Factor Analysis (EFA). The findings revealed that healthier habits like consistent exercise, engaging in outdoor activities, and improved diet quality were positively correlated with perceived health. In contrast, factors such as stress, reduced sleep duration, and frequent illness exhibited negative correlations. Students residing away from their homes indicated notably elevated levels of stress. The analysis through PCA and EFA consistently revealed three primary dimensions affecting self-reported health: routine and nutrition, physical wellness, and psychosocial satisfaction, which together accounted for a significant portion of the variance. The results emphasize the complex relationships among lifestyle choices, emotional health, and daily habits in influencing students' views on health.

KEYWORDS: Health Behavior, Lifestyle Factors, PCA, Student Wellbeing, Stress

1 INTRODUCTION

Students' physical and mental health may suffer as a result of the significant lifestyle changes that come with the adjustment to university life. Stress, unhealthy eating patterns, and a decrease in physical activity are all exacerbated by increased academic pressure, unfamiliar social settings, and the lack of established routines. If continued, these actions may have a long-term impact on students' academic achievement and general well-being.

Few quantitative studies have examined the combined effects of

lifestyle and psychosocial factors on students' self-reported health in Sri Lanka. Multivariate statistical techniques are used in this study to determine the fundamental factors influencing Wayamba University undergraduates' perceptions of their health.

2 LITERATURE REVIEW

2.1 Literature Related to the Study

Research indicates that university students experience lifestyle changes that lead to health problems, including stress, obesity, and irregular dietary habits (Smith & Taylor, 2019). Pereira-Santos

et al. (2016) identified a correlation between elevated fast-food consumption among students and reported declines in health and increases in fatigue. Pop et al. (2021) demonstrated that regular exercise and healthy routines enhance perceived wellbeing, whereas Neuman et al. (2024) underscored the significance of health behaviour in academic performance. In the context of Sri Lanka, there is a scarcity of comprehensive statistical models that identify various health-related factors. This study employs PCA and EFA to identify the primary factors influencing students' health perceptions.

2.2 Theoretical Background

Exploratory Factor Analysis (EFA) and Principal Component Analysis (PCA) are commonly employed methodologies for identifying latent structures within correlated datasets. PCA minimises the number of observed variables while preserving maximum variance, whereas EFA uncovers underlying constructs that account for shared variance. The integration of these methods elucidates behavioural and psychosocial patterns that influence perceived health.

3 METHODOLOGY

Wayamba University of Sri Lanka undergraduates participated in a quantitative cross-sectional study. A structured, anonymous Google Forms questionnaire covering demographics, diet, exercise, sleep, stress, lifestyle choices, and self-reported health on a 5-point Likert scale (1 = Poor, 5 = Excellent) was used to gather the data. Self-reported weight and height were used to compute the Body Mass Index (BMI), which offers an unbiased

assessment of nutritional status. The widely recognized recommendations for factor analysis, which call for a minimum sample size of 100 for stable extraction and at least 5–10 participants per variable, are satisfied by a sample of 100 participants. This sample is suitable for PCA and EFA due to the exploratory nature of the study and sufficient communalities (>0.40).

3.1 Data Analysis

Statistical procedures were conducted using R software for data analysis. Descriptive statistics summarised respondent characteristics, while Pearson correlations identified relationships among lifestyle factors. Principal Component Analysis (PCA) and Exploratory Factor Analysis (EFA) were subsequently conducted to discern latent components influencing health perception. The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett's Test of Sphericity were used to assess the dataset's suitability for factor analysis. The KMO statistic evaluates whether the correlation patterns are sufficiently compact to produce unique and trustworthy factors. A significant result verifies that the variables are sufficiently correlated for factor analysis. Bartlett's Test of Sphericity investigates the null hypothesis that the correlation matrix is an identity matrix, indicating no correlations among variables.

Components with eigenvalues larger than one were kept after data suitability was verified, and Varimax rotation was used to enhance the extracted factor structure's readability and clarity.

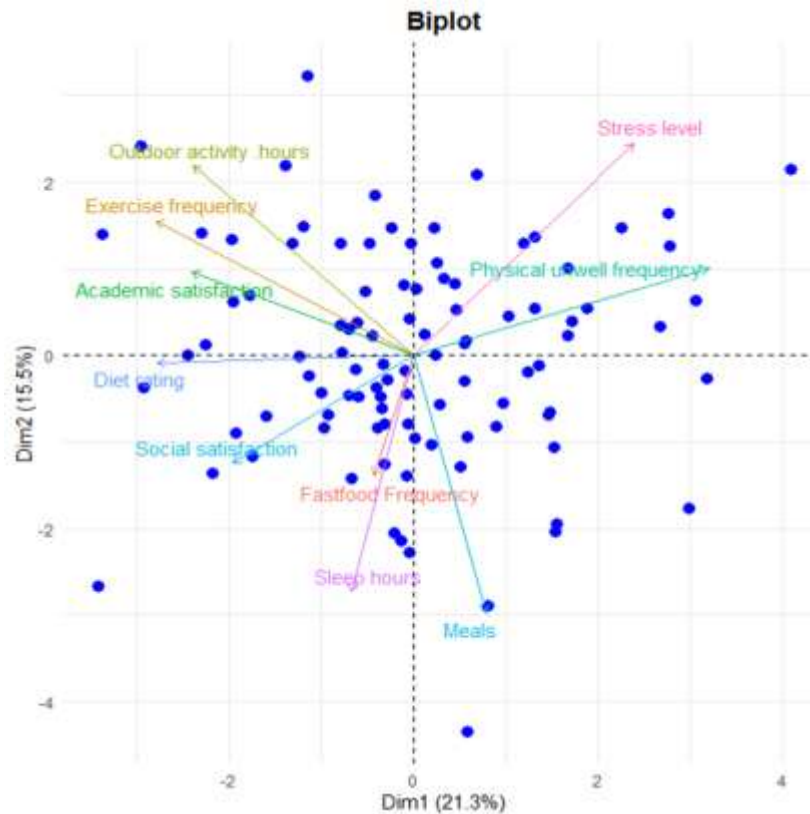


Figure 1: Biplot of Health-Related Factors and Their Interrelationships

4 RESULTS AND DISCUSSION

4.1 Descriptive Analysis

The participants' average height was 162.1 cm, and their average weight was 60 kg, resulting in a mean BMI of 22.7, which is classified as healthy. The majority of students were of normal weight (62%), whereas a minority were underweight (10%) or obese (3%). Cross-tabulation between BMI and self-reported health indicated that students within the normal BMI range were more likely to report good or excellent health compared to those who were overweight or underweight.

4.2 Correlation Analysis

Correlation analysis demonstrated substantial associations among lifestyle, psychological, and academic variables. Exercise frequency exhibited moderate positive correlations with outdoor

activity hours ($p = 0.001$) and diet rating ($p = 0.021$), indicating that physically active students tend to uphold healthier dietary patterns. The stress level exhibited a positive correlation with the frequency of Physical illness ($p = 0.008$) and negative correlations with sleep duration ($p = 0.043$) and social satisfaction ($r = -0.23$). Academic and social satisfaction were positively related ($p = 0.017$), demonstrating the interdependence of academic engagement and social well-being. Self-reported health and gender did not significantly associate, according to chi-Square results ($p = 0.2021$). Nonetheless, there was a significant association between living status and stress level ($p = 0.019$), with students who lived away from family reporting higher levels of stress. BMI did not significantly differ by gender or perceived health categories,

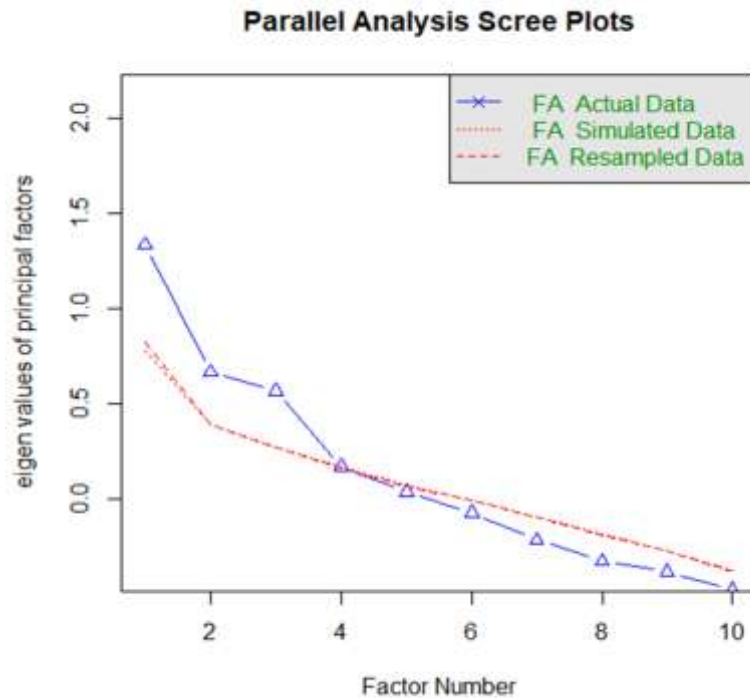


Figure 2: Scree Plot from Parallel Analysis

according to independent t-tests (p value= 0.115) and ANOVA (p value= 0.26), $p > 0.05$, indicating that perceptions of health may be influenced by factors other than BMI.

4.3 Principal Component Analysis

Ten factors pertaining to health and lifestyle were examined. Fifty-one percent (51%) of the variance was explained by the first three principal components.

- PC1-General Wellness- (21%): Indicates general well-being and is strongly correlated with diet, exercise frequency, academic and social satisfaction, and negatively correlated with stress and illness.
- PC2-Routines- (16%): Shows regular stability and is positively correlated with meal frequency and sleep duration.
- PC3-Eating Habits- (11%): Frequent consumption of fast

food predominates, suggesting unhealthy eating habits.

Clustering between negative health indicators (stress, fast-food consumption) and healthy behaviors (exercise, diet, and satisfaction) was demonstrated by the PCA biplot.

4.4 Exploratory Factor Analysis

A three-factor explanation for 33% of the variance was validated by EFA.

Factor 1: Diet and exercise (loadings: outdoor activity 0.62; exercise 0.72).

Factor 2: Stress and routine (stress -0.47), meals- 0.40, and sleep hours- 0.41.

Factor 3: Psychosocial satisfaction is the second factor (social satisfaction = 0.57; academic satisfaction = 0.73).

An acceptable model fit was shown by the fit indices (RMSR = 0.05; RMSEA = 0.06; TLI = 0.80). The findings show that the interaction of

physical, emotional, and lifestyle factors influences students' self-reported health. Overall, these findings collectively support theoretical perspectives that emphasize the multidimensional nature of student health and well-being. These results indicate that interventions aiming to improve student well-being should integrate behavioral, psychological, and social components rather than focusing solely on physical indicators.

5 CONCLUSION

The study emphasizes that a variety of factors, including stress management, physical activity, and psychosocial satisfaction, influence university students' self-reported health in addition to their BMI. Three essential elements were identified by factor analysis: (1) routine and nutrition; (2) psychosocial satisfaction; and (3) physical wellness and activity. These results imply that both physical and psychological wellness techniques should be incorporated into health interventions in academic settings. Students' academic performance and overall well-being can be improved by encouraging regular exercise, healthy eating, sufficient sleep, and stress-reduction techniques. The study is limited by its modest sample size, which may reduce generalizability, and reliance on self-reported data, which may introduce reporting bias. Future studies should include larger and more diverse samples and consider longitudinal designs.

REFERENCES

- Ahmad, S. R., Yunus, T. C., & Quee, D. K. S. (2018). Breakfast consumption and its relationship to sociodemographic and lifestyle factors of undergraduate students in the *School of Health Sciences at the University of KwaZulu-Natal* [Unpublished manuscript].
- Khaled, B., & Iwona, S. (2021). Dietary patterns and associated lifestyle factors among university students in Qatar [PDF].
- López-Olivares, M., De Teresa Galván, C., Nestares, T., Fernández-Gómez, E., & Enrique Mirón, C. (2021). Lifestyle factors influencing dietary patterns of university professors. *International Journal of Environmental Research and Public Health*, 18(18), 9777.
- Meyer, J., Nauck, M., & Musil, V. (2022). Dietary intake and physical activity of German university students. *Journal of Public Health*, 30(5), 123–134.
- Munawar, Y. G., & Lontoh, S. O. (2021). The association between physical activity and the nutritional status of students in the Faculty of Medicine Tarumanagara University in 2019–2020. In *Proceedings of the 1st Tarumanagara International Conference on Medicine and Health (TICMIH 2021)*. Atlantis Press.
- Neuman, J., Ina, E. A., Huq, S. O., Blanca, A., & Petrosky, S. N. (2024). Cross-sectional analysis of the effect of physical activity, nutrition, and lifestyle factors on medical students' academic achievement. *Cureus*, 16(3).
- Pereira-Santos, M., Da Mota Santana, J., Neves De Carvalho, A. C., & Freitas, F. (2016). Dietary patterns among nutrition students at a public university in Brazil. *Revista Chilena de Nutrición*, 43(1), 39–44.
- Pop, L., Iorga, M., Muraru, I., & Petrariu, F. (2021). Assessment of dietary habits, physical activity, and lifestyle in medical university students. *Sustainability*, 13(6), 3572.
- Richardson, G. E., Blacklock, R. G., Wyke, S., & Lean, M. E. J. (2018). Dietary patterns of university students in the UK: A cross-sectional study. *Nutrition Journal*, 17, 90.
- Smith, J., & Taylor, M. (2019). Dietary patterns and lifestyle choices among college students: Implications for health and academic performance. *Nutrition and Health*, 34(2), 101–116.

Faculty of Applied Sciences
Wayamba University of Sri Lanka
Kuliyapitiya

ISBN 978-624-6674-13-7 (Paper)
ISBN 978-624-6674-14-4 (Online)

ISSN 2279-1558 (Paper)
ISSN 2989-0683 (Online)