AI AND ML SOLUTIONS DRIVING MODERN FARMING AND URBAN INNOVATION

Editors:
Suneeta Satpathy
Sachi Nandan Mohanty
Subhendu Kumar Pani

Bentham Books

Applied Artificial Intelligence in Data Science, Cloud Computing and IoT Frameworks

(Volume 3)

AI and ML Solutions Driving Modern Farming and Urban Innovation

Edited by

Suneeta Satpathy
Center For Cybersecurity, SoA Deemed to be University, Bhubaneswar, Odisha, India

Sachi Nandan Mohanty
School of Computer Science & Engineering, VIT-AP University, Andhra Pradesh, India

&

Subhendu Kumar Pani
Computer Science & Engineering Krupajal Engineering College BPUT, Odisha, India

Applied Artificial Intelligence in Data Science, Cloud Computing and IoT Frameworks

(Volume 3)

AI and ML Solutions Driving Modern Farming and Urban Innovation

Editors: Suneeta Satpathy, Sachi Nandan Mohanty and Subhendu Kumar Pani

ISSN (Online): 3029-2247

ISSN (Print): 3029-2255

ISBN (Online): 979-8-89881-210-2

ISBN (Print): 979-8-89881-211-9

ISBN (Paperback): 979-8-89881-212-6

Published by Bentham Science Publishers Pte. Ltd. Singapore, in collaboration with Eureka Conferences, USA.

First published in 2025.

Bentham Science Publishers Pte. Ltd.
No. 9 Raffles Place
Office No. 26-01
Singapore 048619
Singapore
Email: subscriptions@benthamscience.net

CONTENTS

FOREWORD i

PREFACE ii

LIST OF CONTRIBUTORS iii

SECTION 1 MODERN FARMING WITH ADVANCED INTELLIGENCE

CHAPTER 1 A REVIEW ON PADDY PLANT DISEASE DETECTION USING VARIOUS MACHINE LEARNING AND DEEP LEARNING ALGORITHMS 1

Pokkuluri Kiran Sree and *A.L. Kalyani*

INTRODUCTION 1
Fungal Diseases 2
Bacterial Diseases 3
Viral Diseases 4
LITERATURE SURVEY 6
ALGORITHMIC ANALYSIS 8
Image Acquisition and Preprocessing 8
Segmentation 8
K-means Clustering 9
Feature Extraction 9
Image Classification 9
Detection/Post-Processing 13
CONCLUSION 13
REFERENCES 13

CHAPTER 2 REVIEW ON AN INTELLIGENT SYSTEM FOR QUALITY EVALUATION OF DRY FRUITS USING HYPERSPECTRAL IMAGING 16

Shweta Bhelonde and *Manoj B. Chandak*

INTRODUCTION 17
MOTIVATION AND CONTRIBUTIONS 18
REVIEW OF EXISTING MODELS 20
ANALYTICAL COMPARISON OF THE REVIEWED MODELS 30
CONCLUSION AND FUTURE SCOPE 36
REFERENCES 38

CHAPTER 3 DEEP LEARNING FUSION: POTATO LEAF DISEASE CLASSIFICATION THROUGH GENERATIVE DIFFUSION MODEL 42

Virendra Kumar Shrivastava, Prakruthi Ganiga, Rathnakar Achary, Chetan J. Shelke, A. Ezil Sam Leni and *Ramalakshmi K.*

INTRODUCTION 42
RELATED WORK 44
METHODOLOGY 45
Dataset 45
Data Preprocessing 46
Data Augmentation with Diffusion Model 47
Deep Learning Model for Classification of Images 47
RESULT AND DISCUSSION 48
CONCLUSION 50
REFERENCES 50

CHAPTER 4 FUZZIFICATION FOR PRECISION FARMING WITH MINIMAL HUMAN INTERVENTION 54

Sudheer Mangalampalli, Ganesh Reddy Karri, Pelluru Pavan Kumar Reddy, Ramith Yadav S., K. Varada Rajkumar and *Kiran Sree Pokkuluri*
INTRODUCTION 54
LITERATURE REVIEW 55
PROPOSED ARCHITECTURE 56
Control System 56
Fuzzy Logic 57
METHODOLOGY 58
RESULTS AND DISCUSSION 62
CONCLUSION 63
ACKNOWLEDGEMENTS 64
REFERENCES 64

CHAPTER 5 ENHANCING FEEDFORWARD NEURAL NETWORK OPTIMIZER VARIANTS TO STRENGTHEN AQUAPONICS FISH GROWTH PREDICTIONS 66
Safiya Begam, W. Aisha Banu and *Sharon Priya*
INTRODUCTION 66
LITERATURE REVIEW 68
FNN, OPTIMIZATION STRATEGIES AND DATASETS 71
SGD 73
AdaGrad 73
AdaDelta 73
Adam 74
DATASET 74
METHODOLOGY 74
RESULTS AND DISCUSSION 75
CONCLUSION AND FUTURE WORK 78
REFERENCES 79

CHAPTER 6 PREDICTING CARBON DIOXIDE EMISSIONS FROM GREEN WASTE COMPOSTING AND IDENTIFYING KEY FACTORS USING MACHINE LEARNING ALGORITHMS 81
M. Vaishali, M. Monika, M. Atish and *S. Rajarshi*
INTRODUCTION 81
LITERATURE SURVEY 83
PROPOSED METHODOLOGY 84
Dataset Collection 84
Data Preprocessing and Filtering 84
Applying Machine Learning Classifiers for CO2 Prediction from Agri-waste 85
Identifying Key Factors using Exploratory Data Analysis (EDA) 87
RESULTS AND DISCUSSION 88
CONCLUSION AND FUTURE SCOPE 91
REFERENCES 91

CHAPTER 7 CAN ETHIOPIAN COOPERATIVE COFFEE FARMERS PREDICT FINANCIAL MANAGEMENT DECISIONS THROUGH HUMAN CAPITAL AND AI MEDIATION? 93
Mershaye Birhane, Shashi kant, Brehanu Borji and *Chalchisa Amentie*
INTRODUCTION 93
THEORETIC ESTABLISHMENT 94
EMPIRICAL LITERATURE REVIEW 95
Fiscal Supervision and Business Success 95

Management of Finance and Investment in Humans ... 95
Financial Management and AI ... 96
Human Capital and Organization Performance ... 96
AI and Organization Performance ... 97
Human Capital Mediation ... 98
Mediation of AI between FM strategies and OP ... 98
RESEARCH METHODOLOGY ... 99
Description of the Study Area ... 99
Study Population, Sample Size, and Data Sources ... 99
Sampling Practices ... 99
Data Analysis Tool ... 100
Data Adequacy Test ... 100
Confirmatory Factory Analysis ... 100
Model Fitness ... 102
CONCLUDING REMARK ... 102
MANAGERIAL AND PRACTICAL IMPLICATIONS ... 103
AUTHORS' CONTRIBUTIONS ... 104
REFERENCES ... 104

SECTION 2 INTELLIGENT AI & ML SOLUTIONS

CHAPTER 8 SMART CITIES: URBAN PLANNING AND INFRASTRUCTURE MANAGEMENT ... 108
,Aanchal Chaudhary, Harsh Kumar Pandey, Ayush Yadav, Vaibhav Kumar Singh Nishant Singh and *Hitesh Mohapatra*
INTRODUCTION ... 108
Waste Management ... 110
Air Quality ... 110
Noise Monitoring ... 110
Smart Parking ... 110
City Energy Consumption ... 111
LITERATURE REVIEW ... 111
RESEARCH GAP ... 114
Security and Privacy ... 114
Smart Sensors ... 114
Networking ... 115
Big Data Analytics ... 115
PROPOSED STUDY ... 115
CONCLUSION ... 122
ACKNOWLEDGEMENTS ... 123
REFERENCES ... 123

CHAPTER 9 NEXT-GENERATION URBAN MOBILITY MANAGEMENT: A FRAMEWORK FOR INTELLIGENT TRAFFIC CONTROL ... 126
Sujithra Thandapani, Durai Selvaraj, Mohamed Iqbal M., D. Jaganathan and *V.C. Bharathi*
INTRODUCTION ... 127
Shortcomings in the Existing System ... 130
PROPOSED SYSTEM ... 130
Proposed Edge device ... 130
Automatic Green Light Enabler ... 130
Dynamic Signaling ... 133

IMPLEMENTATION AND EXPERIMENTAL RESULTS 135
Automatic Green Light Enabler 135
Dynamic Scheduling 137
CONCLUSION 139
REFERENCES 140

CHAPTER 10 PREDICTING CUSTOMER CHURN IN THE TELECOM INDUSTRY WITH MACHINE LEARNING TECHNIQUES FOR IMPROVED RETENTION STRATEGIES 142
Smital Anandrao Patil, Rushi Patel, Swati Shukla, Yeswanth Satya Prasad and *Yagnesh Challagundla*
INTRODUCTION 143
RELATED WORK 144
METHODOLOGY 145
Data Collection 146
Data Preprocessing 146
Exploratory Data Analysis (EDA) 146
Development of a Model 146
Evaluation of the Model 146
Importance of Features 147
RESULTS 147
Data Analysis 147
Data Preprocessing 148
Construction and Assessment of Models 149
Neural Networks 149
Gradient Boosting 149
Naïve Bayes 149
Principal Discoveries 149
Contract Type 149
Internet Service 149
Summary 149
CONCLUSION 150
REFERENCES 151

CHAPTER 11 A WATER QUALITY PREDICTION AND ASSESSMENT MODEL USING MACHINE LEARNING CLASSIFIERS 153
M. Vaishali, M. Monika, M. Atish and *Lalitha*
INTRODUCTION 154
LITERATURE SURVEY 155
PROPOSED METHODOLOGY 156
Dataset Collection 157
Applying ML Classifiers for Classification and Prediction 157
Naïve Bayes 158
KNN 158
Support Vector Machine (SVM) 158
Decision Tree 159
Random Forest 159
XGBoost 159
Dataset Preparation and Preprocessing 160
Evaluation Criteria 160
RESULTS AND DISCUSSION 161
ROC Curve 161
Statistical Summary 162

Selecting Essential Predictors for the Classifier Models 164
Confusion matrix of WQI model 164
CONCLUSION AND FUTURE SCOPE 166
REFERENCES 166

CHAPTER 12 SYNAPSYS – A TOOL FOR EFFECTIVE VIDEO SUMMARIZATION 169
S. Divya Meena, B.V. Chandrahaas, C. Sai Janakiram, J.L.B. Krishna, M. Rushane and *J. Sheela*
INTRODUCTION 169
Related Work 170
Divya et al. 170
PROPOSED METHODOLOGY 171
Data Collection 171
Synapsys Architecture 171
Whisper 172
Pegasus 174
RESULTS AND DISCUSSION 175
CONCLUSION AND FUTURE SCOPE 177
REFERENCES 178

CHAPTER 13 HDFS: THE BACKBONE OF BIG DATA - A REVIEW OF HIGH AVAILABILITY, SCALABILITY, AND PERFORMANCE USING QUORUM JOURNAL MANAGER AND DYNAMIC FEDERATED METADATA MANAGEMENT 180
Bhagavathi Santos Kumar and *Virendra Kumar Shrivastava*
INTRODUCTION 181
BACKGROUND 182
QUORUM JOURNAL MANAGER 183
HDFS FEDERATION 185
Key Features: 186
DYNAMIC FEDERATED METADATA MANAGEMENT 187
The advantages of DFMM are: 188
Metadata Management Techniques 189
CONTRAST BETWEEN QUORUM JOURNAL MANAGER (QJM) AND DYNAMIC FEDERATED METADATA MANAGEMENT (DFMM) 191
IMPLICATIONS AND FINDINGS 192
Specific Use Cases and Preferences 192
CONCLUSION AND FUTURE DIRECTION 192
REFERENCES 193

CHAPTER 14 OPTIMIZED GOLD PRICE PREDICTION USING PARTICLE SWARM OPTIMIZATION-ENHANCED LSTM NETWORKS 196
Alladi Vennela and *Radha Mohan Pattanayak*
INTRODUCTION 196
LITERATURE REVIEW 197
METHODOLOGY 198
Data Collection 198
Data Preparation 198
Particle Swarm Optimisation (PSO) 199
Train Data & Test Data 200
Long Short-Term Memory (LSTM) network 200
Evaluation and Comparison 201
RESULTS AND DISCUSSIONS 201

Results 201
Performance Metrics 201
Discussion 202
CONCLUSION 202
REFERENCES 203

SUBJECT INDEX 206

FOREWORD

In an era defined by technological convergence, Applied Artificial Intelligence in Data Science, Cloud Computing and IoT Frameworks: Volume 2 – AI and ML Solutions Driving Modern Farming and Urban Innovation emerges as a timely and visionary contribution to the global discourse on intelligent transformation. This volume arrives at a moment when Artificial Intelligence (AI) and Machine Learning (ML) have evolved beyond research laboratories to become foundational forces shaping industries, economies, and societies. Their integration with Cloud Computing and the Internet of Things (IoT) is driving unprecedented innovation, enabling sustainable solutions, and redefining the way humanity interacts with its environment.

The agricultural and urban sectors, in particular, lie at the heart of this transformation. Confronted by challenges such as food insecurity, climate change, urban congestion, and resource inefficiency, these domains demand solutions that blend intelligence with sustainability. AI-powered analytics, when coupled with the scalability of cloud infrastructures and the real-time responsiveness of IoT ecosystems, provide a powerful framework for addressing these issues. From precision agriculture that optimizes yield while conserving resources, to smart city systems that enhance mobility and reduce emissions, the innovations explored within this volume exemplify how data-driven intelligence can catalyze equitable and sustainable progress.

Distinguished by its synthesis of theoretical depth and practical insight, this book bridges the gap between conceptual understanding and real-world implementation. Each chapter presents rigorous analyses alongside applied frameworks, case studies, and experimental outcomes that illuminate the tangible impact of AI and ML technologies across rural and urban landscapes. By integrating cross-disciplinary perspectives, the volume offers researchers, practitioners, and policymakers a comprehensive guide to advancing smarter, greener, and more resilient ecosystems.

Beyond its technical merit, this volume stands as a reflection of collective human ingenuity directed toward the greater good. It envisions a future where innovation serves not only as a tool of progress but as a means of harmonizing technological advancement with environmental and societal well-being. The editors and contributors have assembled a body of work that captures both the promise and responsibility of intelligent systems in shaping the next era of sustainable development.

Shu Hu
School of Applied and Creative Computing
Purdue Machine Learning and Media Forensics (M2) Lab
Purdue University
West Lafayette, Indiana

PREFACE

The rapid evolution of Artificial Intelligence (AI) and Machine Learning (ML) is reshaping the way societies address some of their most pressing challenges. This volume, ***AI and ML Solutions Driving Modern Farming and Urban Innovation***, brings together a diverse collection of research contributions that highlight how intelligent technologies are transforming both agricultural practices and urban living.

Modern farming today demands precision, efficiency, and sustainability. AI and ML are enabling farmers to optimize crop yields, predict diseases, and manage resources with unprecedented accuracy. Similarly, cities around the world are turning to intelligent systems to tackle complex issues such as waste management, water quality monitoring, traffic congestion, and environmental sustainability. Together, these innovations demonstrate the potential of AI and ML to create smarter, healthier, and more resilient communities.

This volume serves as a platform for scholars, practitioners, and innovators to share advancements, methodologies, and applications that bridge the gap between research and real-world deployment. By showcasing interdisciplinary solutions, it emphasizes the societal, local, and national impacts of intelligent technologies in agriculture and smart city ecosystems.

We hope this collection inspires further exploration and collaboration toward building sustainable futures powered by AI and ML.

Suneeta Satpathy
Center For Cybersecurity, SoA Deemed to be University
Bhubaneswar, Odisha, India

Sachi Nandan Mohanty
School of Computer Science & Engineering
VIT-AP University, Andhra Pradesh, India

&

Subhendu Kumar Pani
Computer Science & Engineering, Krupajal Engineering College BPUT
Odisha, India

List of Contributors

Alladi Vennela	School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh, 522237, India
A.L. Kalyani	Department of CSE, Shri Vishnu Engineering College for Women(A), Bhimavaram, Andhra Pradesh, 534202, India
A. Ezil Sam Leni	Centre of Excellence in Computer Vision, Department of Computer Science and Engineering, School of Advanced Computing, Alliance University, Bangalore, Karnataka, 562106, India
Aanchal Chaudhary	School of Computer Engineering, KIIT (Deemed to Be) University, Bhubaneswar, 751024, Odisha, India
Ayush Yadav	School of Computer Engineering, KIIT (Deemed to Be) University, Bhubaneswar, 751024, Odisha, India
Brehanu Borji	Department of Management, Hawasa University, Hawasa, Ethiopia
B.V. Chandrahaas	School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India
Bhagavathi Santos Kumar	Department of Computer Science and Engineering, School of Advanced Computing, Alliance University, Bangalore, Karnataka, 562106, India
Chetan J. Shelke	Centre of Excellence in Computer Vision, Department of Computer Science and Engineering, School of Advanced Computing, Alliance University, Bangalore, Karnataka, 562106, India
Chalchisa Amentie	Department of Management, Ethiopian Civil Service University, Addis Ababa, Ethiopia
C. Sai Janakiram	School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India
Durai Selvaraj	Department of CSE, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India
Ganesh Reddy Karri	SCOPE, VIT-AP University, Andhra Pradesh, India
Harsh Kumar Pandey	School of Computer Engineering, KIIT (Deemed to Be) University, Bhubaneswar, 751024, Odisha, India
Hitesh Mohapatra	School of Computer Engineering, KIIT (Deemed to Be) University, Bhubaneswar, 751024, Odisha, India
J.L.B. Krishna	School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India
J. Sheela	School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India
K. Varada Rajkumar	Department of CSE (AI&ML), MLR Institute of Technology, Telangana, India
Kiran Sree Pokkuluri	Department of CSE, Shri Vishnu Engineering College for Women, Bhimavaram, India
Lalitha	School of Advanced Sciences, VIT-AP University, Andhra Pradesh, India

Manoj B. Chandak	Department of Computer Science and Engineering, Ramdeo Baba University, Nagpur, Maharashtra, India
M. Vaishali	Department of Computer Science and Engineering, Maharishi Markandeshwar Deemed to be University, Ambala, India
M. Monika	Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Navi Mumbai, India
Mershaye Birhane	Department of Management, Bule Hora University, Bule Hora, Ethiopia
Mohamed Iqbal M.	Department of Information Technology, Sri Sivasubramaniya Nadar College of Engineering, Chennai, 603110, India
M. Vaishali	Department of Computer Science and Engineering, Maharishi Markandeshwar Deemed to be University, Ambala, India
M. Monika	Department of Information Technology, Dwarkadas J Sanghvi College of Engineering, Navi Mumbai, India
M. Atish	Department of Computer Science and Engineering, BMS Institute of Technology and Management, Banglore, India
M. Rushane	School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India
Nishant Singh	School of Computer Engineering, KIIT (Deemed to Be) University, Bhubaneswar, 751024, Odisha, India
Prakruthi Ganiga	Centre of Excellence in Computer Vision, Department of Computer Science and Engineering, School of Advanced Computing, Alliance University, Bangalore, Karnataka, 562106, India
Pelluru Pavan Kumar Reddy	SCOPE, VIT-AP University, Andhra Pradesh, India
Rathnakar Achary	Centre of Excellence in Computer Vision, Department of Computer Science and Engineering, School of Advanced Computing, Alliance University, Bangalore, Karnataka, 562106, India
Ramalakshmi K.	Centre of Excellence in Computer Vision, Department of Computer Science and Engineering, School of Advanced Computing, Alliance University, Bangalore, Karnataka, 562106, India
Ramith Yadav S.	SCOPE, VIT-AP University, Andhra Pradesh, India
Rushi Patel	Department of Electronic and Telecommunication Engineering, Indian Institute of Technology - IIT, Kharagpur, 721302, West Bengal, India
Radha Mohan Pattanayak	School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh, 522237, India
Shweta Bhelonde	Department of Computer Science and Engineering, Ramdeo Baba University, Nagpur, Maharashtra, India
Sudheer Mangalampalli	Department of CSE, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India
Safiya Begam	B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India

Sharon Priya	B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India
S. Rajarshi	School of Advanced Sciences, VIT-AP University, Andhra Pradesh, India
Shashi Kant	Department of Management, Bule Hora University, Bule Hora, Ethiopia
S. Divya Meena	School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India
Sujithra Thandapani	Department of CSE, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India
Swati Shukla	School of Electronics Engineering (SENSE), VIT-AP University, Amaravati, 522237, Andhra Pradesh, India
Smital Anandrao Patil	School of Computer Science, KIT's College of Engineering, Kolhapur, 416234, Maharashtra, India
Vaibhav Kumar Singh	School of Computer Engineering, KIIT (Deemed to Be) University, Bhubaneswar, 751024, Odisha, India
V.C. Bharathi	School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India
Virendra Kumar Shrivastava	Department of Computer Science and Engineering, School of Advanced Computing, Alliance University, Bangalore, Karnataka, 562106, India
W. Aisha Banu	B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India
Yeswanth Satya Prasad	School of Computer Science and Engineering (SCOPE), VIT University, Vellore, 632014, Tamil Nadu, India
Yagnesh Challagundla	School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, 522237, Andhra Pradesh, India

Section 1
Modern Farming with Advanced Intelligence

CHAPTER 1

A Review on Paddy Plant Disease Detection Using Various Machine Learning and Deep Learning Algorithms

Pokkuluri Kiran Sree[1,*] and **A.L. Kalyani**[1]

[1] *Department of CSE, Shri Vishnu Engineering College for Women(A), Bhimavaram-534202, Andhra Pradesh, India*

Abstract: Early detection of plant diseases is critical in the agricultural sector to reduce the loss caused by the diseases. Paddy crop diseases impact yields in various ways, given that the crop is cultivated globally. The specific disease manifestations and severity, however, often vary with climatic and geographical conditions. The method used to treat the disease in one area may not be suitable in other areas. The manual detection of diseases by farmers is a very costly, time-consuming, and tedious process. To avoid these issues, researchers have developed various automatic methods. Most models utilize image processing and machine learning, or deep learning algorithms, to achieve the designated output. This study analyzed various algorithms used to automate the plant disease detection process, examining their advantages and disadvantages.

Keywords: Paddy crop diseases, Arrange alphabetically, Machine learning, Deep learning, Image processing.

INTRODUCTION

Agriculture is the foundational pillar on which many human societies have lived for thousands of years. Agriculture is the primary source of nourishment for billions of people worldwide by providing essential foods. Agriculture impacts various aspects of human life, including the economy, employment, and environment, particularly in rural areas. The agricultural sector also includes agribusiness, food processing, distribution, and farming. Out of all the crops growing worldwide, more than 50% of the world's population depends on paddy as a primary food crop. Paddy stands at the forefront of all crops, ensuring food security and providing nutritional food rich in calories, essential for the global population. Paddy crops are the primary income source for most farmers in the countryside, and they also support various industries along the agricultural supply

* **Corresponding author Pokkuluri Kiran Sree:** Department of CSE, Shri Vishnu Engineering College for Women(A), Bhimavaram-534202, Andhra Pradesh, India; E-mail: drkiransree@gmail.com

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

chain. Paddy also holds cultural and traditional significance for worldwide communities. Paddy is a source of Carbohydrates and nutrients are the basic food for most of the global population. Paddy plants can grow in various environments, including lowland flooded areas and highland regions. Additionally, it serves as a good food source for diverse geographical and climatic conditions.

Despite the importance of paddy crops, paddy diseases have a significant impact on production and farmers' lives. It is essential to understand the disease's causes, diagnose its symptoms, and provide better treatment methods in the early stages of the crop. Manual diagnosis of the disease by the farmer's naked eye may not be sufficient, as the farmer must inspect every plant individually, which can be time-consuming and require a significant workforce. This may even lead to the loss of an entire crop if not properly taken care of by the farmer. All these manual processes are being automated using machine learning algorithms, deep learning algorithms, and other advanced technologies. Almost all of the proposed models failed to detect the diseases in the main field. All the models were trained and tested using the controlled images with transparent backgrounds and single leaves. These models should be trained on field images, enabling farmers to utilize them directly to save time and effort and achieve more accurate disease detection results.

Paddy Crop Diseases: Over 50% of the world's population gets food from paddy crops. Paddy crops will be affected by various diseases due to climatic conditions. The diseases vary depending on the climate and geographical conditions. The major diseases that affect the paddy crop are discussed below. Such diseases are mainly classified into:

Fungal Diseases

Blast:

Also called Pyricularia oryzae. The fungus will attack the crop at all growth stages in this disease. The disease symptoms will appear on leaves, nodes, and glumes. Blast is one of the most destructive paddy diseases, and it may cause grain loss of up to 60% -70%. The most favorable conditions are temperatures between 25-28 degrees, high humidity, high nitrogen levels, frequent rainfall, and a cloudy sky.

Brown Spot (Helminthosporium oryzae):

This disease is caused by a fungus that attacks the crop from the seedling stage to the milky stage. The disease symptoms will appear as minute spots on the coleoptile, leaf, and glume. The spots will be cylindrical, which later become circular and dark brown. Spots that appear on glumes cause grain discoloration, leading to seed germination failure and a reduction in grain quality and weight. While Excess nitrogen increases the severity of the disease, humidity over 80% and temperatures between 25-30 degrees are favorable conditions.

Sheath Rot:

This disease is also called Sarocladium oryzae. Its symptoms appear on the upper leaf. The disease will occur in irregular, greyish-brown spots, with an enlarged grey center and brown margins covering significant portions of the leaf. High nitrogen levels, close plantation, and temperatures between 25 and 30 degrees are favorable conditions. Through airborne and seed-borne conidia, the disease spreads.

Sheath Blight (Rhizoctonia solani):

From tillering to the heading stage, the fungus will affect the crop. The disease symptoms will appear on the leaf, typically near the water level. On the leaves, disease symptoms appear as elliptical, oval, or irregular greyish-green spots. The entire plant will result in death if the disease extends to the inner sheaths. Plants affected in the early stage result in poorly filled grain. Closer plantation, with 96% humidity, high nitrogen, and a temperature of 30-32 degrees, are the most favorable conditions.

Bacterial Diseases

Bacterial Leaf Blight (Xanthomonas oryzae):

In this disease, symptoms appear 1-2 weeks after transplantation, including Leaf Lesions, which are visible in seedlings. The disease can occur in earlier stages as well, but is mainly noticed during the heading stage. The increase in disease leads to the death of seedlings. Severe wind, application of heavy nitrogen during late top dressing, heavy rains, and temperatures between 25 °C and 30 °C are the most favorable conditions.

Bacterial Leaf Streak (Xanthomonas oryzae):

This disease manifests as narrow streaks of various lengths, characterized by a dark, greenish color. The lesions enlarge and affect more prominent veins, which turn from yellow-orange to brown. The disease spreads from one planting season to another through infected stubble and seeds. Severe wind, heavy rain, temperatures between 25-30 degrees, and deep irrigation water are the favorable conditions.

CHAPTER 2

Review on an Intelligent System for Quality Evaluation of Dry Fruits Using Hyperspectral Imaging

Shweta Bhelonde[1,*] and **Manoj B. Chandak**[1]

[1] *Department of Computer Science and Engineering, Ramdeo Baba University, Nagpur, Maharashtra, India*

Abstract: In the rapidly evolving field of agricultural technology, the need for advanced, reliable methods for assessing dry fruit quality has become paramount. This paper presents an empirical review of the rapidly increasing application of Machine Learning (ML) and Deep Learning (DL) techniques in hyperspectral imaging for the qualitative analysis of dry fruits and their subtypes. Traditional methodologies, while having laid a robust foundation, often grapple with constraints such as limited precision, suboptimal accuracy, and scalability challenges. Furthermore, these conventional approaches typically exhibit significant delays in processing and struggle with complexity in handling diverse dry fruit categories. Addressing these limitations, this review comprehensively evaluates ML and DL methodologies tailored explicitly for hyperspectral imaging applications in dry fruit quality analysis. The proposed review process stands out by precisely comparing these methods across a spectrum of critical evaluation metrics, such as precision, accuracy, recall, processing delay, complexity, and scalability. This approach not only bridges the gaps identified in existing literature but also lays the groundwork for a more nuanced understanding of the capabilities and limits of these advanced technologies in practical scenarios. This, in turn, has improved quality control measures and enhanced overall productivity levels. In conclusion, this work extends the existing body of knowledge. It sets a new benchmark for applying deep learning techniques in agricultural hyperspectral imaging, thus marking a significant stride forward in pursuing agricultural innovation and excellence.

Keywords: Deep learning, Dry fruit quality, Hyperspectral imaging, Machine learning, Scenarios.

* **Corresponding author Shweta Bhelonde:** Department of Computer Science and Engineering, Ramdeo Baba University, Nagpur, Maharashtra, India; E-mail: shwetabhelonde@gmail.com

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

INTRODUCTION

The introduction of advanced computational techniques, particularly Machine Learning (ML), Deep Learning (DL), and Generative AI [1 - 3] into the domain of agricultural technology has foreshadowed a new era in dry fruit quality assessment. This paper delves into the intersection of these cutting-edge technologies with hyperspectral imaging, a method gaining traction due to its non-destructive nature and efficacy in the quality evaluation of dry fruits. The essence of this study lies in its meticulous review and analysis of existing ML and DL methodologies applied in hyperspectral imaging for dry fruit quality assessment, highlighting their capabilities and demystifying their limitations.

Historically, dry fruit quality assessment has relied heavily on manual inspection and conventional imaging techniques. While fundamental in their contributions, these methods [4 - 6] often falter in precision, scalability, and efficiency, mainly when dealing with large-scale agricultural operations. In contrast, hyperspectral imaging emerges as a powerful alternative, offering detailed insights into the internal quality of dry fruits without physical intrusion. Integrating ML and DL with hyperspectral imaging has amplified its potential, enabling more nuanced, accurate, and rapid assessments.

However, this integration is not without challenges. The complexity inherent in ML and DL algorithms, the diverse nature of dry fruits, and their varying hyperspectral properties pose significant hurdles. Additionally, the vast array of available ML and DL models, each with distinct strengths and weaknesses, necessitates a thorough and comparative analysis to determine the most effective approaches for specific dry fruit quality assessments.

Early dry fruit quality assessment studies predominantly focused on traditional methods, manual inspection, and basic imaging techniques. Though effective to a certain extent, these methods were limited by their subjectivity, lack of scalability, and inability to detect internal defects. The advent of hyperspectral imaging marked a significant leap forward. It provides a non-invasive way to assess dry fruit quality, leveraging the temperature variations and hyperspectral properties of dry fruits. However, interpreting hyperspectral images was initially challenging, often requiring extensive expertise and subjectivity in analysis.

The integration of ML into this domain [10 - 12] marked another pivotal moment. Early research in this area demonstrated that ML algorithms could effectively classify dry fruits based on quality by analyzing patterns in hyperspectral images. These studies laid the groundwork for more sophisticated applications, highlighting the potential of ML in automating and enhancing the accuracy of dry fruit quality assessment.

Subsequent research delved into applying DL, particularly Convolutional Neural Networks (CNNs), in interpreting hyperspectral images. This strand of research revolutionized the field, offering unprecedented accuracy in detecting external and internal defects in dry fruits. The ability of DL models to learn from vast amounts of data and identify intricate patterns undetectable to the human eye was a game-changer.

However, the literature also reveals specific gaps and challenges. One recurring theme is the complexity and resource intensity of DL models, which pose challenges in computational requirements and practical applicability in resource-limited settings. Another area highlighted in the literature is the need for more extensive datasets encompassing a wider variety of dry fruits and conditions to enhance the robustness and generalizability of these models. By systematically evaluating various ML and DL models in the context of hyperspectral imaging for dry fruit quality assessment, this paper aims to provide a clear, structured understanding of the current landscape while identifying areas ripe for future research and development processes.

MOTIVATION AND CONTRIBUTIONS

The Motivation and Contribution section of this paper centers on a central thesis: the integration of Machine Learning (ML), Deep Learning (DL), and Generative AI with hyperspectral imaging in agricultural technology is not merely a progressive step but a necessary evolution in the pursuit of optimal dry fruit quality assessments. This conviction is rooted in a confluence of factors, primarily the escalating demand for high-quality agricultural products and the relentless pursuit of technological innovation in the agriculture scenarios.

For instance, from Fig. (**1**), hyperspectral imaging enables the distinction between good-quality figs and bruised figs by capturing subtle spectral differences that conventional methods often overlook. Good-quality figs display consistent spectral profiles, indicating uniformity in their texture and chemical composition. In contrast, bruised figs exhibit disrupted spectral patterns due to internal damage, moisture variation, or discoloration, manifesting as identifiable anomalies in hyperspectral data. When coupled with advanced ML and DL techniques, these insights allow for highly accurate, non-invasive quality assessments, ensuring that only the best produce reaches consumers.

The motivation for this study emanates from several critical observations. First, the burgeoning global population and the accompanying food demand underscore the need for efficient, scalable, and accurate methods for assessing agriculture quality. Traditional methods, though foundational, are increasingly inadequate in meeting these demands, especially in the context of quality assessment of dry

CHAPTER 3

Deep Learning Fusion: Potato Leaf Disease Classification through Generative Diffusion Model

Virendra Kumar Shrivastava[1,*], **Prakruthi Ganiga**[1], **Rathnakar Achary**[1], **Chetan J. Shelke**[1], **A. Ezil Sam Leni**[1] and **Ramalakshmi K.**[1]

[1] *Centre of Excellence in Computer Vision, Department of Computer Science and Engineering, School of Advanced Computing, Alliance University, Bangalore, Karnataka-562106, India*

Abstract: Potatoes are one of the most highly consumed foods in the world. Due to several environmental and climate changes, potato leaf diseases are emerging frequently, causing severe damage to potato plants and drastically reducing potato crop yields, which results in financial losses to farmers. This research presents a deep learning generative approach to enhance the accuracy of identifying classical diseases in potato leaves using a novel data augmentation method, specifically the diffusion model. By employing diffusion models, researchers created images resembling real ones showing diseases such as Early Blight, Late Blight, and healthy leaves on potato leaves. These generated images are then combined with others to strengthen the training of a deep learning model, enabling it to better distinguish between early blight, late blight, and healthy leaves. This combination of synthetic data enhances the model's robustness. Demonstrates promising advancements in analyzing potato leaf disease. The significance of this study lies in addressing challenges related to data availability while improving the model's ability to generalize across the stages of the disease. The outcomes have implications for agriculture by providing effective disease-identification methods for potato farmers. The proposed CNN model delivered classification accuracy (98.99%) with ResNet18 integrating with the diffusion model on the potato leaf disease image dataset. Integrating diffusion models represents a step towards optimizing crop management and maximizing yields in potato cultivation.

Keywords: Artificial neural network, Credibility limits, Loan approval, Loan prediction, Machine learning.

INTRODUCTION

The potato, scientifically known as *Solanum tuberosum*, is a crop that serves as a food source for billions of people worldwide. In the industry, it is challenging to safeguard potato plants from diseases that can significantly impact their yield and

[*] **Corresponding Author Virendra Kumar Shrivastava:** Centre of Excellence in Computer Vision, Department of Computer Science and Engineering, School of Advanced Computing, Alliance University, Bangalore, Karnataka-562106, India; E-mail: Virendra.shrivastava@alliance.edu.in

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

the quality of the potatoes grown. Identifying and classifying diseases in potato leaves at an early stage is crucial for implementing effective disease management strategies. Recent technological advancements in learning techniques have resulted in automating and improving the accuracy of disease classification in plants. Plant diseases, including those that affect potato leaves, pose a threat to food security. Early blight disease (caused by *Alternaria solani*) and late blight disease in potatoes (caused by Phytophthora infections) are known culprits that negatively affect the health of potato crops [1, 2]. In agriculture, the process of identifying diseases traditionally involves expert inspection. However, this approach is both time-consuming and prone to error. It has become crucial to incorporate technologies to overcome these challenges.

Machine learning is a popular technique for identifying and classifying diseases [3 - 6]. Deep learning, which falls under the umbrella of intelligence, has succeeded in various tasks related to image classification [7 - 10]. One type of learning architecture, called Convolutional Neural Networks (CNNs), has demonstrated the ability to recognize patterns and characteristics within images [11 - 14]. This makes CNNs highly suitable for classifying plant diseases based on leaf images [15, 16].

However, the efficacy of the learning model is primarily dependent on the availability of diverse datasets for training. A significant obstacle in potato leaf disease classification is the scarcity of labelled datasets encompassing different illnesses. Creating and analyzing a well-labeled collection of potato leaf diseases is a challenging and time-consuming task. The limitation [17] dramatically impedes training robust and widely applicable deep learning models. Scientists have explored various data augmentation strategies to address this issue, which involve artificially modifying the training samples to generate a larger dataset. Standard techniques in picture manipulation include flipping, scaling, and rotation. Although these tactics can enhance the generalization capabilities of models to some degree, they may not be sufficient to accurately reflect the complex variations that exist in real-world scenarios [18]. Generative models for data augmentation have made substantial progress thanks to deep learning. Artificial models have been used to create novel products that mimic the functions of living organisms. The goal is to enhance the ease of training deep learning models.

A Generative Adversarial Network is a generative model commonly used to enhance data. It consists of two networks: the generator and discriminator networks. Both components undergo simultaneous training [19, 20]. While the generator creates synthetic data, the discriminator distinguishes between actual samples and fake ones. Synthetic data produced by this adversarial training method is of exceptionally high quality and closely reflects real-world examples [21, 22]. While GANs have shown promise in style transfer and image generation, they may struggle to maintain the authenticity and diversity of the produced data. Synthetic visuals must be accurate and applicable for the model training to be efficient in plant disease classification [23].

In this specific case, generative diffusion models are particularly effective. Drawing inspiration from the physical processes of diffusion, generative diffusion models offer a novel approach to enhancing data. To simulate a diffusion-like effect, the input pictures are gradually exposed to controlled noise during the diffusion process. Deliberately adding noise to synthetic samples makes them appear more realistic, creating several copies of the original images [24 - 26]. Building upon the current foundation of generative diffusion models, this study aims to address the unique challenges of disease classification in potato leaves. The deep learning model must be able to distinguish between Late blight and Early blight, for example, solely by examining the symptoms. This study aims to use generative diffusion models to improve the model's resilience and accuracy in illness classification. Given the dearth of data regarding the classification of diseases affecting potato leaves, our approach carries significant implications for this domain. Generative diffusion models are implemented to augment crop management, increase crop yield, and ensure the long-term viability of the potato farming ecosystem.

RELATED WORK

This study introduces Gaussian noise of varying intensities into training images, enhancing data classification. This approach improves the model's generalizability in classifying disease in the images to encompass robust feature learning. Including chaotic inputs during the training enhances the accuracy and robustness of the model. The diffusion model is indispensable for enhancing accuracy through effective data augmentation and robust feature capture.

Researchers have previously employed datasets, such as PlantVillage and Mendely, to train models and investigate the effects of different potential diseases on potato plants, particularly potato leaf disease. Data augmentation techniques, *viz* a *viz* horizontal rotation, shear, and zoom, were applied on potato leaves to meet the variability in the dataset. K-means segmentation achieved an accuracy of 97% for K=3 [27]. Shin D., Kim T., and Min B. employed a data augmentation model for image-to-image translation. The PlantVillage dataset was used to train the Cycle GAN, which attained an accuracy of 97.74%, after using various deep learning approaches [28]. Among these, the ResNet18 convolutional network demonstrates the highest level of effectiveness. Barman U., Sahu D., and Das J.

CHAPTER 4

Fuzzification for Precision Farming with Minimal Human Intervention

Sudheer Mangalampalli[1,*], **Ganesh Reddy Karri**[2], **Pelluru Pavan Kumar Reddy**[2], **Ramith Yadav S.**[2], **K. Varada Rajkumar**[3] and **Kiran Sree Pokkuluri**[4]

[1] *Department of CSE, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India*

[2] *SCOPE, VIT-AP University, Andhra Pradesh, India*

[3] *Department of CSE (AI&ML), MLR Institute of Technology, Telangana, India*

[4] *Department of CSE, Shri Vishnu Engineering College for Women, Bhimavaram, India*

Abstract: Seamless communication between humans and things like cars and other such machines in important industry sectors is constantly evolving as each day passes. In the 21st century, crucial advancements in several fields have implemented IoT in them because of its numerous merits. When IoT is put into practice alongside Cloud Computing and Data Analytics, it can reduce the gap and enhance cooperation between the physical and digital worlds, which is essential for sustainability in this hyperconnected era. The project involves collecting data through sensors in the field, which is then sent to the IoT analytics and cloud-based platform ThingSpeak. This platform enables data ingestion and storage, allowing farmers to remotely visualize and analyze real-time data in the form of data streams and take action accordingly. Fuzzy logic is one such approach to computing that works on degrees of truth rather than single-valued Boolean logic, thus yielding better and more accurate conclusions for an imprecise spectrum of data. Fuzzy logic is implemented in this project to control water pumping time based on user-defined thresholds. With an increasing population and depleting resources, precision farming has proven to be efficient as it focuses on minimizing waste and is cost- and energy-efficient. The objective of this paper is to develop a system that assists farmers in remotely monitoring data, optimizing resource utilization, and increasing productivity.

Keywords: Data ingestion, Fuzzy logic, IoT, Precision farming, Thing speak.

INTRODUCTION

Today's era has witnessed rapid urbanization and industrialization thanks to ongoing technological evolution and continuous advancements, leading to notable

* **Corresponding author Sudheer Mangalampalli:** Department of CSE, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India, E-mail: ms.sudheer@manipal.edu

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

emerging technologies. India has made progress in practically every industry, but has fallen behind in agriculture, as it still relies on antiquated conventional practices. Agriculture is crucial to a country's development, as it is a significant contributor to the country's GDP and ensures food security for a massive population. Implementing new technologies has the potential to revolutionize farming practices. In traditional agriculture, farmers rely on manual labor for crop and livestock management, often resulting in ineffective resource utilization. Precision farming offers hope for overcoming agricultural challenges. It uses GPS, sensors, and prominent computing technologies like the cloud, IoT, and Fuzzy Logic to monitor crop health remotely and provide precise nutrients and water to reduce waste and increase yields. The Internet of Things (IoT) integrates tangible items into the Internet Protocol (IP) framework, enabling them to connect to the Internet and communicate with other systems and devices. These interconnected devices make intelligent decisions based on gathered data. Domains such as neural networks, AI, and ML play a significant role. The output is automated based on training, and real-time data can be visualized on the cloud from anywhere. The proposed precision farming system is designed to collect data from sensors and transmit it to an Arduino, which in turn sends this large spectrum of data to an IoT-based Cloud platform, ThingSpeak, where the data can be viewed in the form of graphs. This platform comes with integrated MATLAB Analysis and Visualization Apps, where we use the Fuzzy Logic Toolbox to regulate a controlled machine output according to multiple inputs.

LITERATURE REVIEW

Various references have been considered and reviewed to identify existing systems and determine their applications, architectures, and flaws. Xu *et al.* review the introduction and implementation of IoT in agriculture, reporting existing problems and future trends [1]. It focuses on the evolution of communication technologies in agricultural IoT, information processing, and environment monitoring, paving the way for sustainable ideas. Chapter 9 of the book titled "Sustainable Digital Technologies for Smart Cities" describes the role of intelligent predictive analytics and big data in farming, comparing Fuzzy Logic to binary logic and justifying this with performance metrics [2]. A study focuses on using Fuzzy Logic for a controlled water pumping mechanism, but could have used additional sensors, like PH, air quality, and light intensity, to achieve more efficient results [3]. Kagalkar addressed in his study the issue of excess water utilization, but proposes manual handling of the water pump [4]. Other studies suggest frameworks for water pumping based solely on soil moisture, which may yield imprecise outcomes under fluctuating humidity and temperature [5, 6]. Astudy discusses the early implementation of IoT in agriculture, with a focus on data collection and cloud visualization [7]. Pradana *et al.* proposed an irrigation control system using the Tsukamoto Fuzzy Algorithm, which is less accurate than the Mamdani FIS due to a high MAPE value [8, 9]. Katthineni and Reddy presented an architecture for motor control *via* relay using IoT, Cloud, and Machine Learning; however, it relies on predicting using datasets in arrays, which can become complex with large data spectra [10]. Kaswar *et al.* specify a hydroponic system for adaptive control of nutrition, but the models could be more interpretable [11]. A study by Abioye *et al.* outlines a monitoring system and data-driven modeling using drip irrigation, which is not cost-effective and can lead to clogging issues [12]. Benyezza *et al*. propose an innovative greenhouse control platform with a Human-Machine Interface on a Node-RED server for data visualization, but it faces complexity issues in debugging and lacks architectural benefits [13]. Pierre *et al.* propose an irrigation system with low energy consumption [14]. It suggests the inclusion of solar panels, which satisfies the aim, and also includes an XBee module as a communication technology. Xbee, although it uses low power to operate and has a flexible network topology, its transmission distance ranges only from 10 to 100 m, which is not an efficient approach considering the development of Wireless WAN, which has a very long transmission distance.

PROPOSED ARCHITECTURE

Control System

The project comprises functional parts, including a motor/water pump, a GSM/WiFi module, an LCD, a solar panel, and various sensors, such as soil moisture, air humidity, air quality, temperature, and light intensity sensors, all connected to an Arduino Uno, an open-source electronics prototyping platform. This microcontroller can receive input from sensors and initiate actions, like water pumping. The GSM/WiFi module, which uses a SIM card for connection *via* an RS-232 interface, sends and receives messages through GPRS using AT (Attention) Commands. The LCD displays sensor values and is placed near the microcontroller. Solar panels power the project, utilizing renewable energy for long-term efficiency and low operational costs. The soil moisture sensor measures soil water content in percentage volume. The DHT11 sensor records environmental temperature and humidity. The Light Dependent Resistor (LDR) sensor measures light levels around the crop, with resistance decreasing as light intensity increases. The MQ2 air quality sensor evaluates air quality based on chemo-resistive conductivity, identifying gases, like methane and carbon dioxide, and measuring them in ppm (parts per million), with higher ppm values indicating poorer air quality.

CHAPTER 5

Enhancing Feedforward Neural Network Optimizer Variants to Strengthen Aquaponics Fish Growth Predictions

Safiya Begam[1], **W. Aisha Banu**[1,*] and **Sharon Priya**[1]

[1]*B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India*

Abstract: For training a deep learning model, one has to get the parameters necessary to satisfy an objective function. In general, the objective is to minimize the loss encountered throughout the learning process. A type of supervised learning involves feeding data samples and their matching outputs to a model. A framework works to get the generated output nearer to the desired result by comparing its results to the planned outcome and taking into account any variations. Optimization algorithms have been utilized to do this. Multiple cycles of optimization are carried out till the process is completed in order to increase the accuracy of the model. Numerous optimization techniques have been developed to address the challenges associated with the learning process. The techniques addressed involve Adadelta, Adagrad, Adam, and stochastic gradient descent. With Adam and Adagrad, Pond1 has a desirable training result of 0.96 at epoch 200, and Pond2 has an optimal training result of 0.96 at epoch 250. For Ponds 1 and 2, the best Adam test results are 0.9936 and 0.9975, respectively. Except for AdaGrad, which yielded the lowest scores, the results show that Adam-based optimizers gave the best results. Depending on the surroundings, an aquaponics system with a well-optimized FNN may help make more accurate predictions of fish development.

Keywords: Aquaponics, Feedforward neural network, Optimizer variants, Predictive analytics.

INTRODUCTION

Combining hydroponics and aquaculture, aquaponics is a sustainable farming method that offers enhanced resource efficiency, variation, and less negative environmental effects. Most beneficial are urban regions with little arable land and water-stressed places. Aquaponics promotes local food production, reduces reliance on artificial fertilizers and pesticides, and enables year-round production

* **Corresponding author W. Aisha Banu:** B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India; E-mail: aisha@crescent.education

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

of fresh, healthy food. It also reduces the carbon footprint associated with long-distance travel and offers an opportunity for experiential learning [1]. An irrigation system, known as an aquaponics system, combines fish and plants. In this cyclical structure, nitrifying bacteria break down fish excrement, which is then eventually absorbed by the plants as food. Because these aquaponics systems can be installed practically anywhere, they are ideal for urban farming. The study "Food, Fish Growth, and Fish Yield Relationships" looks into the relationships between pond fish productivity, fish development, and food availability. It examines dietary choices, feed mix, and nutritional characteristics, and how these factors affect fish development. Research is essential for managing fisheries and aquaculture [2].

Getting concerned about the water, such systems require a significant initial investment, even though they should consume less energy, water, and fertilizers than conventional agricultural systems. This issue provides important insights for researchers studying fish physiology and aquaculture by examining the endocrinological aspects of fish growth and reproduction. It focuses on the mechanisms of action and effects of different hormones on physiological processes [3]. Knowledge of fish growth is essential for maintaining a healthy aquaponics environment, as overfishing and the biological and atmospheric content of the substrate—which is influenced by pollution among other things—both negatively affect fish growth. Fish populations need to be regularly checked for the aquaponics system to be successful. Diversifying the systems is equally important as raising productivity and/or species diversity to support this expansion. In aquaponics, the total efficiency of the system depends on the fish's ability to develop normally. The entire productivity and sustainability of the setup are directly affected by it. Healthy aquatic environments ensure nutrient-rich waste, which is essential for plant growth, and stable fish growth is a sign of this. This mutually beneficial relationship, in which fish provide vital nutrients to plants and gain from a hygienic and cleansed drinking environment, is the foundation of aquaponics management. Ensuring the health of the fish population requires monitoring and controlling parameters related to fish development, such as temperature, water quality, and feeding. A sustainable supply of protein for human consumption is also provided by fish that grow efficiently, which benefits the system's nutrition cycle [4]. The goal of sustainable development is to strike a balance between human and ecological well-being.

An ecosystem approach to the fishing industry has recently gained recognition as a significant area of study. The article's structure is best described as follows: Section II covers the pertinent research on optimization approaches. The datasets and optimization techniques used in this work are covered in Section III. The algorithms Adam, Adagrad**,** Adadelta, and stochastic gradient descent are covered. The IoT-based Aquaponics datasets for Ponds 1 and 2 are the ones being discussed. In Section IV, the results of training the model with the available optimization techniques on pertinent datasets are compared and elaborated upon. In Section V, the paper is summarized, and future directions are discussed.

LITERATURE REVIEW

The taxonomy presented in the article facilitates a clear understanding of the environment by categorizing different approaches within the field of deep learning. However, the work needs to be improved because it does not adequately address the difficulties in implementing deep learning technologies, new approaches, and their interactions with reinforcement learning, as well as their ethical and societal ramifications.

An extensive overview of deep learning methods, taxonomy, applications, and future possibilities is given in Sarker's 2021 paper. It categorizes approaches and highlights their importance. Future directions and current research trends are also covered in the paper [5].

In 2021, Naseri conducted a study to examine the impact of varying parameters on both traditional and deep neural networks. The study highlights the importance of comprehending how these networks react to modifications in activation functions, learning rates, and network topology for effective task design and fine-tuning [6].

In order to enhance training and convergence, this paper provides a revised version of the Adam optimizer for deep neural networks (DNNs). The enhancement addresses flaws in the first methodology, potentially enhancing the efficacy and efficiency of deep neural network training. Researchers and practitioners interested in machine learning and neural network optimization will find this study to be of interest [7].

This effort aims to study optimizers in scientific computing and deep learning. It discusses several optimization algorithms, including SGD, Adam, and RMSprop, outlining their principles, advantages, and potential applications. Effective neural network training requires an understanding of optimizers, making this study valuable for both practitioners and scholars [8].

The study describes an approach for selecting suitable deep learning algorithms based on specific problem settings or datasets. It utilizes machine learning techniques to analyze the features of the dataset and matches them with the relevant algorithms. Researchers and practitioners interested in enhancing the

CHAPTER 6

Predicting Carbon Dioxide Emissions from Green Waste Composting and Identifying Key Factors Using Machine Learning Algorithms

M. Vaishali[1,*], M. Monika[2], M. Atish[3] and **S. Rajarshi[4]**

[1] *Department of Computer Science and Engineering, Maharishi Markandeshwar Deemed to be University, Ambala, India*

[2] *Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Navi Mumbai, India*

[3] *Department of Computer Science and Engineering, BMS Institute of Technology and Management, Bangalore, India*

[4] *School of Advanced Sciences, VIT-AP University, Andhra Pradesh, India*

Abstract: Today, one of the main concerns that the world faces is climate destruction. The hazardous gases emitted by material waste are a critical component among many that contribute to the constantly changing atmospheric conditions. Waste from agriculture is also included in this group. Controlling the production of these gases largely depends on accurately predicting the amount to be produced. The current study offers new classifiers and prediction models based on machine learning to forecast the generation of hazardous gases from agricultural waste. The Random Forest method achieved 100% classification accuracy after feature optimization and selection were applied to the gathered dataset. Logistic regression, AdaBoost, and Decision Tree ranked second, third, and fourth, respectively, with classification accuracies of 97.19%, 97.18%, and 97.14%. SVM provided the highest prediction accuracy, 98%, in regression models.

Keywords: Artificial intelligence, Environment, Green energy, Machine Learning, Prediction, Renewable energy.

INTRODUCTION

In recent decades, nations worldwide have faced significant challenges in addressing climate change, a formidable threat leading to global warming and other widespread environmental impacts. As a result, numerous academics, ecological scientists, and governmental organizations have been recommending

* **Corresponding author M. Vaishali:** Department of Computer Science and Engineering, Maharishi Markandeshwar Deemed to be University, Ambala, India; E-mail: drvaishaliwadhwa@gmail.com

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

various courses of action to overcome these difficulties. For example, the government is actively attempting to raise public awareness of the seriousness of environmental issues and defacement [1]. In addition, scientists and researchers are working to implement solutions that will reduce the production of harmful gases. To mitigate environmental harm, efforts are also being made to convert these hazardous gases into beneficial energy. Along similar lines, some scientists have suggested transforming greenhouse gases into energy that may be used, such as electricity [2]. Approximately three-fourths of all greenhouse gas emissions originate from the urban sector, which is the primary source of these emissions. Furthermore, this percentage may increase due to the growing urban population, necessitating effective corrective action. While there is no way to prevent or manage agricultural waste, steps may be taken to maximize its use, according to some experts and environmentalists working in this area [3, 4]. For example, agricultural waste can be converted into biogas, which can then be used to generate electricity. Typically, agrarian waste releases CO_2, CH_4, N_2O, and SO_2 gases. These gases are then converted by microorganisms into biogas within biogas plants *via* chemical reactions. The resulting biogas is a renewable soure of energy that can help produce electricity.

Furthermore, it has been noted that, of all the gases, the emission of CO_2 is the highest; this gas is a key contributor to environmental harm. Therefore, it is essential to forecast CO_2 emissions from agri-waste since they significantly impact the prediction of biogas and environmental health [5]. It is important to note that recycling the material helps protect the environment and improves the economy. Therefore, considering the various advantages of generating biogas from agricultural waste, it is essential to capitalize on the technological advancements in this domain. The authors of this work have made an effort to utilize Machine Learning (ML) and Artificial Intelligence (AI) models in this field, building on previous work in this direction [6]. ML models are a practical option in this field because of their enormous potential, which has been demonstrated over many years in information forecasting and categorization. The current efforts aim to evaluate the predictive accuracy of various machine learning models for CO_2 emissions and biogas production. The authors have employed diverse categorization and regression methods on datasets encompassing 192 nations over ten years. Several ML algorithms, such as Random Forest, KNN, Naïve Bayes, Decision Tree, Logistic Regression, SVM, Linear Regression, and ADA Boost, have been used in the experiments. A comparison analysis is then conducted to identify the most effective ML model for this particular case. This research aims to estimate the predictor class using the most widely used regression models and machine learning classifiers, after gathering a dataset with sufficient labeled variables. The study is motivated by the fact that most existing methods, which have been utilized in other studies, have failed to achieve high accuracy levels for the same class of situations [1, 5]. Furthermore, existing investigations do not utilize EDA to gain insight into the problem's parameters and their impact on the obtained outcomes.

The current work is divided into many components. In this case, section 2 outlines the previous relevant research projects completed by various researchers. The implementation process is expounded upon in Section 3. Results are covered in Section 4, and Section 5 concludes with a presentation of future work.

LITERATURE SURVEY

A study suggests that CO_2 is a crucial component produced during the composting of green waste. Its authors researched how to predict CO2 during this process. A comparative analysis of six different methods was conducted. Before applying machine learning (ML) models, preprocessing was performed, which included the removal of outliers. The results of the comparative study show that the random forest algorithm outperforms other models in terms of accuracy, achieving an accuracy of 88% [1].

Additionally, the authors of another study propose using renewable energy sources since they drastically reduce waste and safeguard the environment [2]. For prediction purposes, their study proposes a hybrid model that combines Random Forest with Long Short-Term Memory (LSTM) [10]. The accuracy of this hybrid model is 20% higher than its conventional analytical models. Another study also establishes the use of ML models in the present situation. The authors conducted research with two primary aims: first, to develop a reliable method for estimating waste production; and second, to compare various quantification techniques. To achieve these objectives, they employed four specific models: artificial neural networks, decision trees, grey models, and multiple linear regression [4]. The authors of another study employed a deep neural network to estimate the quantities of SOx and NOx released during the conversion of coal to energy, following a similar approach [5].

Numerous researchers have addressed this problem and proposed several solutions in light of the pervasive and profound effects that toxic gases have on the environment [7 - 9]. A study has proven that an effective model for decision-making is necessary to improve biogas production [3]. Given the input waste values, a graphical user interface is needed to forecast the amount of biogas produced. To achieve this, the authors used various machine learning models, including logistic regression, random forest, support vector machines, gradient boosting, and kNN regression. With an accuracy of 87% for test data, kNN was found to have the best accuracy for evaluating the model efficiency at the Hainan biogas facility.

CHAPTER 7

Can Ethiopian Cooperative Coffee Farmers Predict Financial Management Decisions Through Human Capital and AI Mediation?

Mershaye Birhane[1], **Shashi kant**[1,*], **Brehanu Borji**[2] and **Chalchisa Amentie**[3]

[1] *Department of Management, Bule Hora University, Bule Hora, Ethiopia*

[2] *Department of Management, Hawasa University, Hawasa, Ethiopia*

[3] *Department of Management, Ethiopian Civil Service University, Addis Ababa, Ethiopia*

Abstract: The central theme of the chapter was to demonstrate how Ethiopian cooperative coffee farmers can make informed financial management decisions through the mediation of human capital and AI. The investigators employed a cross-sectional examination using the multistage method with the test group. An Exploratory Factor Analysis (EFA) was conducted to verify whether the data were sufficient. Eigenvalues were utilized to explore the principal proxies in the dataset. A Confirmatory factor analysis was used to assess whether the manifested determinants could appropriately describe the validity of the construct. SEM was used to conduct mediation analysis and assess model fitness. According to the results, the investigation revealed a substantial association, as the p-value was below the algebraic threshold for data adequacy, and the sphericity was 0.893. The chi-square was below 3.0, which the investigators found indicative of model adoption. By Tucker Lewis's fitness, the model was manifested as a fit. The predictive ability of farmers has increased by 75% when coffee cooperatives practice human capital mediation. Consequently, there is a fractional conflict between the sustainability of cooperatives and their management of financial methods, brought about by human capital and AI applications.

Keywords: AI, Coffee cooperatives, Financial management Strategies, Firm performance, Human capital.

INTRODUCTION

The organization's success is influenced by the growth of human capital and technological foundations, such as Artificial Intelligence (AI) [1]. Therefore, human capital development requires careful consideration of financing sources, technological advancements such as AI, and informed investment choices [2]. The

* **Corresponding author Shashi Kant:** Department of Management, Bule Hora University, Bule Hora, Ethiopia; E-mail: skant317@gmail.com

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

company will profit by improving its human capital and artificial intelligence (AI)-based management system and selecting the relevant asset classes to invest in and financing sources [3]. The corporation is expected to see an increase in profits over time from its investments in AI and human resources. Businesses with good investment potential can benefit from rising stock prices and promising prospects, increasing the firm's worth.

Processing businesses enhance their financial success through modern cost management strategies that leverage AI and human capital development [4]. In cooperative settings, adopting effective asset management *via* investments in both human capital and Artificial Intelligence (AI) is expected to boost firm sustainability and market share. This is achieved through improved capital structure predictability, which enables better decision-making, facilitates scaling, and enhances customer satisfaction [5].

Internal controls of cooperatives ensure that companies manage their capital structure in compliance with rules and regulations [6]. Robust internal controls have the potential to enhance operational performance and provide precise financial reporting in the event of internal or external audits [7]. Costing, resource management, inventory agility, and investment in AI and human capital are all aspects of financial management procedures [8]. Community-based businesses are carefully designed to enhance the usefulness, value, and desirability of specific offerings through investments in both human and artificial intelligence [9].

THEORETIC ESTABLISHMENT

Workforce quality is a key component of the broader concept of AI and human capital, encompassing various other aspects [10]. Because they are seen as investors in the company, staff are rewarded with social wealth and the expected return on their assets [9, 10]. Based on a study, three main viewpoints have been presented regarding human capital [11]. According to the principal perspective, capital in the form of humans is the outcome of sharing; as such, its value is assigned to improving physical capabilities, innate intellect, and obtaining knowledge and skills [12, 13].

The Resource-Based View (RBV) theory relates to investing. The subsequent concept of partial output takes human capital into account. It is defined as the specialized AI knowledge, AI knowledge base competence, operational experience, and the expertise of managers, technicians, and inventors [14]. Notably, the perspective on overall output characterizes human capital as the culmination of an individual's physical prowess, knowledge, intellect, and other abilities used in production [15]. The last point of view, which holds that a person's productive potential is derived from their human capital, has gained significant support. It is thought that technical and managerial personnel are not the only ones with access to human capital [16].

EMPIRICAL LITERATURE REVIEW

Fiscal Supervision and Business Success

Every institute needs its management to be prepared to take risks and explore new possibilities to succeed. Providing the required human effort, which is rewarded, entails risk [15]. In contemporary financial management, theoretical and statistical techniques are applied [16]. Employees and other interested parties are invited to provide their comments on prospective subjects of interest during the early phases. The viewpoints inform certain judgments by reviewing all the replies [17].

Financial management is an approach by which an institution uses a range of logical skills to scrutinize historical, current, and future financial results to make informed decisions about its overall progress [18]. The expansion of the economy depends on the business climate [19]. It is precarious to recognize that an economy's future depends entirely on how well its institutions are managed. The financial management plan should be prioritized before deciding on the expected returns [20]. It is divided into short and strategic fiscal prospects [19]. But eventually, both of them reach a point when capital returns surpass expectations [21]. These fundamentals are combined to depict the nature of practices used to manage finances [22]. Including investment in humans, along with balanced investment decisions, leads to optimal conditions [23].

H1: Fiscal Supervision has a statistically noteworthy association with business success.

Management of Finance and Investment in Humans

Human capital refers to an employee's skill and ability value in terms of money [24]. Employers value loyalty, punctuality, skills, intellect, abilities, training, and good physical and mental health [23]. Finance distributes capital to meet corporation objectives by managing spending and revenue. HR recruits, encourages, and inspires employees to attain similar objectives. This department is typically the most expensive part of an organization's HR [25].

In more modern times, the prevailing philosophy, shaped by diverse ideologists, was considered by Becker and Schultz in the 1960s to emphasize the value of human talents [24, 25]. Like any other form of capital, Schultz claims that human capital can be employed to enhance both the quantity and quality of production. This would require investing funds in training, employee development, and

Section 2
Intelligent AI & ML Solutions

CHAPTER 8

Smart Cities: Urban Planning and Infrastructure Management

Aanchal Chaudhary[1], **Harsh Kumar Pandey**[1], **Ayush Yadav**[1], **Vaibhav Kumar Singh**[1], **Nishant Singh**[1] and **Hitesh Mohapatra**[1,*]

[1] *School of Computer Engineering, KIIT (Deemed to Be) University, Bhubaneswar-751024, Odisha, India*

Abstract: The Internet of Things (IoT) can access certain subsets of data for the creation of several digital services, while also integrating a wide range of diverse and heterogeneous end systems in a transparent and seamless manner. This chapter examines innovative solutions to key challenges in smart cities, with an emphasis on ambient noise assessment, air quality monitoring (E-nose system), and smart parking management systems (SPMS). Utilizing servo motors, smartphone apps, and sensors, the SPMS combines Internet of Things technology to maximize parking space distribution, improve user experience, and raise total parking efficiency. An E-nose air quality monitoring system with a sensor array based on the ESP32 can identify gas pollutants, dust, and CO_2. Techniques for processing data, such as threshold alarms and median computation, provide precise and prompt air quality monitoring. Additionally, a noise monitoring system with a Gaussian filter and sound level sensors is demonstrated. The proposed analysis demonstrates how IoT-based solutions may be used to enhance sustainability, manage urban infrastructure more effectively, and improve people's quality of life in smart cities.

Keywords: Internet of Things, Internet of Things network, Monitoring, Smart cities, Urban areas.

INTRODUCTION

The significant development in population density within urban areas has necessitated the provision of substructures and services to fulfill the needs of inhabitants. Since it is now possible to connect devices and enable communication between them *via* the Internet, new smart city solutions are emerging to address these needs.

* **Corresponding author Hitesh Mohapatra:** School of Computer Engineering, KIIT (Deemed to Be) University, Bhubaneswar -751024, Odisha, India; hiteshmahapatra@gmail.com

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

There has been a notable increase in the number and quality of digital devices, including sensors, actuators, smartphones, and smart appliances, which supports the broad commercial objectives of the Internet of Things (IoT) [1]. Thus, the Internet of Things idea promises to make the Internet even more immersive and ubiquitous. Additionally, the Internet of Things (IoT) will facilitate the creation and development of several applications that utilize the potentially massive volume and diversity of data generated by extensive range of devices, such as household appliances, security cameras, monitoring sensors, actuators, vehicles, and many more by facilitating simple access and interaction and thus offering new services to individuals, businesses, and government agencies for public administration [2]. Fig. (**1**) shows some IoT devices that are used to make cities smart.

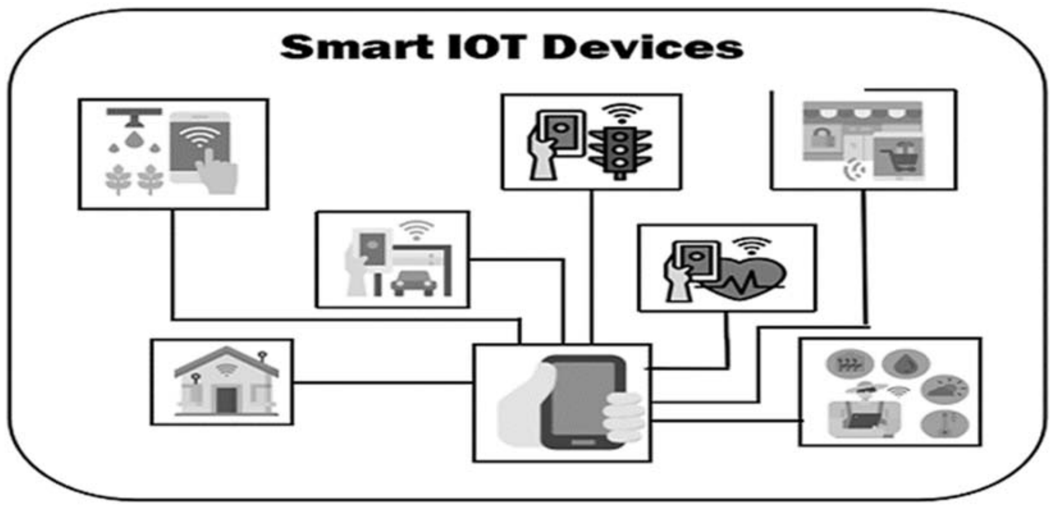

Fig. (1). Smart IoT devices.

It is a difficult task to find solutions that can meet the needs of every potential application scenario in such a diverse field of application. Numerous and, perhaps, incompatible proposals for the practical implementation of IoT systems have proliferated as a result of this difficulty [3]. Due to its novelty and complexity, the implementation of an Internet of Things network, along with the necessary backend network services and devices, still lacks an established best practice from a systems perspective. A well-defined and widely accepted business model that can draw investments to support the deployment of these technologies is lacking, which hinders the adoption of the IoT paradigm, in addition to its technical challenges [4].

Waste Management

Waste management poses a significant barrier for smart city applications, affecting both the environment and the health of our society. Humans and animals produce solid waste, which is usually discarded as useless [5]. At least 33% of the 2.01 billion tons of municipal solid waste produced worldwide each year is not managed in an environmentally responsible manner. By 2050, global waste is estimated to reach 3.40 billion tons, more than doubling population growth over that period [6].

Air Quality

The World Health Organization (WHO) has released a report on air quality that highlights the significant environmental threat that air pollution poses to human health. The research reported nearly 6 million premature deaths as a result of exposure to contaminated air sources [7]. Data from monitoring the busiest metropolitan regions demonstrate that pollution affects not only our own lives but also the lives of future generations. Consequently, it is critical that we make every effort to exercise caution and regulate the hazardous pollutants that we emit into the environment [8]. Installing pollution and air quality sensors throughout the city and making the sensor data accessible to the general public are necessary for the implementation of such a service [9].

Noise Monitoring

Chronic noncommunicable diseases are largely caused by classical risk factors like smoking, diabetes, and arterial hypertension; however, more recent research has indicated that environmental variables also play a role in the development of these diseases. In our industrialized society, environmental stressors, such as air pollution and noise, are becoming increasingly significant. In particular, traffic noise from automobile, airplane, and train travel may be a new risk factor for cardiovascular diseases, such as hypertension, hyperglycemia, hyperlipidemia, and increased blood viscosity and coagulation [10]. In order to measure the amount of noise generated in the locations with these services, an urban IoT can provide a noise monitoring service.

Smart Parking

India is the second most populous country in the world. The growth and development of Indian cities have resulted in a lack of appropriate parking space architecture. As individuals frequently travel for work, there is a high possibility that they may not be familiar with the neighborhood. As a result, they might not be aware of where to park [11]. Road sensors and intelligent displays, which show

Applied Artificial Intelligence in Data Science, Vol. 3, 2025, 126-141

CHAPTER 9

Next-Generation Urban Mobility Management: A Framework for Intelligent Traffic Control

Sujithra Thandapani[1,*], **Durai Selvaraj**[1], **Mohamed Iqbal M.**[3], **D. Jaganathan**[4] and **V.C. Bharathi**[5]

[1] *Department of CSE, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India*

[2] *Department of CSE, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, India*

[3] *Department of Information Technology, Sri Sivasubramaniya Nadar College of Engineering, Chennai, 603110, India*

[4] *Department of Artificial Intelligence, Madanapalle Institute of Technology & Science, Madanapalle, India*

[5] *School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India*

Abstract: Traffic congestion is exacerbated by the proliferation of vehicles, ineffective traffic management practices, and inadequate road infrastructure. Especially in cities, people are struggling to have quick mobility. Improper traffic management leads to longer transit times and contributes to fuel waste, air pollution, and increased accident risk. Conventional traffic control systems set the service time for each lane based on its nature. For example, fast lanes may get more service time than slow ones. It is a static system. It uses fixed timing to service all the lanes irrespective of traffic density. It results in increased traffic in a specific lane. In most countries, automated systems are unavailable to service emergency vehicles (ambulances, firefighting vehicles, mobile medical units, blood donation vehicles, and VIP vehicles) due to the enormous implementation and maintenance costs. In this chapter, a low-cost edge device is proposed for an intelligent traffic control system to address this issue. The proposed system focuses on dynamic signaling and automatic green enablers for emergency vehicles. Dynamic signaling controls the timing of the traffic control system based on the lane's traffic density, which is measured by the You Only Look Once (YOLO V3) algorithm. An automatic green light enabler enables the green light in advance for the lane where the emergency vehicle's presence is 750 meters away from the traffic signal. It deals with various types of emergency vehicles, such as ambulances, firefighting vehicles, and mobile medical units. Computer vision and communication technologies are utilized in designing a cost-effective and innovative traffic control system. It assists

[*] **Corresponding author Sujithra Thandapani:** Department of CSE, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India; E-mail: sujimecse@gmail.com

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

in efficiently managing traffic. It also services emergency vehicles at road intersections. Decreasing the average waiting time for emergency vehicles at road intersections contributes to society by saving lives.

Keywords: Computer vision, Emergency vehicles, Edge devices, IoT, LORA, Traffic.

INTRODUCTION

Due to the rapid evolution of the social economy and the advancement of science and technology, urban areas have undergone significant expansion. Consequently, the issue of traffic congestion has become exceedingly severe. In conventional traffic control systems, traffic lights and signs are limited to providing static and restricted guidance, making it challenging to optimize road resources [1]. We could address the problem by upgrading and optimizing the urban traffic system with the help of effective coordination and cooperation. To scientifically alleviate these pressures and improve convenience for urban residents, it is imperative to swiftly integrate new technologies to transform and enhance traditional traffic modes. Smart city traffic systems represent a significant breakthrough in this regard. Especially with the integration of Internet of Things (IoT) technology, advancements in new smart city traffic systems show significant potential for easing urban traffic congestion. Smart city transportation systems in urban areas address several practical issues by providing information about road conditions, traffic details, alternative routes, and service guidance, among others [2].

Leveraging Internet of Things (IoT) technology, the smart city traffic system adeptly gathers extensive road and vehicle data within the urban landscape. Subsequently, it dynamically calculates optimal route planning to facilitate the smooth flow of urban traffic [3]. Furthermore, integrating IoT technology is pivotal for augmenting the effectiveness of traffic command and scheduling. IoT-enabled traffic command and dispatch can more effectively anticipate and mitigate widespread traffic congestion, consequently ensuring a smoother travel experience for residents [4].

Numerous researchers are currently focusing on advancing the development of smart city transportation infrastructure. Research in this field is still in its early stages, yet it employs cutting-edge technology [5]. By implementing intelligent control systems, urban traffic signal lights can be managed intelligently, reducing traffic congestion [6]. Some notable works in traffic control systems are discussed below. For example, a study [7] reported that students at India's Chandigarh University have developed an intelligent, sensor-based traffic control system designed to give priority passage to ambulances and fire trucks. The presence of emergency service vehicles on the road is identified by RF (Radio Frequency) transmitters and receivers installed in the cars. The sound detection system in ambulances and fire trucks sends signals to the traffic controller, changing the signal to green while keeping the other signals red [8]. The RF receiver can detect the sound frequency of an emergency vehicle within a range of 328-656 feet (100-200m). Chitkara University students, in collaboration with the Mohali traffic police, have developed a 3D innovative traffic signal system aimed at regulating traffic signals using a wireless sensor system and providing a smart bird's-eye view. The system dynamically adjusts traffic movement based on the volume of approaching traffic through a self-sensing technique. It prioritizes green passages for ambulances and fire tenders while managing traffic flow. This innovative system has been deployed at the traffic crossing near Quark City on Airport Road, with an estimated cost of Rs 50,000.

Bengaluru, India, in partnership with the Japan International Cooperation Agency (JICA), successfully implemented an advanced traffic information and management system, involving an investment of Rs 72.86 crore [9]. The system has been deployed at 29 junctions, including prominent locations, such as G Road, Old Madras Road, and Hosur Road. It operates by automatically adjusting signals based on the volume of vehicles at each junction. The system continuously analyses real-time traffic congestion by being equipped with CCTV cameras to gauge vehicle length and size, and sensors positioned at intervals of 50, 100, and 150 meters along roads. It updates the integrated traffic management center accordingly. According to *Traffic Management to Get 'Smarter' in City* (2013) [10], the Hyderabad Traffic Integrated Management System (H-TRIMS) in India implemented intelligent traffic signals at the Jubilee Hills check post in Hyderabad. These signals regulate the duration of red and green lights at traffic junctions based on real-time vehicular traffic density. Cameras installed at this junction continuously monitor traffic density and automatically adjust the duration of green and red lights accordingly. The state government has allocated Rs. 66.50 crore for the H-TRIMS project.

As reported by Shivakumar, in 2019 [11], an intelligent traffic control system was deployed at 12 locations in Chennai city, India, to manage traffic flow efficiently. In partnership with the Japan International Cooperation Agency (JICA), this project has secured funding of Rs 660 crore for the implementation of an Intelligent Transport System. The system utilizes CCTV-based traffic management and incorporates bright signals, which predict traffic volume from upstream intersections to alleviate congestion. These smart signals automatically adjust timings based on traffic volume, effectively reducing congestion. Dubai launched a state-of-the-art, innovative traffic control system powered by artificial intelligence (AI) designed to predict and manage vehicle flow, formulate rapid

CHAPTER 10

Predicting Customer Churn in the Telecom Industry with Machine Learning Techniques for Improved Retention Strategies

Smital Anandrao Patil[1], **Rushi Patel**[2], **Swati Shukla**[3,*], **Yeswanth Satya Prasad**[4] and **Yagnesh Challagundla**[5]

[1] *School of Computer Science, KIT's College of Engineering, Kolhapur-416234, Maharashtra, India*

[2] *Department of Electronic and Telecommunication Engineering, Indian Institute of Technology - IIT, Kharagpur-721302, West Bengal, India*

[3] *School of Electronics Engineering (SENSE), VIT-AP University, Amaravati- 522237, Andhra Pradesh, India*

[4] *School of Computer Science and Engineering (SCOPE), VIT University, Vellore- 632014, Tamil Nadu, India*

[5] *School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, 522237, Andhra Pradesh, India*

Abstract: In the commercial world, customers are paramount, and understanding their behavior patterns can lead to highly impactful business decisions. Customer churn, defined as the rate at which consumers switch from one business to another, presents a significant obstacle, especially for businesses that rely on subscription models, such as telecommunications companies. Accurately forecasting customer turnover enables firms to effectively implement targeted retention initiatives and minimize income loss. The objective of this study is to employ logistic regression in order to forecast customer attrition for a telecoms company. We utilized a dataset that comprised comprehensive client information, encompassing demographics, service usage, and payment records. Through a thorough exploratory data analysis (EDA), we identified important patterns and trends that are directly linked to customer attrition. We performed data preprocessing by addressing missing values, encoding category variables, and normalizing numerical features. A robust statistical technique, specifically designed for binary classification, was employed to construct a predictive model. The discovered key predictive variables included tenure, monthly charges, and contract type. Our data indicate that clients who have been with us for a shorter period have higher monthly rates, and those on month-to-month contracts are more likely to discontinue their services. The regression model attains the highest accuracy, surpassing other machine learning techniques in terms of overall predictive capability.

* **Corresponding authors Swati Shukla:** School of Electronics Engineering (SENSE), VIT-AP University, Amaravati- 522237, Andhra Pradesh, India; E-mail: swati.shukla@vitap.ac.in

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

This study highlights the efficacy of this instrument as a valuable tool for forecasting customer turnover, enabling firms to enhance customer satisfaction and reduce churn rates by making informed decisions based on data.

Keywords: Classification models, Customer churn, Customer retention, Data analysis, Machine learning, Predictive modeling, Telecom industry.

INTRODUCTION

Customer churn, also known as attrition, is the act of losing clients or customers and is a vital measurement for businesses that rely on subscriptions, such as telecommunications companies. Anticipating customer attrition can have a substantial influence on a company's capacity to maintain its client base. The objective of this study is to forecast customer attrition for a telecommunications business by employing logistic regression. The data utilized in this investigation was obtained from a Kaggle dataset, consisting of 7043 distinct client entries with 21 characteristics, encompassing demographic details, services utilized, and monthly costs. Precise churn prediction enables telecom businesses to formulate focused strategies for retaining customers who are likely to leave, resulting in decreased turnover rates and enhanced profitability. This project encompassed various crucial stages: data acquisition, preprocessing, conversion, and modeling. Initially, the data underwent a process of cleaning and recoding to rectify any discrepancies and missing data. Descriptive statistics and visualizations provided a comprehensive analysis of the data's distribution, revealing any potential relationships between variables. After conducting the exploratory data analysis, we used logistic regression to create a model for predicting the binary outcome of churn. Logistic regression was used due to its straightforwardness and efficacy in managing categorical outcome variables. This approach enabled us to measure the magnitude of the influence of different predictor variables, such as tenure, contract type, and total costs, on the probability of customer turnover. The performance of the logistic regression model was assessed using common measures, including accuracy, precision, recall, and the confusion matrix.

The primary objective of this research is to identify the key indicators of customer churn and develop a reliable model that accurately predicts churn occurrences. Telecom firms may effectively manage customer churn by comprehending the reasons that lead to it. This understanding enables them to take proactive measures to address issues and create retention tactics, thereby ensuring a stable client base. This study demonstrates the practical application of logistic regression in predictive modeling and highlights the importance of making informed decisions based on data in the telecommunications industry.

RELATED WORK

Smith *et al.* [1] proposed a machine learning approach for forecasting customer turnover in the telecommunications industry. Their research explored various algorithms and emphasized the importance of selecting the optimal one based on specific data attributes. Lee *et al.* [2] performed a comparative analysis of machine learning techniques for predicting customer attrition in telecommunications networks. Their study evaluated the efficacy of several algorithms and provided valuable insights into their advantages and limitations for this specific application. Garcia [3] highlighted the significance of machine learning in enhancing client retention within the telecoms sector. This study examined the overarching benefits of employing machine learning to predict customer churn and refine strategies for customer retention. Chen and Zhang [4] concentrated on enhancing customer lifetime value in the telecommunications industry by utilizing machine learning to forecast churn. Their research examined the application of attrition prediction algorithms to enhance customer value for telecommunications firms. Khan *et al.* [5] presented a thorough examination of how artificial intelligence (AI) is used in managing customer turnover in the telecom industry. This paper examined various artificial intelligence (AI) methods used in churn prediction and their impact on client retention efforts. Another study [6] investigated the application of machine learning techniques to identify consumers at a higher risk of churn in telecom firms. At a conference, Patel (2023) [7] presented a discussion on the use of data-driven techniques for customer retention and the application of machine learning to anticipate churn in the telecommunications industry. Miller *et al.* [8] emphasized the cost-saving potential of machine learning in predicting customer turnover in the telecommunications industry, particularly highlighting its financial advantages for telecom corporations. Brown (2022) [9] explored the application of machine learning in the telecom industry for the purpose of customer retention. The author presented it as a means to anticipate and forecast future customer behavior. Williams (2020) [10] highlighted the significance of machine learning in ensuring customer satisfaction in the telecom industry by accurately anticipating customer turnover and proactively resolving any future problems. In their study, Gupta and Singh (2022) [11] investigated the application of deep learning techniques to predict customer churn in the telecommunications industry, highlighting the significant capabilities of deep learning in this specific field. Yang *et al.* (2022) [12] presented a cost-sensitive machine learning method for predicting customer attrition in telecommunications services. This study examined the financial implications of customer attrition and sought to refine prediction models accordingly. Manzoor *et al.* (2024) [13] presented an extensive examination of machine learning techniques for predicting customer attrition, providing suggestions for business professionals.

CHAPTER 11

A Water Quality Prediction and Assessment Model using Machine Learning Classifiers

M. Vaishali[1,*], **M. Monika**[2], **M. Atish**[3] and **Lalitha**[4]

[1] *Department of Computer Science and Engineering, Maharishi Markandeshwar Deemed to be University, Ambala, India*

[2] *Department of Information Technology, Dwarkadas J Sanghvi College of Engineering, Navi Mumbai, India*

[3] *Department of Computer Science and Engineering, BMS Institute of Technology and Management, Banglore, India*

[4] *School of Advanced Sciences, VIT-AP University, Andhra Pradesh, India*

Abstract: Ensuring safe and clean water availability is vital to the health of not only human beings but also all species. Additionally, it is crucial for the sustainability of the environment. With the emergence of advanced technologies like machine learning, predictive models can significantly contribute to assessing and managing water quality. Current research proposes a methodology that predicts water quality using several machine learning classifiers on a dataset comprising diverse parameters, such as pH levels, dissolved oxygen, turbidity, and other pollutants, collected from multiple water sources. Initially, the data were preprocessed to remove missing values and outliers. Feature engineering was employed to identify the most relevant parameters that contribute to water quality. Several popular machine learning classifiers, including Random Forest, Support Vector Machines, Decision Trees, and XGBoost, were evaluated and compared for their performance in predicting water quality. The trained models were validated and tested using cross-validation techniques to ensure generalizability and resilience. The research findings demonstrated that the proposed method is effective in accurately forecasting water quality levels. The XGBoost, in particular, exhibited superior performance with high accuracy and minimal overfitting. Additionally, feature importance analysis revealed key factors influencing water quality, providing valuable insights for policymakers and environmentalists.

Keywords: Cross-validation, Machine learning classifiers, ROC curve, Water quality assessment.

* **Corresponding author M. Vaishali:** Department of Computer Science and Engineering, Maharishi Markandeshwar Deemed to be University, Ambala, India; E-mail: drvaishaliwadhwa@gmail.com

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

INTRODUCTION

The primary source of clean and safe drinking water for residential homes, businesses, and agriculture is groundwater, which is sourced from rivers, streams, dams, lakes, reservoirs, creeks, and wetlands [1]. Groundwater is severely impacted by pollution from unrestricted human activities, including industrialization and agriculture, despite its essential role in the survival of humans and other ecosystems, such as aquatic plants and wildlife. Nearly 90% of untreated sewage from underdeveloped nations enters water sources, according to the 2015 United Nations World Water Development Report [2]. In the meantime, approximately 300–400 megatons of waste from the sector are dumped into water sources annually, according to data from the United Nations Educational, Scientific, and Cultural Organization (UNESCO) [3].

Most nations, including India, also have high rates of uncontrolled sewage treatment or discharge from manufacturing and agriculture [4]. River pollution is primarily caused by aluminum, ammoniacal nitrogen (NH3-N), suspended solids (SS), cadmium, chloramine, and biochemical oxygen demand (BOD). High BOD can be caused by wastewater and inadequate sewage treatment from the industry and agro-based sectors. Meanwhile, elevated NH 3-N may be caused by animal husbandry and unchecked household sewage. Furthermore, inappropriate earthworks and land-clearing practices may affect elevated SS [5]. To identify which areas are contaminated and require treatment, ongoing data on the water quality of these river pollution hotspots are needed. The evaluation of Indian water quality has been extensively covered in previous research. However, a new, effective method to track pollution hotspots has been developed due to the growth and complexity of massive data from ambiguous water quality metrics. To effectively identify hotspots for river pollution, this study suggests employing a range of techniques and comparative analyses that utilize both supervised and unsupervised Machine Learning (ML) algorithms. ML algorithms and water quality have been the subject of several studies [6 - 8]. However, to our knowledge, there has not been a thorough comparison research in the literature on using ML with datasets used to assess river quality. Hence, seven popular algorithms, including supervised and unsupervised ML, were utilized to classify the pollution hotspots in the Indian water. These include Support Vector Classifiers (SVC), K-Nearest Neighbors (KNN), AdaBoost (AB), Naïve Bayes (NB), Logistic Regression (LR), Random Forest (RF), Gradient Boost (XG), and Decision Tree (DT).

Additionally, we adjusted and modified every parameter in ML algorithms to validate the accuracy of the algorithms. This procedure is essential to prevent underfitting and overfitting in ML algorithms [9]. Overall, the accuracy of the river pollution categorization was used to compare and evaluate each ML method in detail in this chapter. The preprocessing techniques and algorithm tuning procedures yielded the highest accuracy.

LITERATURE SURVEY

Different nations and organizations have produced many WQIs, such as the surface quality water index and groundwater quality index, to achieve particular objectives [10]. Many factors, such as (i) uncertainty concerns, (ii) model reliability, (iii) transparency, and (iv) model sensitivity, have been brought up against this technique. WQI model created a significant amount of uncertainty in the final score, according to recent studies (Juwana *et al.*, 2016; Sutadian *et al.*, 2018; Uddin *et al.*, 2021) [1, 10, 11]. Furthermore, several studies have recently shown that the water quality index model employs several classification techniques [12, 13]. It is recommended to use various qualitative measurements, such as “excellent,” “good,” “bad,” “very bad,” “poor,” “marginal,” “higher,” and “lower,” when interpreting the WQI score. As a result, many approaches offer a variety of interpretations for comparable water qualities, which adds to the significant uncertainty surrounding the accurate categorization of water quality.

Recent research indicates that, due to these problems, the water resource managers receive unclear information from the current WQI model, which makes it difficult for bodies to act promptly [1]. As previously stated, the ultimate objective of the WQI model is to utilize a classification system to categorize water quality [3] [6]. The outcome of the WQI model is often a numerical number, also referred to as an index score, that ranges from 0 to 100, where 100 represents “good” water quality and zero means “worst” water quality [7, 14]. Furthermore, according to numerous recent studies, a significant degree of uncertainty has been introduced into the WQI system due to the improper application of classification techniques. Therefore, for comparable groups of water quality indicators, the traditional classification algorithms may yield conflicting results in the final water quality evaluation [2]. Thus, in an earlier work [1, 11], the authors presented a universal classification scheme (see Table **3**) for assessing coastal and transitional water quality to minimize erroneous assessment and optimize the metaphor difficulties of existing classification systems. According to Uddin *et al.* [1], the universal scheme outcomes for water quality could be helpful in accurately representing various scenarios of water quality [1]. This chapter contains details about the classification scheme and development process. This study employed a grading scheme to assess India's current water quality status. Several ML classifier techniques were used to assess the WQI models' performance after obtaining water quality classifications. To reduce model uncertainty and improve

CHAPTER 12

Synapsys – A Tool for Effective Video Summarization

S. Divya Meena[1], **B.V. Chandrahaas**[1], **C. Sai Janakiram**[1], **J.L.B. Krishna**[1], **M. Rushane**[1] and **J. Sheela**[1,*]

[1] *School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India*

Abstract: In the 21st century, the internet has become a ubiquitous platform for exchanging information, contributing to an extreme daily influx of approximately 2.5 Exabytes of data. Statistics show that a substantial 1/5th of the data transmitted over the internet comprises videos. With 30,000 hrs. of videos uploaded on YouTube every hour, the need to evaluate and derive valuable insights from this massive amount of video data becomes apparent in a dynamic landscape. Particularly for professionals engaged in internet exploration for various purposes, such as prominent data analysts, data scientists, and cybercrime analysts, it becomes apparent that they need to have an understanding of it. Though traditional machine learning-based systems exist, they have become obsolete due to modern times' issues, including decreased performance, greater computing complexity, and resource-intensive processes. GenAI has effectively addressed these problems, which is considered a transformative force with enhanced accuracy and semantic comprehension. This chapter focusses on the task of video summarization, which has been poised to supplant antiquated methodologies with the cutting-edge capabilities of GenAI. We have also applied the concepts of hybridization and fine-tuning of models in machine learning to create an improved version of generative models. The proposed methodology has demonstrated better performance in obtaining information, understanding computational complexity, *etc.*

Keywords: Abstractive summarization, Extractive summarization, Generative AI, Video summarization.

INTRODUCTION

Video summarization has become prevalent in the 21st century due to the inflow of data over the internet, not only for gaining valuable insights from videos, but every person has unique needs with the video content. Also, the versatility of the format of the videos makes them even more popular, and psychologically, people

* **Corresponding author J. Sheela:** School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India; E-mail: Sheela.j@vitap.ac.in

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

understand the content more accurately when it has been provided to them in the form of videos. Considering the above scenario, we can understand the need for video summarization, which helps the users gain insights without watching the entire content. This enables the respective platforms not only to reduce the rate of user interaction but also allow users to consume information in a very minimal amount of time. Now, the question of where this task of video summarization can be found useful often arises. The state-of-the-art video summarization can be used by platforms that host video content, commoners who consume video content, and students who usually rely on video content for knowledge and information. Also, the current hot topic of 2023 is TikTok resumes, where job applicants make a short video about themselves and forward it to the companies. In this case, multinational companies can also use such summarizer systems to screen candidates more quickly and efficiently. From the above context, we can see that these systems have a broader scope. A video summary can be understood as a concise paragraph of the entire video, highlighting the key information mentioned in it. The current system provides such a service with the help of generative artificial intelligence models, which consist of attention mechanisms as their backbone and can be deployed in scenarios where a video consists of background audio and both contexts are overlapped.

Related Work

Video summarization has witnessed significant developments, encompassing various methodologies and technologies. This section provides a comprehensive overview of related works, highlighting key studies, methodologies, and technological advancements that have contributed to the current state of the art. Early research on video summarization focused on extracting keyframes to condense video content effectively. Seminal works [1 - 4] introduced foundational concepts in keyframe-based summarization, laying the groundwork for subsequent developments. In a paper [5], this approach was extended by incorporating content-based retrieval techniques, enabling more nuanced video analysis. Another study [6] explored the application of visual saliency for summarizing soccer highlights, introducing the idea of leveraging visual attention to identify key events.

Divya et al.

A study [7] proposed a hierarchical approach to story-based video summarization, emphasizing the importance of organizing content at different levels of granularity. The advent of deep learning techniques has revolutionized video summarization. A study [8] surveyed the utilization of deep learning for image and video description, showcasing the potential of neural networks in understanding complex visual information. Another paper [9] introduced "BertSum," a model that utilizes pre-trained encoders like BERT for abstractive text summarization, influencing the text summarization component of our proposed Video Summarization System. Another study [10] proposed a multimodal approach integrating vision and text information for unsupervised video summarization. Their work highlights the significance of combining audio-visual cues and textual content to create more comprehensive and contextually rich summaries. Recent studies have explored the integration of transformer models for text summarization, showcasing their ability to generate coherent and contextually relevant summaries. A study [11] introduced GPT-3, a language model that exhibits few-shot learning capabilities, opening new possibilities for enhancing the text summarization aspect of our proposed Video Summarization System. The above works have inspired this study and aim to bring advancements in content-based video retrieval and the exploration of novel technologies for improving the overall efficiency and effectiveness of video summarization.

PROPOSED METHODOLOGY

Data Collection

Data collection becomes tedious while working with complex architectures like transformers, since they are well known for the errors in training cases with very little data. To avoid these problems, we have tried to collect as much data as possible and fine-tune the models with it so that the system performs better. We have used C4, HugeNews, and LibriSpeech datasets to fine-tune the models to create the system. C4 stands for Colossal Clean Crawled Corpus, a text dataset made by web crawling. It is a 750GB textual dataset for training LLMs and chatbots. HugeNews is a dataset comprising 1.5 billion news articles. The LibriSpeech corpus is a collection of approximately 1,000 hours of audiobooks that are a part of the LibriVox project. Most of the audiobooks come from Project Gutenberg. The training data is split into three partitions of 100hr, 360hr, and 500hr sets, while the dev and test data are divided into the 'clean' and 'other' categories, respectively.

Synapsys Architecture

After finding gaps in the previous literature mentioned above and getting exposure to the new technologies and methodologies available during the execution of this project, we came up with a method that integrates the power of generative AI models along with the traditional theory of hybrid models from machine learning. Above is the representation of the method in a flow diagram as shown in Fig. (**1**). Summarizing the flow diagram, initially, the user will approach our Synapsys platform with their already downloaded video or any other video

CHAPTER 13

HDFS: The Backbone of Big Data - A Review of High Availability, Scalability, and Performance Using Quorum Journal Manager and Dynamic Federated Metadata Management

Bhagavathi Santos Kumar[1] and **Virendra Kumar Shrivastava**[1,*]

[1] *Department of Computer Science and Engineering, School of Advanced Computing, Alliance University, Bangalore, Karnataka 562106, India*

Abstract: Working in tandem, the Hadoop Distributed File System (HDFS) plays a crucial role in managing large-scale data across distributed clusters. The basic structure of HDFS comprises the NameNode, which manages the file system namespace and client interactions, and the Data Nodes, which handle storage and data arrangement duties. Although HDFS faces constraints, particularly regarding scaling the NameNode, which is responsible for managing metadata, these constraints impact namespace scalability and performance. Therefore, this paper discusses the two architectures, the Quorum Journal Manager (QJM) components and the active and standby NameNodes, which enable the sharing of edit logs between them. To mitigate the risk of metadata corruption from split-brain scenarios, the system ensures that only a single NameNode writes to the JournalNodes. Through the utilization of the QJM, HDFS attains High Availability (HA), enabling swift failover in the event of machine crashes or scheduled maintenance. Dynamic Federated Metadata Management (DFMM) addresses the limitations of the current Hadoop architecture by distributing metadata management across multiple federated components. It entails dispersing metadata management among various federated components. HDFS can overcome the problems of connected block storage, namespace scalability, and performance bottlenecks by dynamically managing metadata on an exabyte scale. This work summarizes the key findings and presents a comparative analysis of both architectures. QJM performs a vital function that ensures the high availability, reliability, and fault tolerance of the NameNode in Hadoop's HDFS. DFMM enhances the fault tolerance, scalability, and performance of distributed file systems by distributing metadata across multiple servers or nodes.

Keywords: DFMM, Dynamic federated metadata management, Hadoop, HDFS, HDFS federation, QJM, Quorum journal manager.

[*] **Corresponding Author Virendra Kumar Shrivastava:** Department of Computer Science and Engineering, School of Advanced Computing, Alliance University, Bangalore, Karnataka-562106, India; E-mail: Virendra.shrivastava@alliance.edu.in

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

INTRODUCTION

Hadoop and HDFS play a crucial role in managing tasks, providing expandable, resilient storage and processing capabilities [1] across distributed computer clusters. Hadoop is an open-source framework that facilitates the distributed handling of large datasets over computer clusters using a simplified programming model. The fundamental structure consists of Hadoop, which comprises HDFS [2] for storage, MapReduce for concurrent processing, and Yet Another Resource Negotiator (YARN) for resource management, facilitating adaptable and robust data management. The architectural design of the HDFS [3] is structured with an enslaver/agent configuration, wherein the NameNode functions as the primary server, managing the file system's namespace and client file access. The worker nodes, known as DataNodes, manage storage on the physical devices they control and divide data into blocks to enhance processing efficiency. A Secondary Name Node [4] is included in the design to assist with administrative tasks for the Name Node, including checking file system information at checkpoints. By enhancing HDFS's fault tolerance, this setup ensures that data remains available throughout the entire hardware cluster. The inability of HDFS to scale the NameNode, which is responsible for managing metadata [5], in the same manner as it scales the DataNodes is a drawback. The inability of the current Hadoop architecture to scale Name Nodes highlights the importance and challenges of managing metadata at the Exabyte scale. Consequently, this leads to closely interconnected block storage, constraints on namespace scalability, and performance [6] bottlenecks. Several limitations constrain the NameNode within HDFS. A single point of vulnerability results in system downtime when it fails or requires maintenance. Memory limitations restrict its capacity to accommodate files, particularly impacting performance when dealing with small files that may create bottlenecks. The centralized structure hinders scalability, while HDFS lacks the capabilities for low-latency data access, which is essential for real-time applications. Thus, architectures are available to mitigate the limitations of the NameNode, which are discussed in sections **3** and **4** below.

The subsequent parts of this paper are organized as follows. Section **2** outlines the main points of QJM and DFMM, discussing the importance of these systems in managing data and ensuring high availability. Section **3** provides a brief overview of QJM, and Section **4** discusses the HDFS Federation. Section **5** provides a brief overview of the DFMM and discusses the two main metadata management techniques in brief. Section **6** presents a contrast between QJM and DFMM. Section 7 discusses the implications of the findings. Section **8** describes the conclusion and future research prospects.

BACKGROUND

Table **1** presents a summary of potential related research papers. QJM plays a pivotal core component across the HDFS framework developed by Apache Hadoop. It aims to secure HA for the HDFS NameNode, addressing the SPOF challenge in standard HDFS setups. QJM achieves this by replicating the NameNode's edit log, which contains metadata changes, within a configurable quorum of nodes across the cluster. If the Active NameNode [13] becomes unavailable or is undergoing maintenance, a different node in the quorum will automatically assume control, maintaining uninterrupted access to the file system. DFMM [14] was developed to enhance the scalability [15] and performance of expensive distributed file systems, specifically when traditional metadata methods become bottlenecks. DFMM [16] distributes metadata and partitions among several metadata servers or nodes in a federated system. DFMM enhances the efficiency of metadata access and modification by distributing the workload across multiple servers, thereby improving scalability constraints associated with a centralized metadata server. This approach balances workloads, strengthens resilience against failures, and allows for seamless adjustment to evolving workload patterns and system conditions. These systems regularly integrate techniques. These systems regularly incorporate optimization techniques to ensure peak performance in dynamic computing environments. QJM and DFMM are vital for managing data and maintaining HA in distributed file systems such as HDFS.

Table 1. A summary of similar work that has been reviewed.

Reference	Approach	Description
[7]	High Availability	Proposed a hierarchical architecture of Hadoop Nodes for HA HDFS, eliminating single-point node failure and establishing a model for optimal data processing and storage.
[8]	Chord protocol-based Architecture	Reduces the reliance on the metadata volume on the NameNode and provides a highly scalable and fault-tolerant HDFS. The Chord protocol directly connects to the HDFS to address scalability issues. However, it also increases the complexity of the Single HDFS NameNode Architecture.
[9]	Quorum Journal Manager	Manages Journal Nodes carrying edit logs of the NameNode, mitigating the SPOF.
[10]	Triple H	A Hybrid Design (Triple H) for HDFS on the HPC cluster was proposed to enhance HDFS's efficiency.
[11]	HDFS Federation	This proposed architecture mitigates SPOF.
[12]	Dynamic Federated Metadata Management (DFMM)	This proposed architecture mitigates all the limitations of NameNode.

CHAPTER 14

Optimized Gold Price Prediction Using Particle Swarm Optimization-Enhanced LSTM Networks

Alladi Vennela[1] and **Radha Mohan Pattanayak**[1,*]

[1] *School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh-522237, India*

Abstract: Gold has long been regarded as a valuable asset and a reliable haven for investors, particularly during periods of economic uncertainty. Gold price forecasting is difficult due to the market's intrinsic volatility and sensitivity to various factors, including global economic trends, inflation rates, currency value fluctuations, and geopolitical events. This study presents a novel strategy for predicting gold prices that combines Long Short-Term Memory (LSTM) neural networks and Particle Swarm Optimization (PSO). Traditional models struggle with these complications of uncertainty, but the PSO-LSTM technique improves LSTM hyperparameters for greater accuracy. The PSO-LSTM model beats both standalone deep Learning models and traditional forecasting, based on performance criteria such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE).

Keywords: Deep learning, Gold price, LSTM, Machine learning, Optimization, Time series.

INTRODUCTION

Gold has been a fundamental pillar in the intricate web of global economies and financial markets for a long time, symbolizing prosperity, firmness, and security. It is of great significance as one of the valuable metals utilized for financing trade transactions. In countries like the United States, India, the United Arab Emirates, China, and many others, it is highly esteemed for its role in jewelry. Furthermore, gold is commonly used as a valuable present or symbol of memory, and gold ornaments are frequently included in matrimonial agreements [1]. As more people become aware of and believe that gold can be owned as a low-risk investment asset, demand for gold is increasing daily. To mitigate financial losses, it is crucial to predict the value of gold [2] accurately.

* **Corresponding author Radha Mohan Pattanayak:** School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh-522237, India; E-mail: radhamohan.pattanayak@gmail.com

Suneeta Satpathy, Sachi Nandan Mohanty & Subhendu Kumar Pani (Eds.)

Researchers are currently observing a growing interest in gold acquisition, focusing on comprehending the specific characteristics that influence price fluctuations. Previously, numerous approaches were studied to predict the future value of gold, including Machine Learning models [3] such as ARIMA [4, 5], Least Square Support Vector Machine [6], Tree based models [7, 8], Multi Linear Regression Model [9, 10], Multivariate Stochastic Model [11, 12], Extreme Learning Machine [13, 14] and Fuzzy set-based models have also been widely used for time series predictions, such as high-order fuzzy time series [15, 16], hesitant fuzzy time series [17, 18], and neutrosophic fuzzy models [19, 20], demonstrating their effectiveness in forecasting complex datasets.

The proposed PSO-LSTM model aims to assist investors in making informed decisions by accurately predicting future gold prices. This article explains how PSO is combined with LSTM, describes the dataset and implementation steps, and reviews the experimental setup and performance results. The findings are discussed, and suggestions for future research in gold price prediction are provided.

The paper is organized as follows: Section **2** discusses earlier gold price forecast methodologies. Section **3** describes the methods employed in this investigation. Section **4** contains the findings and discussions, which compare the proposed model to benchmarks. Finally, Section **5** presents the key findings and offers suggestions for future research topics.

LITERATURE REVIEW

Estimating the prices of gold is of great importance to investors, financial institutions, and policymakers, leading to extensive studies in this area. This section provides a comprehensive review of the existing literature on gold price forecasting, encompassing methodologies, strategies, and key findings from previous studies. Traditional econometric models, such as ARIMA [21], have been extensively used for forecasting gold prices. Tanattrin Bunnag utilized ARIMA models to forecast gold prices based on historical data, achieving reasonable accuracy in short-term predictions [22].

Wenjing Fang proposed three methods for predicting gold prices: regression analysis, back propagation neural network (BP neural network), and time series analysis—the gold price dataset is used for prediction with high reliability and accuracy [23]. Minal Ghute and Mridula S. Korde analyzed economic variables and gold prices using machine learning algorithms. It compares four algorithms —decision tree, linear regression, random forest, support vector machine, and ridge regression —to predict future gold prices accurately [24].

Using machine learning algorithms, D. Nanthiya, S. B. Gopal, S. Balakumar, M. Harisankar, and S. P. Midhun. developed a model to predict daily and future gold prices. The study evaluated the performance of Linear Regression, ARIMA, and Random Forest models. The gold price dataset was used to assess accuracy, yielding the following results: Linear Regression achieved an MAE of 19.82 and an RMSE of 24.41, the ARIMA model had an MAE of 0.040 and an RMSE of 0.046, and the Random Forest model recorded an MAE of 0.150 and an RMSE of 0.156 [25].

Dasari Siva Sankar and Harish Sahu suggested a Random Forest regressor approach for predicting gold prices. The R-squared error was utilized to study the influence of the number of estimators on accuracy and prediction time, and the Random Forest model was trained on a dataset with a relative squared error of 0.99077 [26]. Ziyang Yuan implemented KNN, XGBoost, and Light GBM models to predict Bitcoin and Gold prices. The OpenCV dataset was utilized for accuracy in predicting Bitcoin and gold prices. Results guide investors and offer ML insights for price forecasting [27].

METHODOLOGY

Data Collection

The dataset from Kaggle contains historical gold prices with key attributes: Date, Price (closing), Open, High, Low, Volume, and Change Percentage. The 'Date' column represents the trading day, while 'Price' shows the closing price. 'Open,' 'High,' and 'Low' reflect the session's opening, highest, and lowest prices. 'Volume' indicates the total gold traded, and 'Change %' represents the price change compared to the previous day. The data can be collected from Kaggle platforms that provide historical gold price records over a desired period. Fig. (**1**) below represents the flowchart of the proposed model.

Data Preparation

The data must be pre-processed before being fed into the prediction model to ensure its quality and usefulness.

- Data Cleaning: Handling missing values or outliers that may exist in the dataset.
- Normalization: Scaling the data within a specific range, commonly between 0 and 1, to help the LSTM model perform optimally.
- Feature Engineering: Extracting essential features from the raw data to improve the model's predictive power. For instance, additional features could be created based on moving averages, momentum indicators, or lagged price data.

SUBJECT INDEX

A

AdaBoost 81, 82, 87, 90, 154
AdaDelta 66, 73, 75, 77, 78
Adagrad 66, 73, 75, 76, 77, 78
Adam 48, 66, 72, 73, 74, 75, 76, 77, 78, 79, 149
  Adam Optimizer 48, 73
Agriculture 1, 6, 17, 18, 19, 27, 29, 32, 54, 55, 57, 63, 64, 82, 103, 104, 154, 197
  Precision Farming 54, 55, 57
  Sustainable Agriculture 27, 54, 55
Air Quality 54, 56, 57, 59, 62, 108, 110, 113, 114, 117, 118, 119, 122
  Monitoring 108, 110, 113, 117, 122
  Sensor 56, 59, 122
Almond Leaf Scorch Disease (ALSD) 26, 27
Ambulance 126, 127, 128, 130, 131, 132, 133, 135, 137
Aquaponics 66, 67, 74, 75, 76, 79
Arduino 55, 56, 57, 119, 121, 135, 137
Artificial Intelligence (AI) 1, 17, 18, 27, 55, 68, 70, 81, 82, 93, 94, 96, 97, 98, 100, 102, 103, 104, 128, 129, 144, 183, 185, 193
  Generative AI (GenAI) 27, 169, 172
AUC (Area Under the Curve) 89, 147, 150, 161, 162
Augmentation 42, 43, 44, 45, 47, 50, 148, 149
  Data 42, 43, 44, 45, 47, 50, 57
  PCA Color 45
  Positional 45

B

Bacterial Diseases 3
Bagging 47, 87, 159
Bayesian 6, 7, 69, 70, 71, 145, 164
  Classifier 6
  Optimization 70, 71, 145
Big Data 55, 115, 129, 180, 181, 192
Biogas 82, 83, 87, 91
Binary Classification 85, 86, 87, 90, 142, 143, 146, 156, 158
Blynk 118, 119, 122
Boosting 47, 83, 87, 159
Brown Spot 2, 6, 7
Business Performance 93, 95, 96, 97, 98, 100, 102, 143, 149, 151
  Organization Performance 96, 97, 98, 172
  Firm Performance 93, 97, 98, 102, 181

C

Carbon Dioxide (CO_2) 56, 81, 82, 87, 89, 91, 108, 117, 118, 122
CFA (Confirmatory Factor Analysis) 93, 100
Clustering 6, 7, 8, 9, 25, 47, 107
  K-Means Clustering 6, 7, 9, 45
  Spectral Clustering 25, 33, 34
CM (Confusion Matrix) 50, 143, 146, 156, 160, 164
CNN (Convolutional Neural Network) 6, 7, 10, 18, 43, 45, 48, 50, 70, 71, 79, 122, 134, 138, 139, 140
  Deep CNN 7, 43
  Fully Convolutional Neural Networks 21
  Spatial-Spectral Two-Branch CNN 27, 35
  Improved EfficientNet 28, 36
Color Correction 8, 10
Composting 81, 83
Conductivity 157
Contamination 25, 33, 34
  Aflatoxin 25, 33, 34
Contract Type 142, 143, 147, 149, 151
Convergence 68, 70, 71, 73, 75, 77, 78, 79, 144
Cost Management 94, 96, 101
Cross-Validation 47, 153
Customer Churn 142, 143, 144, 145, 146, 147, 149, 150, 151
  Turnover 142, 143, 144, 146, 147, 151

D

Data Acquisition 8, 119
Data Analytics 54, 55, 59, 103, 115
Data Augmentation 42, 43, 44, 45, 47, 50
  Diffusion Model 42, 43, 45, 47, 50
Data Ingestion 54
Data Preprocessing 46, 47, 50, 142, 146, 148, 198
Data Quality 27, 46, 153, 154, 160, 162, 164, 165
Data Resizing 8, 46, 75
Data Science 1, 17, 26, 29, 36, 42, 54, 66, 81, 93, 108, 126, 142, 153, 169, 180, 196
Decision Making 93, 94, 96, 97, 103, 142, 143, 147
Decision Tree (DT) 21, 81, 83, 86, 87, 90, 149, 151, 153, 154, 155, 158, 159, 164, 165, 198
Defuzzification 57, 58, 60
Detection 1, 6, 7, 8, 10, 21, 22, 25, 27, 28, 33, 34, 35, 42, 45, 50, 56, 60, 64, 70, 71, 83, 108, 110, 113, 114, 117, 118, 121, 122, 126, 129, 130, 135, 137, 140, 170
  Disease 1, 6, 7, 10, 27, 42, 45, 50, 122
  Edge 50, 121
  Fraud 96
  Intrusion 70
  Object 8, 134, 138
  Traffic 135, 137
Dissolved Oxygen (DO) 74, 153, 157
Distributed File System (DFS) 180, 181, 182, 183, 185, 187, 188, 191, 192, 193
  Hadoop (HDFS) 180, 181, 182, 183, 185, 186, 187, 191, 192, 193
  HDFS Federation 180, 186, 187
DSL (Digital Subscriber Line) 147, 149, 151
Dry Fruit 17, 18, 19, 21, 22, 25, 27, 28, 29, 32

E

Early Blight 42, 43, 46, 47, 48, 50
EFA (Exploratory Factor Analysis) 93, 100
Edge Computing 130, 135, 137, 139, 193
EDA (Exploratory Data Analysis) 83, 87, 142, 143, 146, 147, 150
E-nose 108, 117, 118, 119, 122
Ensemble Learning 87, 159, 160
Entropy 86, 87, 91, 98
Environmental Monitoring 103, 111, 122, 123
Evapotranspiration (ETa) 29, 35
Extractive Summarization 169, 172
Extreme Gradient Boosting (XGBoost) 145, 153, 154, 155, 158, 159, 160, 161, 164, 166

F

F1 Score 7, 87, 90, 143, 146, 147, 150, 157, 158, 160, 161, 164
Fault Tolerance 180, 181, 182, 183, 185, 186, 188, 191, 192, 193
Feature Extraction 6, 7, 8, 9, 10, 27, 42, 45, 50, 85, 153, 198
Feature Importance 87, 153, 154, 160
Feature Selection 7, 85, 153
Feedforward Neural Network (FNN) 66, 71, 72, 74, 75, 79
Fuzzy Logic 54, 55, 57, 58, 59, 60, 63
  Inference System (FIS) 57, 58, 60
Fuzzification 57, 58

G

GAN (Generative Adversarial Network) 43, 44, 45, 169, 170, 172, 174, 177
Gaussian Blur 8, 120
Gaussian Filtering 50, 120
Gaussian Naïve Bayes 149
Gaussian Noise 44
Genetic Algorithms 70, 71, 103
Gini Index 87, 159
GLCM (Grey Level Co-occurrence Matrices) 6, 9
GNSS (Global Navigation Satellite System) 129
Gold Price 196, 197, 198, 200, 201, 202, 203
Gradient Boosting (XGBoost) 83, 145, 153, 154, 155, 158, 159, 160, 161, 164, 166
Greenhouse Gases 81, 82, 87, 89, 91

H

Hadoop 180, 181, 182, 183, 186, 188, 191, 192, 193
  Distributed File System (HDFS) 180, 181, 182, 183, 185, 186, 187, 191, 192, 193
  HDFS Federation 180, 186, 187
Hashing Technique 189, 190, 191
Hazelnut 21, 22, 25, 26, 32, 33, 34, 35

High Availability (HA) 180, 182, 183, 185, 191, 192, 193
Human Capital 93, 94, 95, 96, 97, 98, 100, 102, 103
Hybrid Model 83, 145, 169, 172, 177, 197, 202, 203
Hyperparameter Optimization 45, 70, 71, 196, 199, 200, 202, 203
Hyperspectral Imaging 17, 18, 19, 27, 28, 29, 32, 35, 36, 37

I

Image Processing 1, 6, 8, 10, 26, 45, 46, 48, 50, 114, 115
Information Gain 86, 87, 159
IoT (Internet of Things) 54, 55, 56, 62, 74, 108, 109, 110, 111, 113, 114, 115, 117, 119, 121, 122, 126, 127, 135, 192
  E-parking System 115, 117
  Solutions 108, 113, 122
  Urban 110, 113, 135

J

JournalNode 180, 182, 183, 185

K

KDE Plot 87, 88
KNN (K-Nearest Neighbors) 6, 11, 82, 86, 90, 154, 155, 158, 159, 164, 166, 198
K-Means Clustering 6, 7, 9, 45

L

Late Blight 42, 43, 46, 47, 48, 50
Learning Rate 68, 73, 75, 76, 78, 149, 199
Long Short-Term Memory (LSTM) 83, 196, 197, 198, 199, 200, 202, 203
LoRa (Long Range) 126, 130, 135
Loss Function 48, 66, 72, 75, 77, 78, 79, 196, 201

M

MAE (Mean Absolute Error) 196, 198, 201
Mamdani FIS 55, 57, 60
MAPE (Mean Absolute Percentage Error) 55, 196, 201, 202
MapReduce 181
MATLAB 55, 60
Mediation 93, 98, 102, 103
Metadata Management 180, 181, 182, 186, 187, 188, 191, 192, 193
  Dynamic Federated 180, 182, 187, 188, 189, 191, 192, 193
Metrics 17, 19, 35, 37, 55, 66, 70, 89, 90, 114, 122, 143, 146, 156, 160, 161, 164, 175, 176, 196, 201, 202
  Performance 17, 19, 35, 37, 66, 114, 122, 143, 146, 156, 160, 161, 164, 175, 176, 196, 201, 202
Microorganisms 82
MongoDB 121
MQTT (Message Queuing Telemetry Transport) 116, 121, 122
Multi-class Classification 86, 156, 159, 162, 164, 166

N

Naïve Bayes (NB) 82, 86, 90, 149, 150, 154, 155, 156, 158, 159, 164, 165
NameNode 180, 181, 182, 183, 185, 186, 187, 191, 192, 193
Natural Language Processing (NLP) 99, 175, 177
Natural Gas 113
Neural Networks 55, 66, 68, 70, 71, 72, 74, 75, 79, 83, 149, 150, 196, 200, 203
  Artificial 6, 83, 142
  Convolutional 6, 7, 10, 18, 43, 45, 48, 50, 70, 71, 79, 122, 134, 138, 139, 140
  Deep 68, 69, 71, 83
  Feedforward 66, 71, 72, 74, 75, 79
  Long Short-Term Memory (LSTM) 83, 196, 197, 198, 199, 200, 202, 203
  Recurrent 200, 202
  Self-Organizing Map (SOM) 22, 26, 31
NIR (Near-Infrared) Spectroscopy 21, 22, 31, 32
Noise Monitoring 108, 110, 113, 114, 119, 121, 122
Normalization 46, 153, 160, 198

O

Optimization 45, 66, 68, 69, 70, 71, 72, 73, 75, 76, 77, 78, 79, 81, 82, 85, 108, 110, 143, 196, 197, 199, 200, 202, 203
  Hyperparameter 45, 70, 71, 196, 199, 200, 202, 203
  Particle Swarm 196, 197, 199, 200, 202, 203
Outliers 83, 87, 90, 153, 198
Overfitting 45, 47, 48, 76, 85, 87, 154, 159

P

Paddy Crop Diseases 1, 6
Parking Management System (SPMS) 108, 115, 116, 117, 121, 122
Particle Swarm Optimization (PSO) 196, 197, 199, 200, 202, 203
PCA (Principal Component Analysis) 25, 45, 85, 91, 100, 160
PEGASUS 172, 174, 175, 176
pH Level 56, 74, 153, 162
Pistachio 26
Precision 17, 19, 45, 69, 87, 90, 143, 146, 147, 150, 151, 154, 157, 158, 160, 161, 164, 175, 196, 203
  Average 21, 31, 139
Preprocessing 6, 7, 8, 46, 47, 50, 83, 85, 119, 121, 142, 146, 148, 153, 155, 160, 198
Propane 56, 117, 118, 122

Q

QJM (Quorum Journal Manager) 180, 182, 183, 185, 186, 191, 192, 193
Quorum 180, 182, 183, 185, 191, 192, 193
  Nodes 183, 185

R

Random Forest (RF) 21, 26, 28, 81, 82, 83, 86, 87, 90, 149, 151, 153, 154, 155, 158, 159, 164, 165, 198, 202
Raissin 21, 22, 26, 32
Recall 19, 45, 87, 90, 143, 146, 147, 150, 157, 158, 160, 161, 164, 175, 176
Regression 27, 29, 72, 79, 81, 82, 83, 85, 91, 99, 142, 143, 146, 147, 149, 150, 151, 153, 154, 159, 197, 198, 202
  Linear 82, 85, 91, 198, 202
  Logistic 7, 81, 83, 85, 86, 90, 142, 143, 146, 154, 159
  Partial Least Squares 21, 29, 36
  Stacking 27, 35
  Support Vector Machine 21, 22, 31, 36
ResNet 42, 45, 48, 50
  ResNet-18 42, 44, 48, 50
RFID (Radio Frequency Identification) 113, 116, 121
RMSE (Root Mean Square Error) 22, 32, 35, 87, 90, 196, 198, 201, 202
ROC (Receiver Operating Characteristic) Curve 87, 153, 156, 160, 161, 162
ROUGE (Recall-Oriented Understudy for Gisting Evaluation) 175, 176

S

Scalability 10, 17, 18, 55, 87, 103, 114, 122, 140, 180, 181, 182, 186, 188, 191, 192, 193
Segmentation 6, 7, 8, 26, 27, 31, 35, 45, 50, 116, 151
  Image 8, 50, 116
  Semantic 8, 9
Sensors 54, 55, 56, 57, 59, 62, 63, 64, 74, 108, 110, 113, 114, 115, 116, 117, 118, 119, 121, 122, 127, 129, 135, 192
  Air Quality 56, 59, 122
  Acoustic 37
  LDR 56
  Smart 114
  Soil Moisture 55, 56, 57, 60, 62
  Temperature 57, 60, 118
Sheath Blight 3, 7
Sheath Rot 3
Smart Cities 108, 109, 110, 111, 112, 113, 114, 115, 116, 122, 127
Smart Parking 108, 110, 111, 113, 114, 115, 116, 121, 122
SPOF (Single Point of Failure) 182, 183, 192
SGD (Stochastic Gradient Descent) 66, 69, 70, 72, 73, 75, 76, 77, 78, 79
Subtree Partitioning 189, 191
Summarization 169, 170, 171, 172, 174, 175, 176, 177, 178
  Abstractive 169, 171, 174
  Extractive 169, 172
  Video 169, 170, 171, 172, 177

Support Vector Machine (SVM) 6, 81, 82, 86, 90, 149, 150, 153, 154, 155, 158, 159, 164, 165, 198, 202

T

TD-NMR (Time-Domain Nuclear Magnetic Resonance) 25, 33, 34
Telecommunications (Telecom) 142, 143, 144, 147, 149, 151
Tenure 142, 146, 147, 149, 151
THz (Terahertz) Radiation 22, 32
ThingSpeak 54, 55, 57, 62, 121
Time Series 196, 197, 200
Traffic Control System 126, 127, 128, 129, 130, 134, 135, 137, 139, 140
Traffic Density 126, 127, 128, 130, 133, 134, 135, 138, 139
Traffic Management 108, 126, 127, 128, 129, 130, 135, 137, 139, 140
Transfer Learning 6, 7, 11, 45

U

UAV (Unmanned Aerial Vehicle) 21, 22, 26, 32
Underfitting 154
Unsupervised Learning 9, 149, 154

V

Validation 46, 49, 50, 70, 75, 83, 90, 153, 157, 160, 161, 164, 165, 174, 192, 200
VGG (Visual Geometry Group) 45, 48
Viral Diseases 4
VIS (Visible) Spectroscopy 21, 22, 31
Visualization 55, 56, 62, 87, 108, 113, 114, 119, 121, 122, 143, 146, 147, 148, 169, 170
Voting 47, 146, 159, 166

W

Water Pumping Time (WPT) 59, 60, 62
Water Quality (WQ) 74, 153, 154, 155, 156, 157, 160, 162, 164, 166
  Index (WQI) 87, 155, 156, 157, 160, 162, 164, 166
Weight Initialization 70
Whisper 172, 174
Wind Power 70

Y

YOLO (You Only Look Once) 21, 28, 36, 126, 130, 134, 138, 139
YARN (Yet Another Resource Negotiator) 181

Suneeta Satpathy

Dr. Suneeta Satpathy is an Associate Professor at the Center for Cyber Security, SOA University, Bhubaneswar, with a Ph.D. in Computer Science from Utkal University (2015) supported by the Directorate of Forensic Sciences, MHA scholarship. Her research focuses on Computer Forensics, Cyber Security, Data Fusion, Data Mining, Big Data Analysis, and Decision Mining. She has edited books with Springer, Wiley, CRC AAP, and NOVA Publications, published widely in reputed international journals and conferences, and guided numerous graduate and postgraduate students. Actively engaged in professional service, she is a reviewer or editorial board member for leading journals such as Robotics and Autonomous Systems (Elsevier), Computational and Structural Biotechnology Journal (Elsevier), and Journal of Big Data (Springer). She is also an active member of CSI, ISTE, OITS, IE, Nikhil Bharat Shiksha Parisad, and IEEE.

Sachi Nandan Mohanty

Dr. Sachi Nandan Mohanty, ranked among the Top 2% World Scientists by Stanford University and Elsevier (2023–2025), completed his postdoctoral research at IIT Kanpur and Ph.D. at IIT Kharagpur with an MHRD fellowship. He has published over 240 research papers and authored/edited 42 books with leading publishers including Springer, Wiley, CRC Press, and IEEE-Wiley. His research covers Data Mining, Big Data Analytics, Cognitive Science, Fuzzy Decision Making, Brain–Computer Interface, and Computational Intelligence. Recipient of multiple awards including the Best Ph.D. Thesis Award (2015) from the Computer Society of India and four Best Paper Awards, he has guided nine Ph.D. scholars and 23 postgraduates. A Fellow of IEI and IETE, Senior Member of IEEE, and Ambassador of EAI, he also serves as Editor-in-Chief of EAI Transactions on Intelligent Systems and Machine Learning Applications.

Subhendu Kumar Pani

Dr. Subhendu Kumar Pani is a Professor at Krupajal Engineering College, Biju Patnaik University of Technology, Odisha, and serves as Editor-in-Chief of multiple international book series with Routledge (Taylor & Francis), CRC Press, Wiley, Bentham Science, and NOVA Publishers. He has also been a guest editor for journals such as the MDPI Journal of Sensors. His multidisciplinary research spans Artificial Intelligence, Health Informatics, IoT, Collective Computational Intelligence, Robotics, Social Computing, Web Intelligence, Web Services, and Data Mining, with over 300 research publications, including 40+ books and 80+ Scopus-indexed works. Actively engaged in global academia, he has organized numerous conferences, contributed to over 150 international program committees, and served on several editorial boards. Recognized with five prestigious research awards, Dr. Pani earned his Ph.D. in Computer Science from Utkal University (2013) and M.Tech. from KIIT University (2007). He is a Fellow of the Scientific Society of Advanced Research and Social Change and a life member of IE, ISTE, ISCA, OBA, OMS, SMIACSIT, SMUACEE, and CSI.