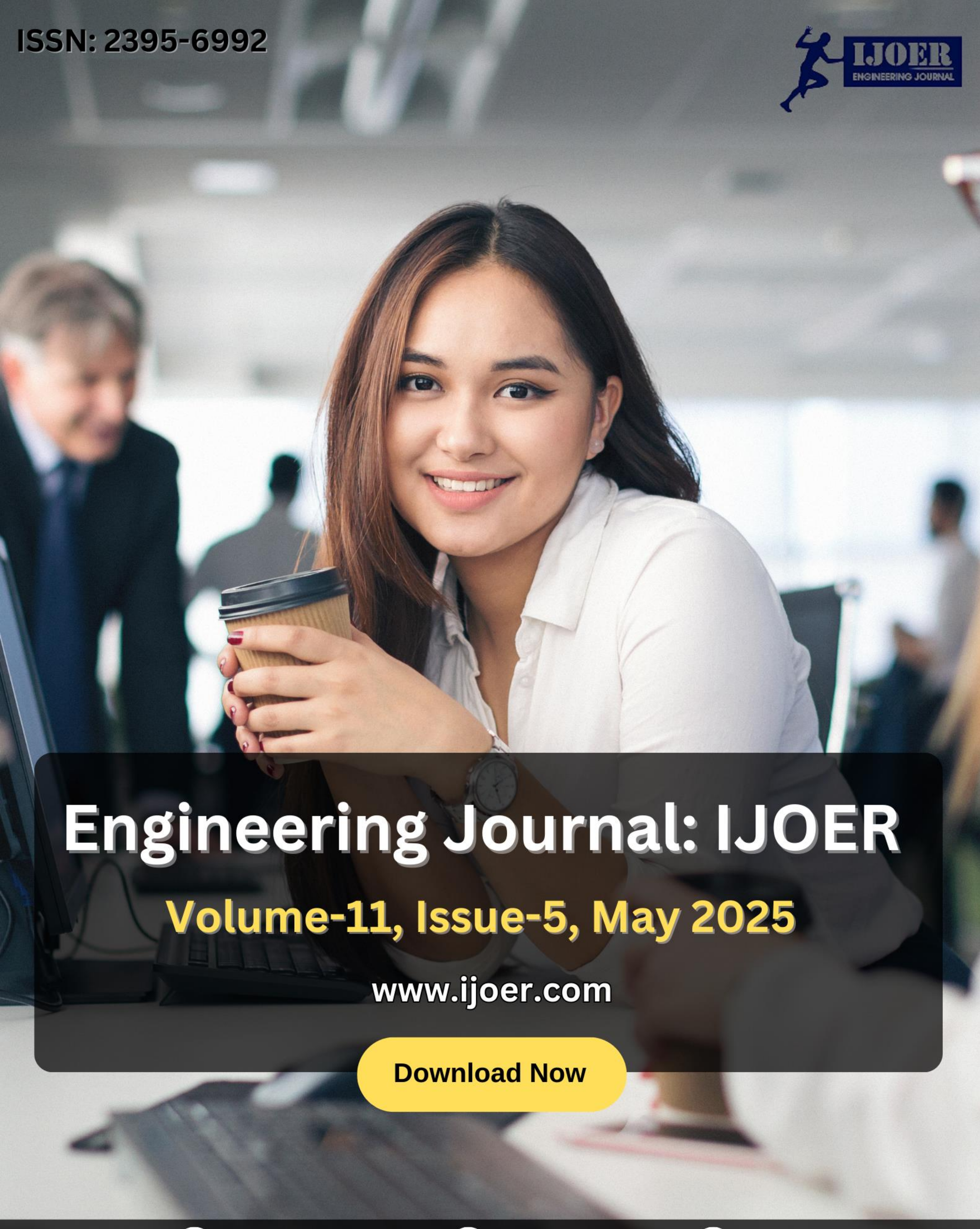
ISSN: 2395-6992



Engineering Journal: IJOER

Volume-11, Issue-5, May 2025

www.ijoer.com

Download Now

Contact us +91-7665235235 www.ijoer.com info@ijoer.com

Preface

We would like to present, with great pleasure, the inaugural volume-11, Issue-5, May 2025, of a scholarly journal, *International Journal of Engineering Research & Science*. This journal is part of the AD Publications series *in the field of Engineering, Mathematics, Physics, Chemistry and science Research Development*, and is devoted to the gamut of Engineering and Science issues, from theoretical aspects to application-dependent studies and the validation of emerging technologies.

This journal was envisioned and founded to represent the growing needs of Engineering and Science as an emerging and increasingly vital field, now widely recognized as an integral part of scientific and technical investigations. Its mission is to become a voice of the Engineering and Science community, addressing researchers and practitioners in below areas:

Chemical Engineering	
Biomolecular Engineering	Materials Engineering
Molecular Engineering	Process Engineering
Corrosion Engineering	
Civil Engineering	
Environmental Engineering	Geotechnical Engineering
Structural Engineering	Mining Engineering
Transport Engineering	Water resources Engineering
Electrical Engineering	
Power System Engineering	Optical Engineering
Mechanical Engineering	
Acoustical Engineering	Manufacturing Engineering
Optomechanical Engineering	Thermal Engineering
Power plant Engineering	Energy Engineering
Sports Engineering	Vehicle Engineering
Software Engineering	
Computer-aided Engineering	Cryptographic Engineering
Teletraffic Engineering	Web Engineering
System Engineering	
Mathematics	
Arithmetic	Algebra
Number theory	Field theory and polynomials
Analysis	Combinatorics
Geometry and topology	Topology
Probability and Statistics	Computational Science
Physical Science	Operational Research
Physics	
Nuclear and particle physics	Atomic, molecular, and optical physics
Condensed matter physics	Astrophysics
Applied Physics	Modern physics
Philosophy	Core theories

Chemistry	
Analytical chemistry	Biochemistry
Inorganic chemistry	Materials chemistry
Neurochemistry	Nuclear chemistry
Organic chemistry	Physical chemistry
Other Engineering Areas	
Aerospace Engineering	Agricultural Engineering
Applied Engineering	Biomedical Engineering
Biological Engineering	Building services Engineering
Energy Engineering	Railway Engineering
Industrial Engineering	Mechatronics Engineering
Management Engineering	Military Engineering
Petroleum Engineering	Nuclear Engineering
Textile Engineering	Nano Engineering
Algorithm and Computational Complexity	Artificial Intelligence
Electronics & Communication Engineering	Image Processing
Information Retrieval	Low Power VLSI Design
Neural Networks	Plastic Engineering

Each article in this issue provides an example of a concrete industrial application or a case study of the presented methodology to amplify the impact of the contribution. We are very thankful to everybody within that community who supported the idea of creating a new Research with IJOER. We are certain that this issue will be followed by many others, reporting new developments in the Engineering and Science field. This issue would not have been possible without the great support of the Reviewer, Editorial Board members and also with our Advisory Board Members, and we would like to express our sincere thanks to all of them. We would also like to express our gratitude to the editorial staff of AD Publications, who supported us at every stage of the project. It is our hope that this fine collection of articles will be a valuable resource for *IJOER* readers and will stimulate further research into the vibrant area of Engineering and Science Research.

Mukesh Arora

(Chief Editor)

Board Members

Dr. M. Varatha Vijayan

Annauniversity Rank Holder, Commissioned Officer Indian Navy, Ncc Navy Officer (Ex-Serviceman Navy), Best Researcher Awardee, Best Publication Awardee, Tamilnadu Best Innovation & Social Service Awardee From Lions Club.

Dr. Mohamed Abdel Fatah Ashabrawy Moustafa

PhD. in Computer Science - Faculty of Science - Suez Canal University University, 2010, Egypt.

Assistant Professor Computer Science, Prince Sattam bin AbdulAziz University ALkharj, KSA.

Prof.S.Balamurugan

Dr S. Balamurugan is the Head of Research and Development, Quants IS & CS, India. He has authored/co-authored 35 books, 200+ publications in various international journals and conferences and 6 patents to his credit. He was awarded with Three Post-Doctoral Degrees - Doctor of Science (D.Sc.) degree and Two Doctor of Letters (D.Litt) degrees for his significant contribution to research and development in Engineering.

Dr. Mahdi Hosseini

Dr. Mahdi did his Pre-University (12^{th}) in Mathematical Science. Later he received his Bachelor of Engineering with Distinction in Civil Engineering and later he Received both M.Tech. and Ph.D. Degree in Structural Engineering with Grade "A" First Class with Distinction.

Dr. Anil Lamba

Practice Head – Cyber Security, EXL Services Inc., New Jersey USA.

Dr. Anil Lamba is a researcher, an innovator, and an influencer with proven success in spearheading Strategic Information Security Initiatives and Large-scale IT Infrastructure projects across industry verticals. He has helped bring about a profound shift in cybersecurity defense. Throughout his career, he has parlayed his extensive background in security and a deep knowledge to help organizations build and implement strategic cybersecurity solutions. His published researches and conference papers has led to many thought provoking examples for augmenting better security.

Dr. Ali İhsan KAYA

Currently working as Associate Professor in Mehmet Akif Ersoy University, Turkey.

Research Area: Civil Engineering - Building Material - Insulation Materials Applications, Chemistry - Physical Chemistry – Composites.

Dr. Parsa Heydarpour

Ph.D. in Structural Engineering from George Washington University (Jan 2018), GPA=4.00.

Dr. Heba Mahmoud Mohamed Afify

Ph.D degree of philosophy in Biomedical Engineering, Cairo University, Egypt worked as Assistant Professor at MTI University.

Dr. Kalpesh Sunil Kamble (Ph.D., P.Eng., M.Tech, B.E. (Mechanical))

A distinguished academic with a Ph.D. in Mechanical Engineering and 13 Years of extensive teaching and research experience. He is currently a Assistant professor at the SSPM's COE, Kankavli and contributes to several undergraduate and masters programs across Mahrashtra, India.

Dr. Aurora Angela Pisano

Ph.D. in Civil Engineering, Currently Serving as Associate Professor of Solid and Structural Mechanics (scientific discipline area nationally denoted as ICAR/08"–"Scienza delle Costruzioni"), University Mediterranea of Reggio Calabria, Italy.

Dr. Faizullah Mahar

Associate Professor in Department of Electrical Engineering, Balochistan University Engineering & Technology Khuzdar. He is PhD (Electronic Engineering) from IQRA University, Defense View, Karachi, Pakistan.

Prof. Viviane Barrozo da Silva

Graduated in Physics from the Federal University of Paraná (1997), graduated in Electrical Engineering from the Federal University of Rio Grande do Sul - UFRGS (2008), and master's degree in Physics from the Federal University of Rio Grande do Sul (2001).

Dr. S. Kannadhasan

Ph.D (Smart Antennas), M.E (Communication Systems), M.B.A (Human Resources).

Dr. Christo Ananth

Ph.D. Co-operative Networks, M.E. Applied Electronics, B.E Electronics & Communication Engineering Working as Associate Professor, Lecturer and Faculty Advisor/ Department of Electronics & Communication Engineering in Francis Xavier Engineering College, Tirunelveli.

Dr. S.R.Boselin Prabhu

Ph.D, Wireless Sensor Networks, M.E. Network Engineering, Excellent Professional Achievement Award Winner from Society of Professional Engineers Biography Included in Marquis Who's Who in the World (Academic Year 2015 and 2016). Currently Serving as Assistant Professor in the department of ECE in SVS College of Engineering, Coimbatore.

Dr. Balasubramanyam, N

Dr.Balasubramanyam, N working as Faculty in the Department of Mechanical Engineering at S.V.University College of Engineering Tirupati, Andhra Pradesh.

Dr. PAUL P MATHAI

Dr. Paul P Mathai received his Bachelor's degree in Computer Science and Engineering from University of Madras, India. Then he obtained his Master's degree in Computer and Information Technology from Manonmanium Sundaranar University, India. In 2018, he received his Doctor of Philosophy in Computer Science and Engineering from Noorul Islam Centre for Higher Education, Kanyakumari, India.

Dr. M. Ramesh Kumar

Ph.D (Computer Science and Engineering), M.E (Computer Science and Engineering).

Currently working as Associate Professor in VSB College of Engineering Technical Campus, Coimbatore.

Dr. Maheshwar Shrestha

Postdoctoral Research Fellow in DEPT. OF ELE ENGG & COMP SCI, SDSU, Brookings, SD Ph.D, M.Sc. in Electrical Engineering from SOUTH DAKOTA STATE UNIVERSITY, Brookings, SD.

Dr. D. Amaranatha Reddy

Ph.D. (Postdocteral Fellow, Pusan National University, South Korea), M.Sc., B.Sc. : Physics.

Dr. Dibya Prakash Rai

Post Doctoral Fellow (PDF), M.Sc., B.Sc., Working as Assistant Professor in Department of Physics in Pachhuncga University College, Mizoram, India.

Dr. Pankaj Kumar Pal

Ph.D R/S, ECE Deptt., IIT-Roorkee.

Dr. P. Thangam

PhD in Information & Communication Engineering, ME (CSE), BE (Computer Hardware & Software), currently serving as Associate Professor in the Department of Computer Science and Engineering of Coimbatore Institute of Engineering and Technology.

Dr. Pradeep K. Sharma

PhD., M.Phil, M.Sc, B.Sc, in Physics, MBA in System Management, Presently working as Provost and Associate Professor & Head of Department for Physics in University of Engineering & Management, Jaipur.

Dr. R. Devi Priya

Ph.D (CSE), Anna University Chennai in 2013, M.E, B.E (CSE) from Kongu Engineering College, currently working in the Department of Computer Science and Engineering in Kongu Engineering College, Tamil Nadu, India.

Dr. Sandeep

Post-doctoral fellow, Principal Investigator, Young Scientist Scheme Project (DST-SERB), Department of Physics, Mizoram University, Aizawl Mizoram, India- 796001.

Dr. Roberto Volpe

Faculty of Engineering and Architecture, Università degli Studi di Enna "Kore", Cittadella Universitaria, 94100 – Enna (IT).

Dr. S. Kannadhasan

Ph.D (Smart Antennas), M.E (Communication Systems), M.B.A (Human Resources).

Research Area: Engineering Physics, Electromagnetic Field Theory, Electronic Material and Processes, Wireless Communications.

Mr. Bhavinbhai G. Lakhani

An expert in Environmental Technology and Sustainability, with an M.S. from NYIT. Their specialization includes Construction Project Management and Green Building. Currently a Project Controls Specialist Lead at DACK Consulting Solutions, they manage project schedules, resolve delays, and handle claim negotiations. Prior roles as Senior Project Manager at FCS Group and Senior Project Engineer at KUNJ Construction Corp highlight their extensive experience in project estimation, resource management, and on-site supervision.

Mr. Omar Muhammed Neda

Department of Electrical Power Engineering, Sunni Diwan Endowment, Iraq.

Mr. Amit Kumar

Amit Kumar is associated as a Researcher with the Department of Computer Science, College of Information Science and Technology, Nanjing Forestry University, Nanjing, China since 2009. He is working as a State Representative (HP), Spoken Tutorial Project, IIT Bombay promoting and integrating ICT in Literacy through Free and Open Source Software under National Mission on Education through ICT (NMEICT) of MHRD, Govt. of India; in the state of Himachal Pradesh, India.

Mr. Tanvir Singh

Tanvir Singh is acting as Outreach Officer (Punjab and J&K) for MHRD Govt. of India Project: Spoken Tutorial - IIT Bombay fostering IT Literacy through Open Source Technology under National Mission on Education through ICT (NMEICT). He is also acting as Research Associate since 2010 with Nanjing Forestry University, Nanjing, Jiangsu, China in the field of Social and Environmental Sustainability.

Mr. Abilash

MTech in VLSI, BTech in Electronics & Telecommunication engineering through A.M.I.E.T.E from Central Electronics Engineering Research Institute (C.E.E.R.I) Pilani, Industrial Electronics from ATI-EPI Hyderabad, IEEE course in Mechatronics, CSHAM from Birla Institute Of Professional Studies.

Mr. Varun Shukla

M.Tech in ECE from RGPV (Awarded with silver Medal By President of India), Assistant Professor, Dept. of ECE, PSIT, Kanpur.

Mr. Shrikant Harle

Presently working as a Assistant Professor in Civil Engineering field of Prof. Ram Meghe College of Engineering and Management, Amravati. He was Senior Design Engineer (Larsen & Toubro Limited, India).

Mr. Zairi Ismael Rizman

Senior Lecturer, Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM) (Terengganu) Malaysia Master (Science) in Microelectronics (2005), Universiti Kebangsaan Malaysia (UKM), Malaysia. Bachelor (Hons.) and Diploma in Electrical Engineering (Communication) (2002), UiTM Shah Alam, Malaysia.

Mr. Ronak

Qualification: M.Tech. in Mechanical Engineering (CAD/CAM), B.E.

Presently working as a Assistant Professor in Mechanical Engineering in ITM Vocational University, Vadodara. Mr. Ronak also worked as Design Engineer at Finstern Engineering Private Limited, Makarpura, Vadodara.

Table of Contents

Volume-11, Issue-5, May 2025

Enhancing Financial Fraud Detection using XGBoost, LSTM, and KNN with SMOTE for Imbalanced Datasets

Godfred Antwi Koduah[1*]; Jinguo Lian[2]

Department of Mathematics and Statistics, University of Massachusetts Amherst
*Corresponding Author

Received: 08 May 2025/ Revised: 15 May 2025/ Accepted: 24 May 2025/ Published: 05-06-2025

Abstract— *The surge in digital financial activity has led to increasingly sophisticated forms of fraud, creating serious challenges for financial institutions. One of the core obstacles in fraud detection is the substantial class imbalance present in transactional datasets, where fraudulent records represent a small minority. This study presents a robust machine learning framework that integrates the Synthetic Minority Over-sampling Technique (SMOTE) with three distinct classifiers—XGBoost, Long Short-Term Memory (LSTM), and K-Nearest Neighbors (KNN)—to enhance the detection of fraudulent activities. Using a real-world dataset of six million banking transactions, we assess each model's performance through accuracy, precision, recall, F1-score, and both PR and ROC AUC metrics. Our findings show that SMOTE significantly boosts model recall and AUC scores. Among the models, XGBoost consistently delivers superior results with near-perfect metrics, while KNN maximizes recall, albeit at a slight cost to precision. LSTM produces more moderate but stable performance. Visual diagnostics, such as ROC/PR curves and confusion matrices, further confirm the reliability of XGBoost when combined with SMOTE. Overall, the integration of data balancing with advanced classifiers proves to be a powerful approach for real- time fraud detection.*

Keywords— Financial fraud detection, imbalanced datasets, machine learning, XGBoost, SMOTE.

I. INTRODUCTION

With the rapid evolution of digital banking, the financial industry faces a mounting threat from fraudulent transactions. Fraud not only leads to significant monetary losses but also undermines consumer confidence in online financial platforms [1]. According to a 2022 report by the Association of Certified Fraud Examiners (ACFE), global fraud resulted in losses exceeding $3.6 billion, underscoring the urgent need for effective detection systems [2].

A major challenge in identifying fraudulent behavior lies in the highly skewed nature of fraud datasets, where valid transactions vastly outnumber fraudulent ones. This skewness often leads to machine learning models performing poorly on minority classes, resulting in high false negative rates and overlooked fraud [3].

To improve detection, a wide range of techniques has been investigated—ranging from rule-based heuristics to deep learning systems. While machine learning excels at identifying complex, non-linear patterns in high-dimensional data, its effectiveness is hindered by class imbalance [4]. To mitigate this, methods like the Synthetic Minority Over-sampling Technique (SMOTE) have been employed to create a more balanced training distribution by synthesizing new examples from the minority class [5].

In this study, we propose a data-centric approach to financial fraud detection that integrates SMOTE with three different learning models: XGBoost, LSTM, and KNN. These models were selected based on their complementary strengths: XGBoost for structured data classification, LSTM for temporal pattern recognition in transaction sequences, and KNN for localized anomaly detection. Our contributions are summarized as follows:

- **Synthetic Oversampling:** We apply SMOTE to increase the representation of fraudulent cases in the training data, improving recall and reducing bias toward the majority class.
- **Model Diversity:** We implement and compare three distinct algorithms—XGBoost, LSTM, and KNN—each optimized for specific aspects of the data.
- **Comprehensive Analysis:** We evaluate all models using six key metrics and visualize results through confusion matrices, ROC/PR curves, and histograms to provide a thorough performance assessment.

The remainder of this paper is organized as follows: Section 2 provides a detailed description of the dataset used in our experiments, including its features and preprocessing steps. Section 3 presents some machine learning models and data balancing methods. Section 4 provides evaluation and performance comparison of some machine learning models. Finally, Section 5 concludes the paper and suggests the best model for detecting fraud based on the results in section 4.

II. DATA DESCRIPTIONS

This study utilizes the Bank Account Fraud dataset, which is composed of six distinct subsets, each containing one million records, leading to a cumulative dataset of six million transaction entries. These include a base dataset along with five variant datasets, labeled I through V.

The base dataset, along with variants I, II, and IV, contains 32 features per transaction. Variants III and V expand slightly, including 34 features. The attributes across these datasets span a wide range of transaction characteristics, including:

- Demographic indicators (e.g., customer age, income level)
- Behavioral features (e.g., transaction frequency, session duration)
- Risk-related metrics (e.g., credit risk rating, proposed credit limits)
- Transactional data points (e.g., number of bank branches used, days since account activity)

A defining trait of the dataset is its significant class imbalance. The overwhelming majority of transactions are legitimate (class 0), while only a small fraction are labeled as fraudulent (class 1). Figure 1 in the paper presents a histogram illustrating this disproportion, emphasizing the difficulty of training fraud detection models on such imbalanced data.

A major challenge in fraud detection research is the limited availability of large-scale, real-world datasets, particularly those focused on new bank account (NBA) fraud. The dataset used here is one of the few publicly accessible and comprehensive datasets in this space. As such, it serves as a valuable benchmark for developing and evaluating the performance of machine learning models intended for fraud detection in the financial domain.

TABLE 1
DATA DESCRIPTION OF THE BASE DATASET

count	mean	std	min	25%	50%	75%	max	
fraud bool	1000000	0.01	0.1	0	0	0	0	1
income	1000000	0.56	0.29	0.1	0.3	0.6	0.8	0.9
name_email_similarity	1000000	0.49	0.29	0	0.23	0.49	0.76	1
prev_address_months_count	1000000	16.72	44.05	-1	-1	-1	12	383
current_address_months_count	1000000	86.59	88.41	-1	19	52	130	428
customer_age	1000000	33.69	12.03	10	20	30	40	90
days_since_request	1000000	1.03	5.38	0	0.01	0.02	0.03	78.46
intended balcon amount	1000000	8.66	20.24	-15.53	-1.18	-0.83	4.98	112.96
zip_count_4w	1000000	1572.69	1005.37	1	894	1263	1944	6700
velocity_6h	1000000	5665.3	3009.38	-170.6	3436.37	5319.77	7680.72	16715.57
velocity_ 24h	1000000	4769.78	1479.21	1300.31	3593.18	4749.92	5752.57	9506.9
velocity 4w	1000000	4856.32	919.84	2825.75	4268.37	4913.44	5488.08	6994.76
bank_branch_count_8w	1000000	184.36	459.63	0	1	9	25	2385
date_of_birth_distinct_emails_4w	1000000	9.5	5.03	0	6	9	13	39
credit risk score	1000000	130.99	69.68	-170	83	122	178	389
email is free	1000000	0.53	0.5	0	0	1	1	1
phone_home_valid	1000000	0.42	0.49	0	0	0	1	1
phone mobile valid	1000000	0.89	0.31	0	1	1	1	1
bank months count	1000000	10.84	12.12	-1	-1	5	25	32
has other cards	1000000	0.22	0.42	0	0	0	0	1
proposed_credit_limit	1000000	515.85	487.56	190	200	200	500	2100
foreign request	1000000	0.03	0.16	0	0	0	0	1
session_length_in_minutes	1000000	7.54	8.03	-1	3.1	5.11	8.87	85.9
keep alive session	1000000	0.58	0.49	0	0	1	1	1
device_distinct_emails_8w	1000000	1.02	0.18	-1	1	1	1	2
device_fraud_count	1000000	0	0	0	0	0	0	0
month	1000000	3.29	2.21	0	1	3	5	7

TABLE 2
DATA DESCRIPTION OF THE VARIANT I DATASET

count	mean	std	min	25%	50%	75%	max	
fraud bool	1000000	0.01	0.1	0	0	0	0	1
income	1000000	0,56	0.29	0.1	0.3	0.6	0.8	0.9
name_email_similarity	1000000	0.49	0.29	0	0.23	0.49	0.76	1
prev_address_months_count	1000000	16.96	43.87	-1	-1	-1	15	399
current address months count	1000000	83.59	86.46	-1	18	50	124	429
customer age	1000000	31.97	10.9	10	20	30	40	90
days since request	1000000	1.05	5.46	0	0.01	0.02	0.03	76.64
intended balcon amount	1000000	8.72	20.21	-15.74	-1.18	-0.83	6.22	113.12
zip count 4w	1000000	1574.5	1003.7	1	893	1270	1952	6678
velocity_6h	1000000	5661.9	3010.9	-174.11	3431.2	5300	7692.3	16817.8
velocity_ 24h	1000000	4767.1	1481.6	1322.3	3587	4745.6	5753.2	9539.36
velocity_4w	1000000	4857.2	919.76	2855.2	4269.2	4913.8	5488.6	7019.2
bank _branch_count_8w	1000000	181.17	457.64	0	1	9	24	2386
date _of_birth_distinct_emails_4w	1000000	9.86	5	0	6	9	13	39
credit risk score	1000000	129.41	69.07	-191	82	121	176	388
email is free	1000000	0.53	0.5	0	0	1	1	1
phone home valid	1000000	0.4	0.49	0	0	0	1	1
phone mobile valid	1000000	0.9	0.3	0	1	1	1	1
bank months count	1000000	10.8	12.12	-1	-1	5	25	32
has other cards	1000000	0.22	0.41	0	0	0	0	1
proposed_credit_limit	1000000	507.16	481.46	190	200	200	500	2100
foreign request	1000000	0.03	0.16	0	0	0	0	1
session length in minutes	1000000	7.46	7.95	-1	3.09	5.08	8.76	85.57
keep alive session	1000000	0.58	0.49	0	0	1	1	1
device distinctemails 8w	1000000	1.02	0.18	-1	1	1	1	2
device fraud count	1000000	0	0	0	0	0	0	0
month	1000000	3.29	2.21	0	1	3	5	7

TABLE 3
DATA DESCRIPTION OF VARIANT II DATASET

Column 01	count	mean	std	min	25%	50%	75%	max
fraud_bool	1000000.0	0.01	0.1	0.0	0.0	0.0	0.0	1.0
income	1000000.0	0.57	0.29	0.1	0.3	0.6	0.8	0.9
name_email_similarity	1000000.0	0.49	0.29	0.0	0.21	0.49	0.75	1.0
prev_address_months_count	1000000.0	14.82	43.23	-1.0	-1.0	-1.0	-1.0	399.0
current_address_months_count	1000000.0	99.38	94.56	-1.0	26.0	64.0	154.0	429.0
customer_age	1000000.0	41.3	13.8	10.0	30.0	50.0	50.0	90.0
days since_request	1000000.0	0.91	4.99	0.0	0.01	0.02	0.03	76.58
intended_balcon_amount	1000000.0	8.64	20.57	-15.74	-1.18	-0.83	0.08	112.7
zip_count_4w	1000000.0	1567.4	1009.62	1.0	901.0	1236.0	1909.0	6650.0
velocity_6h	1000000.0	5685.1	3001.71	-174.11	3470.24	5408.43	7653.99	16801.34
velocity_24h	1000000.0	4787.41	1470.37	1322.33	3628.56	4765.97	5750.78	9539.36
velocity_4w	1000000.0	4860.39	916.81	2870.59	4271.19	4919.35	5489.47	7019.2
bank_branch_count_8w	1000000.0	202.46	474.13	0.0	1.0	10.0	32.0	2377.0
date of birth distinct emails_4w	1000000.0	7.95	4.96	0.0	4.0	7.0	11.0	39.0
credit risk_score	1000000.0	137.46	72.2	-191.0	87.0	128.0	187.0	388.0
email_is_free	1000000.0	0.52	0.5	0.0	0.0	1.0	1.0	1.0
phone_home_valid	1000000.0	0.49	0.5	0.0	0.0	0.0	1.0	1.0
phone_mobile_valid	1000000.0	0.86	0.35	0.0	1.0	1.0	1.0	1.0
bank_months_count	1000000.0	11.2	12.11	-1.0	1.0	6.0	25.0	32.0
has_other_cards	1000000.0	0.24	0.43	0.0	0.0	0.0	0.0	1.0
proposed credit_limit	1000000.0	558.75	513.85	190.0	200.0	200.0	1000.0	2100.0
foreign_request	1000000.0	0.02	0.16	0.0	0.0	0.0	0.0	1.0
session_length_in_minutes	1000000.0	7.91	8.34	-1.0	3.21	5.28	9.42	87.24
keep alive_session	1000000.0	0.56	0.5	0.0	0.0	1.0	1.0	1.0
device_distinct_emails_8w	1000000.0	1.02	0.2	-1.0	1.0	1.0	1.0	2.0
device_fraud count	1000000.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
month	1000000.0	3.29	2.21	0.0	1.0	3.0	5.0	7.0

TABLE 4
DATA DESCRIPTION OF VARIANT III DATASET

Column 01	count	mean	std	min	25%	50%	75%	max
fraud_bool	1000000.0	0.01	0.1	0.0	0.0	0.0	0.0	1.0
income	1000000.0	0.58	0.29	0.1	0.3	0.6	0.8	0.9
name_email_similarity	1000000.0	0.49	0.29	0.0	0.21	0.49	0.75	1.0
prev_address_months_count	1000000.0	14.67	43.02	-1.0	-1.0	-1.0	-1.0	399.0
current_address_months_count	1000000.0	99.23	94.07	-1.0	27.0	64.0	154.0	429.0
customer_age	1000000.0	41.34	13.77	10.0	30.0	50.0	50.0	90.0
days_since_request	1000000.0	0.9	5.01	0.0	0.01	0.02	0.03	76.58
intended balcon_amount	1000000.0	8.55	20.52	-15.74	-1.18	-0.83	-0.07	112.7
zip_count_4w	1000000.0	1517.66	965.03	1.0	886.0	1208.0	1844.0	6650.0
velocity_6h	1000000.0	5489.73	2940.94	-174.11	3332.99	5188.16	7367.06	16754.96
velocity 24h	1000000.0	4660.88	1451.48	1322.33	3503.01	4640.4	5591.86	9539.36
velocity 4w	1000000.0	4733.57	871.23	2870.59	4238.23	4813.0	5331.57	7019.2
bank_branch_count_8w	1000000.0	201.15	473.59	0.0	1.0	10.0	31.0	2377.0
date of birth distinct emails 4w	1000000.0	7.77	4.82	0.0	4.0	7.0	11.0	39.0
credit risk score	1000000.0	139.29	71.43	-177.0	90.0	130.0	188.0	388.0
email_is_free	1000000.0	0.52	0.5	0.0	0.0	1.0	1.0	1.0
phone_home_valid	1000000.0	0.49	0.5	0.0	0.0	0.0	1.0	1.0
phone mobile valid	1000000.0	0.86	0.35	0.0	1.0	1.0	1.0	1.0
bank months_count	1000000.0	11.14	12.13	-1.0	1.0	6.0	25.0	32.0
has_other_cards	1000000.0	0.25	0.43	0.0	0.0	0.0	0.0	1.0
proposed_credit_limit	1000000.0	551.69	506.66	190.0	200.0	200.0	1000.0	2100.0
foreign_request	1000000.0	0.02	0.15	0.0	0.0	0.0	0.0	1.0
session_length_in_minutes	1000000.0	7.81	8.23	-1.0	3.15	5.25	9.37	85.57
keep alive session	1000000.0	0.56	0.5	0.0	0.0	1.0	1.0	1.0
device_distinct emails_8w	1000000.0	1.02	0.19	-1.0	1.0	1.0	1.0	2.0
device_fraud_count	1000000.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
month	1000000.0	3.66	2.12	0.0	2.0	4.0	5.0	7.0
x1	1000000.0	0.01	1.01	-4.98	-0.67	0.01	0.69	6.43
x2	1000000.0	0.01	1.01	-4.85	-0.67	0.01	0.68	6.54

TABLE 5
DATA DESCRIPTION OF VARIANT IV DATASET

	count	mean	std	min	25%	50%	75%	max
fraud_bool	1000000.00	0.01	0.10	0.00	0.00	0.00	0.00	1.00
income	1000000.00	0.58	0.29	0.10	0.30	0.60	0.80	0.90
name_email_similarity	1000000.00	0.49	0.29	0.00	0.21	0.49	0.75	1.00
prev_address_months_count	1000000.00	14.68	43.01	-1.00	-1.00	-1.00	-1.00	399.00
current_address_months count	1000000.00	99.21	94.08	-1.00	27.00	64.00	154.00	429.00
customer_age	1000000.00	41.34	13.78	10.00	30.00	50.00	50.00	90.00
days_since_request	1000000.00	0.90	5.01	0.00	0.01	0.02	0.03	76.58
intended_balcon_amount	1000000.00	8.55	20.52	-15.74	-1.18	8	-0.07	112.70
zip_count_4w	1000000.00	1517.55	964.96	1.00	886.00	1208.00	1844.00	6650.00
velocity_6h	1000000.00	5489.69	2940.44	-174.11	3333.59	5188.38	7366.62	16754.96
velocity_24h	1000000.00	4660.86	1451.39	1322.33	3502.92	4640.63	5591.77	9539.36
velocity_4w	1000000.00	4733.55	871.21	2870.59	4238.22	4813.07	5331.50	7019.20
bank branch count 8w	1000000.00	201.00	473.48	0.00	1.00	10.00	31.00	2377.00
date_of_birth_distinct_emails_4w	1000000.00	7.78	4.82	0.00	4.00	7.00	11.00	39.00
credit_risk_score	1000000.00	139.30	71.45	177.00	90.00	130.00	188.00	388.00
email is free	1000000.00	0.52	0.50	0.00	0.00	1.00	1.00	1.00
phone_home_valid	1000000.00	0.49	0.50	0.00	0.00	0.00	1.00	1.00
phone_mobile_valid	1000000.00	0.86	0.35	0.00	1.00	1.00	1.00	1.00
bank_months_count	1000000.00	11.14	12.13	-1.00	1.00	6.00	25.00	32.00
has_other_cards	1000000.00	0.25	0.43	0.00	0.00	0.00	0.00	1.00
proposed_credit_limit	1000000.00	551.73	506.71	190.00	200.00	200.00	1000.00	2100.00
foreign_request	1000000.00	0.02	0.15	0.00	0.00	0.00	0.00	1.00
session_length_in_minutes	1000000.00	7.81	8.23	-1.00	3.15	5.25	9.37	87.24
keep_alive_session	1000000.00	0.56	0.50	0.00	0.00	1.00	1.00	1.00
device distinct emails 8w	1000000.00	1.02	0.19	-1.00	1.00	1.00	1.00	2.00
device_fraud_count	1000000.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
month	1000000.00	3.66	2.12	0.00	2.00	4.00	5.00	7.00

TABLE 6
DATA DESCRIPTION OF VARIANT V DATASET

	count	mean	std	min	25%	50%	75%	max
fraud_bool	1000000.00	0.01	0.10	0.00	0.00	0.00	0.00	1.00
income	1000000.00	0.58	0.29	0.10	0.30	0.60	0.80	0.90
name_email_similarity	1000000.00	0.49	0.29	0.00	0.21	0.49	0.75	1.00
prev_address_months_count	1000000.00	14.74	43.13	-1.00	-1.00	-1.00	-1.00	384.00
current address months count	1000000.00	99.19	94.07	-1.00	27.00	64.00	154.00	426.00
customer_age	1000000.00	41.35	13.75	10.00	30.00	50.00	50.00	90.00
days_since_request	1000000.00	0.92	5.07	0.00	0.01	0.02	0.03	77.85
intended_balcon_amount	1000000.00	8.57	20.54	-15.71	-1.18	-0.83	-5	113.09
zip_count_4w	1000000.00	1517.47	965.95	1.00	885.00	1208.00	1846.00	6830.00
velocity_6h	1000000.00	5490.94	2940.12	-143.65	3332.98	5190.72	7371.56	16802.05
velocity_24h	1000000.00	4661.53	1450.44	1297.72	3505.06	4641.57	5593.34	9585.10
velocity_4w	1000000.00	4733.51	870.56	2858.75	4238.38	4813.95	5331.50	7019.20
bank_branch_count_8w	1000000.00	201.08	473.74	0.00	1.00	10.00	31.00	2404.00
date_of_birth_distinct_emails_4w	1000000.00	7.77	4.82	0.00	4.00	7.00	11.00	37.00
credit_risk_score	1000000.00	139.30	71.43	-177.00	90.00	130.00	188.00	388.00
email_is_free	1000000.00	0.52	0.50	0.00	0.00	1.00	1.00	1.00
phone_home_valid	1000000.00	0.49	0.50	0.00	0.00	0.00	1.00	1.00
phone_mobile_valid	1000000.00	0.86	0.35	0.00	1.00	1.00	1.00	1.00
bank_months_count	1000000.00	11.14	12.12	-1.00	1.00	6.00	25.00	32.00
has_other_cards	1000000.00	0.25	0.43	0.00	0.00	0.00	0.00	1.00
proposed_credit_limit	1000000.00	551.04	506.53	190.00	200.00	200.00	1000.00	2100.00
foreign_request	1000000.00	0.02	0.15	0.00	0.00	0.00	0.00	1.00
session_length_in_minutes	1000000.00	7.82	8.26	-1.00	3.15	5.25	9.36	83.21
keep_alive_session	1000000.00	0.56	0.50	0.00	0.00	1.00	1.00	1.00
device_distinct_emails_8w	1000000.00	1.02	0.19	-1.00	1.00	1.00	1.00	2.00
device_fraud_count	1000000.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
month	1000000.00	3.66	2.12	0.00	2.00	4.00	5.00	7.00
X1	1000000.00	0.01	1.01	-4.98	-0.67	0.01	0.68	6.43
x2	1000000.00	0.01	1.01	-4.85	-0.67	0.00	0.68	6.54

The histogram below shows the fraud class distribution of the Bank Account Fraud dataset.

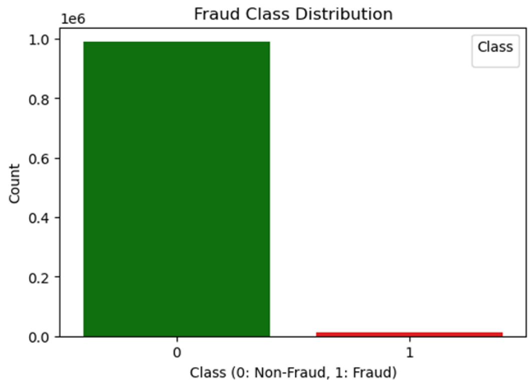


FIGURE 1: Fraud Class Distribution

III. MACHINE LEARNING MODELS AND DATA BALANCING METHODS

3.1 Machine Learning Models:

This section outlines the core models employed for fraud detection—XGBoost, LSTM, and KNN—along with the data balancing technique SMOTE used to mitigate class imbalance. Each model was selected based on its strengths in dealing with structured data, temporal dependencies, and anomaly patterns.

3.1.1 XGBoost:

XGBoost (Extreme Gradient Boosting) is a highly efficient implementation of the gradient boosting algorithm, known for its scalability and predictive accuracy. It constructs an ensemble of decision trees, where each successive tree is trained to correct the errors of the previous ones. This additive training approach minimizes a loss function and optimizes model performance over time.

Key features of XGBoost include:

- Built-in regularization to prevent overfitting
- Native support for handling missing values
- High compatibility with structured/tabular data

Due to its robustness and precision, XGBoost is commonly adopted in fraud detection tasks, especially where speed and accuracy are critical. Important tuning parameters include the number of trees (estimators), learning rate, and tree depth.

3.1.2 LSTM:

Long Short-Term Memory (LSTM) networks are a specialized form of Recurrent Neural Networks (RNNs) designed to learn long-term dependencies in sequential data. LSTMs incorporate gated memory cells that control the flow of information, allowing the network to retain or forget past data as needed.

This architecture makes LSTM ideal for analyzing time-series data, such as transaction histories, where order and temporal patterns are essential. In fraud detection, LSTMs help capture subtle behavioral sequences that could indicate anomalous activity.

Critical hyperparameters include:

- Number of LSTMunits
- Number of hidden layers
- Dropout rate for regularization

3.1.3 KNN:

The K-Nearest Neighbors (KNN) algorithm is a non-parametric, instance-based learning method used for classification and regression. It classifies a data point based on the majority label of its ’k’ closest neighbors in the feature space.

KNNis particularly useful in scenarios with imbalanced data, as it can detect localized anomalies—clusters of fraudulent transactions that deviate from the norm. It also requires no training phase, making it computationally simple, though costly at prediction time.

Key factors influencing KNN performance include:

- The value of k (set to 5 in this study)
- The choice of distance metric (e.g., Euclidean or Manhattan)
- Weighting schemes for neighbours (uniform or distance-based)

3.2 Data Balancing with SMOTE:

Due to the severe imbalance in the dataset, standard training processes would result in models heavily biased toward the dominant (non-fraud) class. To correct this, we apply the Synthetic Minority Over-sampling Technique (SMOTE).

SMOTE addresses class imbalance by generating new synthetic instances of the minority class. Rather than duplicating existing samples, it interpolates between neighboring minority instances to create plausible new examples. This method enhances the model’s exposure to fraudulent behavior during training, improving its ability to generalize and detect rare events.

3.3 Integration of SMOTE with Machine Learning Models:

Before model training, SMOTE is applied to the training set to ensure a balanced representation of both classes. This preprocessing step ensures that the models—XGBoost, LSTM, and KNN—learn from a more equitable sample distribution.

As shown in later sections, applying SMOTE results in significant improvements across multiple performance metrics, particularly recall, F1 score, and AUC values, all of which are critical for identifying fraud cases.

3.4 Summary:

This section introduced the core components of our fraud detection framework: three machine learning models—XGBoost, LSTM, and KNN—and the SMOTE oversampling method for addressing class imbalance. The combination of these tools serves as the foundation for the evaluation and analysis described in the next section.

IV. MODEL EVALUATION AND PERFORMANCE COMPARISON:

This section presents a comparative analysis of the three selected machine learning models—XGBoost, LSTM and KNN—evaluated both with and without SMOTE. The assessment is based on six key performance metrics. In addition, we include visual tools such as ROC and PR curves, histograms, and confusion matrices to provide deeper insight into each model's behavior.

4.1 Evaluation Metrics:

To quantify the models' performance, we employ the following evaluation criteria:

4.1.1 Accuracy:

Accuracy measures the ratio of total predictions that are correct. It is computed as:

$$\textbf{Accuracy} = \frac{\text{True Positives (TP)} + \text{True Negatives (TN)}}{\text{TP} + \text{TN} + \text{False Positives (FP)} + \text{False Negatives (FN)}} \tag{1}$$

However, in highly skewed datasets, accuracy can be misleading as it may reflect the dominance of the majority class rather than true performance on the minority (fraud) class.

4.1.2 Precision:

Precision measures the proportion of predicted fraud cases that are actually fraudulent:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{2}$$

High precision is desirable when false positives are costly, such as in fraud investigations that demand manual follow-up.

4.1.3 F1 Score:

The F1 score is the harmonic mean of precision and recall, offering a balanced view of a model's performance:

$$F1\ Score = \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

In fraud detection, where both missed fraud (FN) and false alarms (FP) carry consequences, F1 is an important metric.

4.1.4 Recall:

Recall, captures the proportion of actual fraud cases correctly identified:

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

A high recall is particularly important in fraud detection, as missing fraudulent cases can lead to significant financial and reputational losses.

4.1.5 PR AUC:

The Precision-Recall Area under the Curve (PR AUC) evaluates the trade-off between precision and recall across different thresholds. It is particularly useful in skewed datasets where traditional ROC curves might be less informative.

$$PR\ AUC = \int Precision(r)\ dr \tag{5}$$

Where r represents recall. A perfect model achieves a PR AUC of 1, while random classifiers typically score close to the class prevalence.

4.1.6 ROC AUC:

The Receiver Operating Characteristic Area under the Curve (ROC AUC) shows the model's ability to differentiate between the positive (fraud) and negative (non-fraud) classes at various thresholds. A higher ROC AUC indicates better classification performance.

4.2 Performance Comparison with and without SMOTE:

Table 7 summarizes the models' performance with and without the application of SMOTE. As observed, using SMOTE leads to substantial gains in recall, F1 score, PR AUC, and ROC AUC for all models.

TABLE 7
PERFORMANCE METRICS FOR XGBOOST, LSTM, AND KNN WITH AND WITHOUT SMOTE

Model	Accuracy	Precision	F1 Score	Recall	PR AUC	ROC AUC
XGBoost (without SMOTE)	0.9916	0.8127	0.4237	0.2865	0.55	0.64
XGBoost (with SMOTE)	0.9928	0.9943	0.9891	0.984	1	1
LSTM (without SMOTE)	0.9917	0.8627	0.4099	0.2688	0.57	0.63
LSTM (with SMOTE)	0.8949	0.7676	0.8617	0.9822	0.95	0.98
KNN (without SMOTE)	0.9905	0.9371	0.2201	0.1247	0.54	0.56
KNN (with SMOTE)	0.9489	0.8671	0.9288	1	0.98	0.99

KNN experiences the most dramatic improvement in recall and F1 score, indicating its increased ability to detect fraud once the dataset is balanced. XGBoost, when paired with SMOTE, achieves perfect AUC values (both ROC and PR) and maintains strong precision and recall. LSTM also benefits from SMOTE, particularly in recall, though its overall performance remains more moderate.

4.3 Visualizing Model Performance with and without SMOTE:

To better understand how each model behaves, we provide several forms of visualization.

4.3.1 Histograms:

Figures 2 and 3 display the distribution of evaluation metrics before and after SMOTE is applied. These visualizations make it evident that SMOTE leads to a notable boost in recall, F1 score, and AUC values, while having minimal impact on accuracy and precision.

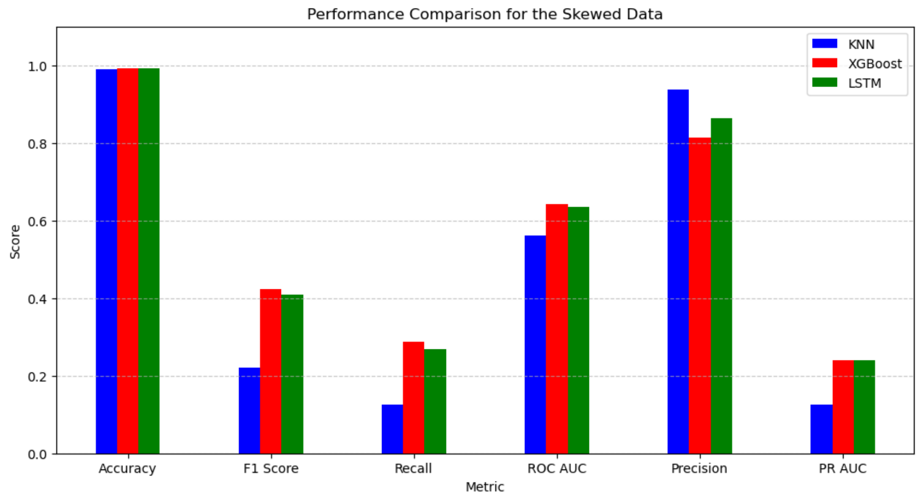


FIGURE 2: Distribution without SMOTE

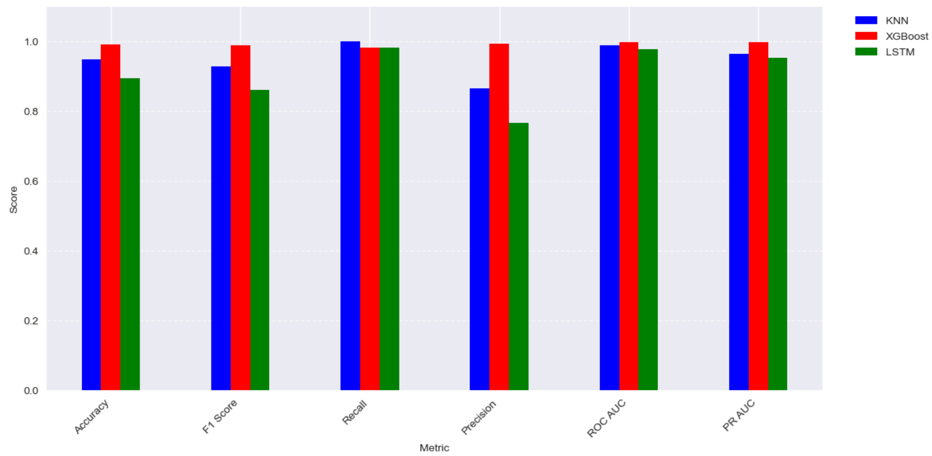


FIGURE 3: Distribution with SMOTE.

4.3.2 ROC Curves:

Figures 4 and 5 illustrate ROC curves for all three models under both conditions (with and without SMOTE). The curves confirm that balancing the dataset enhances the models' ability to differentiate between fraudulent and legitimate transactions, as reflected in higher AUC scores.

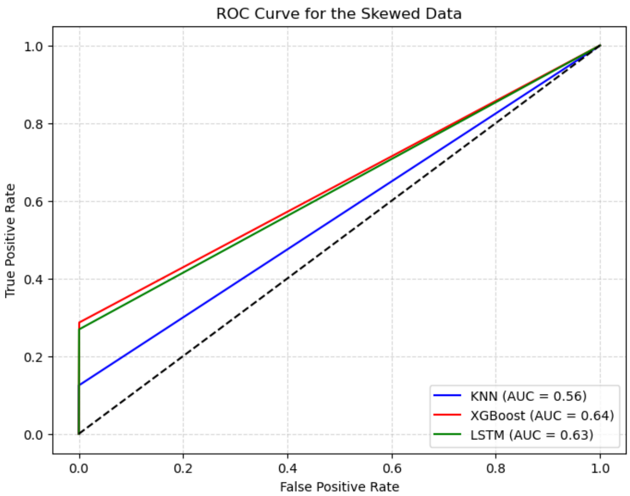


FIGURE 4: ROC curves without SMOTE

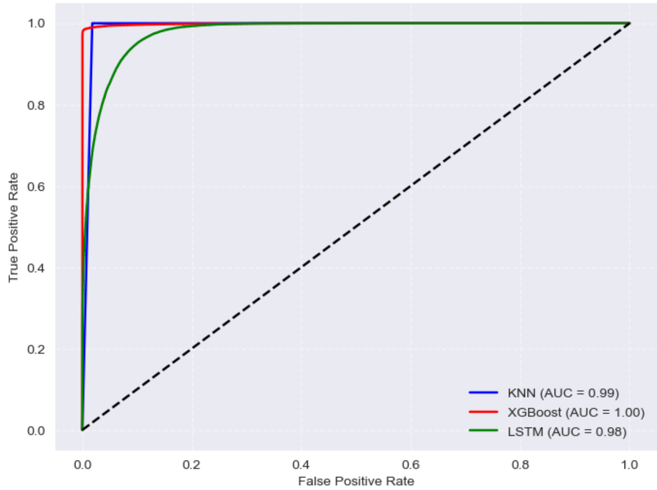


FIGURE 5: ROC curves with SMOTE

4.3.3 PR Curves:

Figures 6 and 7 present the Precision-Recall curves. These plots provide insight into the trade-offs models make as decision thresholds shift. Post-SMOTE results demonstrate improved recall while maintaining precision, particularly for XGBoost and LSTM.

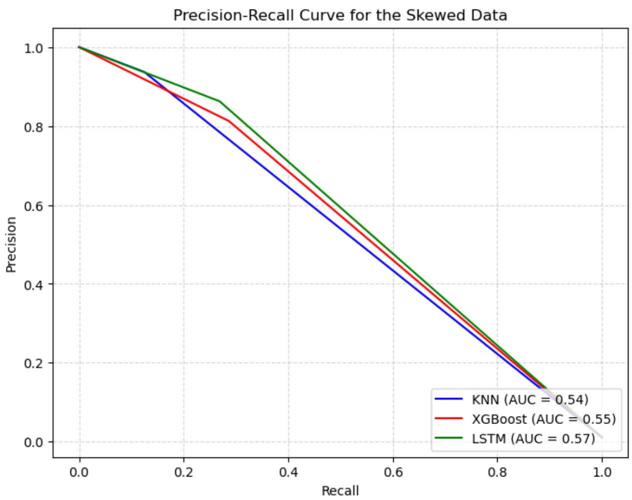


FIGURE 6: PR curves without SMOTE

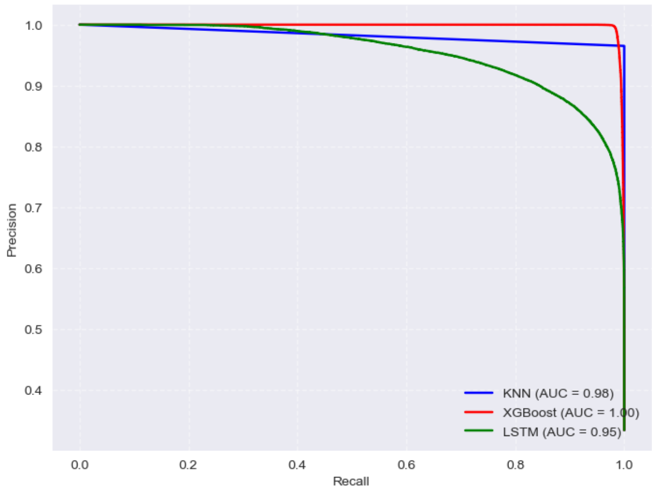


FIGURE 7: PR curves with SMOTE

4.3.4 Confusion Matrices:

Figures 8 and 9 show confusion matrices, offering a granular view of model classification results. After applying SMOTE, there is a visible reduction in false negatives (FN) across all models, especially in KNN, which shifts to identifying nearly all fraudulent transactions.

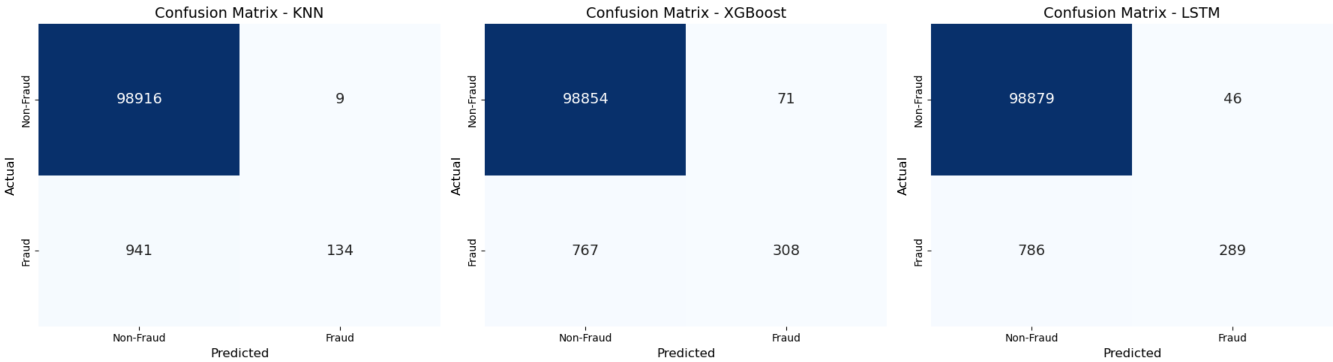


FIGURE 8: Confusion Matrix without SMOTE

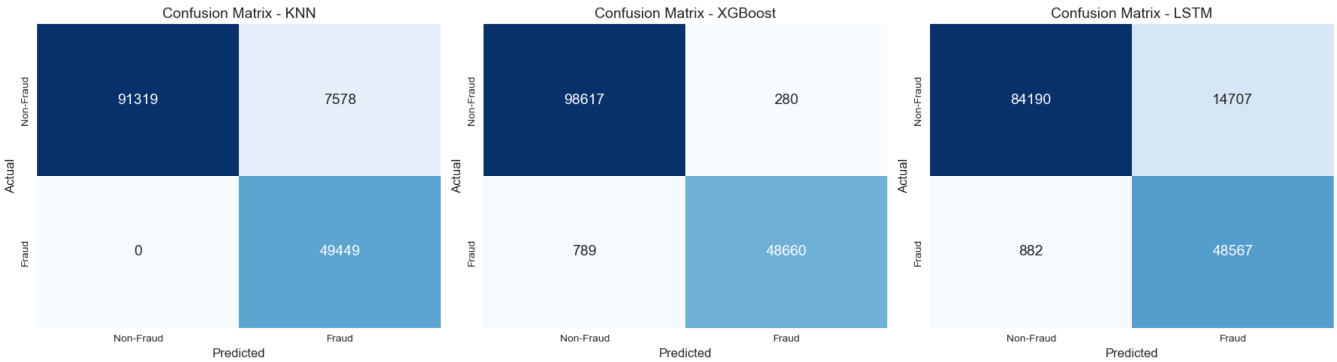


FIGURE 9: Confusion Matrix with SMOTE

4.4 Summary

The experiments confirm that SMOTE is an effective technique for improving fraud detection in imbalanced datasets. All three models benefit from its use, particularly in terms of recall and F1 score. XGBoost remains the top-performing model, delivering high precision and near-perfect AUC scores. KNN, while achieving perfect recall, sacrifices precision and suffers from higher false positives. LSTM shows a good balance but does not match XGBoost in overall performance.

V. CONCLUSION

This study explored the application of three machine learning models—XGBoost, KNN, and LSTM—for detecting fraudulent financial transactions, with a particular focus on addressing class imbalance through the use of SMOTE.

Our findings reveal clear differences in model behavior both before and after applying SMOTE. Without balancing, all models—especially KNN and LSTM—struggled with recall due to the scarcity of fraudulent instances in the training data. Among the unbalanced results, XGBoost stood out with the fewest false negatives, but still suffered from reduced sensitivity overall.

The introduction of SMOTE significantly improved each model's ability to detect fraudulent transactions. Recall increased across the board, most notably for KNN (from 0.1247 to 1.0000) and LSTM (from0.2688to0.9822), as evidenced by confusion matrices. However, these gains were accompanied by a rise in false positives, particularly for KNN, reflecting the classic precision-recall trade-off that arises in imbalanced classification problems.

Despite these trade-offs, XGBoost with SMOTE consistently emerged as the best-performing model, achieving outstanding results across all key metrics. It reached perfect PR AUC and ROC AUC scores (1.00) and maintained a strong balance between precision (0.9943) and recall (0.9840). While KNN achieved flawless recall, it did so at the cost of precision (0.8671) and overall F1 score stability. LSTM demonstrated considerable improvement post-SMOTE, but still lagged behind in precision and balanced performance.

These results were further reinforced by visual tools such as ROC and PR curves, histograms, and confusion matrices. The consistency of XGBoost's dominance across all metrics and visual diagnostics makes it a robust and scalable solution for real-world fraud detection systems, especially when augmented with data balancing techniques like SMOTE.

SOURCE OF DATA

The Bank Account Fraud data used in this research is publicly available at https://github.com/feedzai/bank-account-fraud.

REFERENCES

[1] T. Ashfaq et al., "A machine learning and blockchain based efficient fraud detection mechanism," *Sensors*, vol. 22, no. 19, p. 7162, Sep. 2022.

[2] ACFE, "Association of Certified Fraud Examiners (ACFE) 2022 Report to the Nations," 2022. [Online]. Avail- able: https://legacy.acfe.com/report-to-the-nations/2022/[Accessed: 2023]; E. Eber- lein et al., "Mathematics in Financial Risk Management," *Research Gate*, vol. 4, no. 1, pp. 1–26, 2007.

[3] N. S. Alfaiz and S. M. Fati, "Enhanced credit card fraud detection model using machine learning," *Electronics*, vol. 11, no. 4, p. 662, 2022.

[4] P. Vanini et al., "Online payment fraud: From anomaly detection to risk management," *Financial Innovation*, vol. 9, no. 1, p. 66, Mar. 2023, doi: 10.1186/s40854-023-00470-w.

[5] D. Gorton, "Modeling fraud prevention of online services using incident response trees and value at risk," in *Proc. 10th Int. Conf. Availability, Reliability and Security*, Toulouse, France, Aug. 2015, pp. 149–158, doi: 10.1109/ARES.2015.17.

Electric Vehicles in India: A Boon for Transportation or a Challenge for Consumer

Preet Gill[1*]; Dr. Kavita Dua[2]

[*1]PhD Scholar, Department of Resource Management and Consumer Science I.C. College of Community Science, CCS HAU HISAR (Haryana)

[2]Associate Professor, Department of Resource Management and Consumer Science I.C. College of Community Science, CCS HAU HISAR (Haryana)

*Corresponding Author

Received: 11 May 2025/ Revised: 19 May 2025/ Accepted: 27 May 2025/ Published: 05-06-2025

Abstract— *India, like many nations, faces the dual challenge of increasing transportation demands and growing environmental concerns. In this context, electric vehicles (EVs) have emerged as a potential game-changer, promising a cleaner and more sustainable alternative to traditional fossil fuel-powered vehicles. This study explores perceptions and experiences of 30 electric vehicle (EV) users, focusing on socio-demographics, usage patterns, challenges and satisfaction. The sample was predominantly male (83.3%), with a balanced age distribution and a high proportion of private-sector employees. Most respondents owned electric two-wheelers, auto-rickshaws and cars, with over 40% using EVs for more than three years. Cost savings on fuel emerged as the primary motivation for EV adoption. While most users reported home charging as convenient, issues such as high initial costs, limited fast-charging infrastructure and range anxiety were noted. Respondents generally recognized EVs' environmental benefits and lower running costs. A majority expressed satisfaction with their EV experience and recommended them to others. The findings highlight the importance of enhancing affordability, technological reliability and policy support to accelerate EV adoption in India.*

Keywords— Electric Vehicles, Technology, Environment, Cost Effective, Government Policies.

I. INTRODUCTION

India, like several other countries, is grappling with rising transportation needs alongside escalating environmental issues. In this scenario, electric vehicles (EVs) have surfaced as a promising solution, offering a cleaner and more sustainable alternative to conventional fossil fuel driven transportation. With advancements in technology and growing public awareness, EVs are gaining momentum as a viable option for future mobility. Their adoption is further supported by government initiatives, financial incentives, and the increasing need to reduce carbon emissions and dependence on fossil fuels. EVs offer a multitude of potential benefits for India. They hold the promise of reducing air pollution in congested urban centers, decreasing the nation's reliance on imported oil and mitigating the effects of climate change.

1.1 Electric Vehicle Types:

1.1.1 Vehicles powered by batteries (BEVs):

Battery electric vehicles are full electric cars without an exhaust pipe, fuel storage or a petrol or diesel engine that run solely on electricity. They are also referred to as "plug-in electric vehicles (PEVs)" since they use an external charging outlet to charge the battery. BEVs come in a variety of forms, including electric vehicles, trucks many more.

1.1.2 Hybrid electric vehicle:

A hybrid electric vehicle improves fuel economy and operates at peak efficiency by producing far fewer emissions than a pure gasoline-powered vehicle. Additionally, plug in hybrid vehicles (PHEVs) exist. Even so, they make less noise than fully hybrid cars.

1.2 India's Battery Technology:

1.2.1 Lead-Acid Battery:

Most automakers choose to employ lead-acid battery technology in their cars because of its low cost and great efficiency, which lowers the vehicle's overall cost and increases customer profitability. Lead-acid batteries composition includes a number of environmentally hazardous chemical compounds. Researchers created novel batteries such as nickel-cadmium, nickel-metal hydride, sulphur-containing lithium, air containing lithium, lithium-ion, etc. to get around these drawbacks. However, scientists began employing.

1.2.2 Lithium-ion Battery:

Rechargeable batteries are made of lithium-ion material. Most electric vehicles in India and other nations utilise it because of its enormous capacity. Due to the high cost and weight of lithium-ion batteries, the price of electric vehicles rises.

1.2.3 Fuel Cell:

A fuel cell is an electrical device made up of an anode and a cathode. Fuel and oxide make up the majority of the fuel cell, which uses a redox reaction to produce energy. Fuel cells were employed by Korea for their e-buses, which had a positive effect on the auto sector. India also introduced a fuel cell-powered electric bus in 2018; it debuted in New Delhi.

1.3 Policy Initiatives:

1.3.1 FAME India Scheme: April 1, 2015

The Faster Adoption and Manufacturing of Hybrid and Electric Vehicles (FAME) scheme provides financial incentives for EV buyers and manufacturers.

1.3.2 Battery Waste Management Rules, 2022

Mandates Extended Producer Responsibility (EPR), requiring manufacturers to take back used batteries and ensure their safe disposal or recycling.

1.3.3 PM E-Drive (Electric Drive Revolution in Innovative Vehicle Enhancement):

Launched in October 2024, this ₹10,900 crore scheme offers upfront purchase subsidies for electric two-wheelers, three-wheelers, buses, trucks, and ambulances. It also supports the development of EV charging infrastructure and aims to replace polluting vehicles with electric alternatives.

1.4 Disposal of lithium Iron batteries in India:

Over 90% of India's battery recycling occurs in the informal sector, using rudimentary techniques like open burning and acid leaching. These practices are harmful to workers' health and the environment. Companies like Attero Recycling and Lohum Cleantech are setting up advanced facilities for LIB recycling. These companies use environmentally friendly processes to extract valuable metals and ensure safe disposal of waste.

In Hisar, Haryana, notable electric vehicle brands include Zelio (Eeva ZX, Eeva, Gracy, Gracy i, Gracy Pro, Speed X), Elesco (Aqua, Big Show, Broch, Classic, Craze, Eroyal, Eroyal Pro, Etorq, Thunder), and YUKIE (specific model names not listed). In the electric three-wheeler segment, SKYRIDE is active, though specific model names are not detailed. While there are no electric four-wheeler manufacturers directly in Hisar, Haryana is home to brands like Mahindra Electric (e-Verito, e-Supro) and Okinawa (Ridge+, iPraise+, Praise Pro, Lite, R30, Okhi-90, Ridge100, Dual100).

II. MATERIAL AND METHODOLOGY

2.1 Locale of study:

For the study, one district from Haryana state were selected randomly for carrying out the research objectives.

2.2 Sample procedure:

30 people were purposively selected for the study who own electric vehicles from selected district. Results Table 1: Demographics Summary.

III. RESULTS

TABLE 1
DEMOGRAPHICS SUMMARY

Variable	Frequency (n)	Percentage (%)
Gender		
Male	25	83.30%
Female	5	16.70%
Age Group		
Below 25 years	8	26.70%
25-40 years	12	40%
Above 40	10	33.30%
Occupation		
Student	6	20%
Private Sector	17	56.60%
Government Employee	7	23.30%
Household Income		
Below ₹40,000	5	16.70%
₹40,000-₹80,000	12	40%
Above ₹80,000	13	43.30%
Type of EV Owned		
Electric Auto-rickshaw	10	33.30%
Electric Bike/Scooter	10	33.30%
Electric Car	10	33.30%

In table 1 showed that the study sample consisted of 30 respondents, predominantly male (83.3%) with a smaller proportion of females (16.7%). The age distribution was fairly balanced, with 26.7% below 25 years, 40% between 25–40 years, and 33.3% above 40 years. In terms of occupation, the majority were employed in the private sector (56.6%), followed by government employees (23.3%) and students (20%). Most respondents belonged to medium- to high-income households, with 40% earning between ₹40,000–₹80,000 and 43.3% earning above ₹80,000 per month. Electric vehicle ownership was evenly split, with 33.3% each owning an electric auto-rickshaw, an electric bike/scooter, and an electric car, providing a diverse perspective on EV usage across different vehicle types.

TABLE 2
EV OWNERSHIP AND USAGE

Variable	Frequency (n)	Percentage (%)
Duration of EV Ownership		
Less than 1 year	8	26.70%
1 year – 3 years	10	33.30%
More than 3 years	12	40.00%
Primary Reason for Purchase		
Environmental Concerns	9	30.00%
Cost Savings on Fuel	12	40.00%
Government Subsidy	6	20.00%
New Technology Enthusiast	3	10.00%
Frequency of Driving		
Daily	18	60.00%
2-3 times a week	10	33.30%
Occasionally (weekends)	2	6.70%
Charging Location		
Home Charging	25	83.30%
Public Charging Stations	2	6.60%
Workplace Charging	3	10.00%

Table 2 showed that the duration of EV ownership, the majority (40%) had been using their EVs for more than 3 years, followed by 33.3% who owned them for 1–3 years, and the lowest (26.7%) had less than 1 year of ownership. As for the primary reason for EV purchase, cost savings on fuel was the leading factor (40%), followed by environmental concerns (30%), government subsidies (20%), and the lowest (10%) were motivated by interest in new technology. In terms of driving frequency, most respondents (60%) drove their EVs daily, followed by 33.3% who drove 2–3 times a week, and the least (6.7%) used them occasionally on weekends. Among the 30 respondents, the majority (83.3%) reported charging their electric vehicles at home, followed by 10% who used workplace charging, and the lowest, 6.6%, relied on public charging stations.

TABLE 3
SUSTAINABLE TRANSPORTATION AND ENVIRONMENTAL BENEFITS

Statement	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)	WMS
EVs reduce air pollution in India	2 (6.7%)	3 (10.0%)	6 (20%)	12 (40%)	7 (23.3%)	3.6 (III)
Lower carbon footprint with EV use	1 (3.3%)	1 (3.3%)	20 (66.6%)	3 (10.0%)	5 (16.7%)	3.3 (VI)
EVs combat climate change effectively	1 (3.3%)	4 (13.3%)	8 (26.7%)	11 (36.7%)	6 (20.0%)	3.5 (IV)
EVs reduce fossil fuel dependency	2 (6.7%)	1 (3.3%)	7 (23.3%)	14 (46.7%)	6 (20.0%)	3.7 (II)
EVs reduce noise pollution	1 (3.3%)	3 (10.0%)	4 (13.3%)	13 (43.3%)	9 (30.0%)	3.8 (I)
Government promotes EV adoption adequately	5 (16.7%)	7 (23.3%)	9 (3%)	6 (20.0%)	3 (10.0%)	2.8 (VII)
EVs help achieve India's sustainability goals	2 (6.7%)	4 (13.3%)	8 (26.7%)	12 (40.0%)	4 (13.3%)	3.4 (V)

Table 3 indicate generally positive perceptions toward electric vehicles (EVs) in terms of environmental benefits, though opinions vary across specific aspects. A majority (40%) agreed that EVs reduce air pollution in India, with 23.3% strongly agreeing, while only a small portion (6.7%) strongly disagreed. On the statement that EVs help reduce fossil fuel dependency, 46.7% agreed and 20% strongly agreed, making it one of the most supported views, with minimal disagreement. Similarly, the idea that EVs reduce noise pollution was supported by 43.3% agreeing and 30% strongly agreeing. However, the statement regarding EVs lowering carbon footprints received the highest neutrality (66.6%), suggesting uncertainty or lack of awareness, though 16.7% did strongly agree and 10.0% were agree. Regarding the role of EVs in combating climate change, 36.7% agreed and 20% strongly agreed, with limited disagreement. The statement on EVs supporting India's sustainability goals saw 40% agreement and 13.3% strong agreement, with 26.7% remaining neutral. In contrast, the government's role in promoting EV adoption was viewed more critically—only 20% agreed and 10% strongly agreed, while 40% either disagreed or strongly disagreed, indicating public skepticism about policy effectiveness.

TABLE 4
CHALLENGES IN EV ADOPTION

Statement	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)	WMS
High initial EV cost	2 (6.7%)	4 (13.3%)	5 (16.7%)	12 (40%)	7 (23.3%)	3.6 (I)
Government subsidies make EVs affordable	3 (10.0%)	4 (13.3%)	13 (43.3%)	7 (23.3%)	3 (10.0%)	3.1(IV)
Difficulty finding suitable EV models	2 (6.7%)	3 (10%)	8 (26.7%)	11 (36.7%)	6 (20.0%)	3.5 (II)
Uncertain resale value of EVs	4 (13.3%)	5 (16.7%)	6 (20%)	10 (33.3%)	5 (16.7%)	3.2 (III)
Inadequate EV charging infrastructure	1 (3.3%)	3 (10.0%)	13 (43.3%)	7 (23.3%)	6 (20.0%)	3.5 (II)
Difficulty finding fast- charging stations	2 (6.7%)	3 (10.0%)	15 (50.0%)	6 (20.0%)	4 (13.3%)	3.2 (III)
Home charging facilities are convenient	2 (6.7%)	4 (13.3%)	6 (20.0%)	12 (40.0%)	6 (20.0%)	3.5 (II)

Table 4 reveal varying levels of agreement regarding the challenges and conveniences associated with electric vehicles (EVs). For the statement on high initial EV cost, the majority (40%) agreed, followed by 23.3% who strongly agreed, while the lowest (6.7%) strongly disagreed, indicating that cost is widely seen as a significant barrier. On whether government subsidies make EVs affordable, 43.3% remained neutral—the highest response—while the lowest (10%) strongly agreed or strongly disagreed, showing mixed sentiments. Regarding difficulty in finding suitable EV models, the majority (36.7%) agreed, followed by 26.7% who were neutral, while the lowest (6.7%) strongly disagreed. On the uncertainty of EV resale value, the highest proportion (33.3%) agreed, followed by 20% who were neutral, while the lowest (13.3%) strongly disagreed. Concerning inadequate charging infrastructure, 43.3% remained neutral—the majority—followed by 23.3% who agreed, and the least (3.3%) strongly disagreed, suggesting uncertainty or mixed experiences. For difficulty in finding fast-charging stations, 50% were neutral, the majority, followed by 20% who agreed, and the lowest (6.7%) strongly disagreed. Lastly, when asked about the convenience of home charging, the majority (40%) agreed, 20% strongly agreed, and the lowest (6.7%) strongly disagreed, reflecting overall positive sentiment toward home-based charging facilities.

TABLE 5
INFRASTRUCTURE AND CHARGING CHALLENGES

Statement	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)	WMS
Experience range anxiety (fear of running out of charge) while driving an EV	3 (10.0%)	5 (16.7%)	6 (20.0%)	10 (33.3%)	6 (20.0%)	3.3 (II)
EV charging time is too long	2 (6.7%)	4 (13.3%)	7 (23.3%)	7 (23.3%)	11 (36.7%)	3.6 (I)
Charging costs are reasonable compared to fuel costs	4 (13.3%)	5 (16.7%)	6 (20.0%)	10 (33.3%)	5 (16.7%)	3.2 (III)
Lack of charging stations discourages long-distance travel	3 (10.0%)	4 (13.3%)	9 (30.0%)	8 (26.7%)	6 (20.0%)	3.3 (II)
EV maintenance costs are lower than fuel-based vehicles	2 (6.7%)	4 (13.3%)	5 (16.7%)	12 (40.0%)	7 (23.3%)	3.6 (I)
Battery replacement costs are a major concern for EV owners	4 (13.3%)	5 (16.7%)	8 (26.7%)	10 (33.3%)	3 (10.0%)	3.1 (IV)

Table 5 reflect diverse user experiences and concerns regarding the practicality and economics of electric vehicles (EVs). On the issue of range anxiety, the majority (33.3%) agreed they experience it, followed by 20% who were neutral or strongly agreed, while the lowest (10%) strongly disagreed. Regarding the statement that EV charging time is too long, the majority (36.7%) strongly agreed, followed by 23.3% who were neutral, and the lowest (6.7%) strongly disagreed, indicating that slow charging remains a concern for many users. On whether charging costs are reasonable compared to fuel costs, 33.3% agreed—the highest—followed by 20% neutral and 13.3% strongly disagreeing, showing that while many find charging economical, some remain unconvinced. Concerning the lack of charging stations discouraging long-distance travel, 30% remained neutral—the majority—followed by 26.7% agreeing, and the lowest (10%) strongly disagreed, reflecting moderate concern. On EV maintenance costs being lower than fuel-based vehicles, 40% agreed—the highest—followed by 23.3% strongly agreeing, while the lowest (6.7%) strongly disagreed, suggesting positive perceptions about maintenance savings. Finally, regarding battery replacement costs being a major concern, 33.3% agreed—the majority— followed by 26.7% neutral, and the lowest (10%) strongly agreed, showing it remains a financial worry for many owners.

TABLE 6
TECHNOLOGICAL AND PERFORMANCE CONCERNS

Statement	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)	WMS
EV performance is satisfactory in extreme weather conditions (hot/cold)	4 (13.3%)	5 (16.7%)	9 (30.0%)	7 (23.3%)	5 (16.7%)	3.1 (IV)
Lack of trained mechanics for EV repairs is a problem	3 (10.0%)	5 (16.7%)	4 (13.3%)	14 (46.6%)	4 (13.3%)	3.3 (II)
I face issues with EV software updates and battery management	4 (13.3%)	5 (16.7%)	7 (23.3%)	9 (30.0%)	5 (16.7%)	3.2 (III)
EVs offer a comfortable and smooth driving experience	2 (6.7%)	6 (20.0%)	12 (40%)	4 (13.3%)	6 (20%)	3.2 (III)
I have experienced a significant reduction in running costs after switching to an EV	3 (10%)	5 (16.7%)	7 (23.3%)	11 (36.7%)	4 (13.3%)	3.2 (III)
The government should increase investment in EV charging infrastructure	2 (6.7%)	4 (13.3%)	5 (16.7%)	12 (40%)	7 (23.3%)	3.6 (I)

Table 6 highlight mixed perceptions regarding electric vehicle (EV) performance, support infrastructure and overall experience. For the statement on EV performance in extreme weather conditions, the majority (30%) remained neutral, followed by 23.3% who agreed, while the lowest (13.3%) strongly disagreed, indicating uncertainty or variable experiences. Regarding the trained mechanics for EV repairs, the highest proportion (46.6%) agreed. On issues with EV software updates and battery management, 30% agreed the majority followed by 23.3% who were neutral, and the lowest (13.3%) strongly disagreed. In terms of driving experience, 40% remained neutral the highest response followed by 20% who strongly agreed, while only 6.7% strongly disagreed, showing general satisfaction with EV driving comfort. For reduction in running costs after switching to EVs, 36.7% agreed the majority followed by 23.3% who were neutral, and the lowest (10%) strongly disagreed, indicating a largely positive cost-saving perception. Lastly, on whether the government should increase investment in EV charging infrastructure, the majority (40%) agreed, followed by 23.3% who strongly agreed, while the lowest (6.7%) strongly disagreed, highlighting strong public support for infrastructure development.

TABLE 7
POLICY AND AWARENESS ISSUES

Statement	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)	WMS
Public awareness campaigns about EV benefits should be improved	2 (6.7%)	4 (13.3%)	5 (16.7%)	12 (40%)	7 (23.3%)	3.6 (III)
Financial incentives (subsidies, tax benefits) are crucial for increasing EV adoption	3 (1%)	5 (16.7%)	6 (20%)	6 (20%)	10 (33.3%)	3.5 (IV)
India is ready for a complete transition to electric mobility	4 (13.3%)	6 (20%)	8 (26.7%)	9 (30%)	3 (10%)	3.0 (VII)
The resale market for EVs needs improvement to make them more attractive	3 (10%)	5 (16.7%)	6 (20%)	11 (36.7%)	5 (16.7%)	3.3 (VI)
The government should introduce stricter policies to phase out petrol/diesel vehicles	2 (6.7%)	4 (13.3%)	10 (33.3%)	7 (23.3%)	7 (23.3%)	3.4 (V)
I would recommend an EV to my friends and family	1 (3.3%)	3 (1%)	5 (16.7%)	12 (40.0%)	9 (30.0%)	3.8 (I)
More automakers should invest in affordable EV models	2 (6.7%)	4 (13.3%)	6 (20.0%)	7 (23.3%)	11 (36.7%)	3.7 (II)
Despite challenges, I am satisfied with my decision to own an EV	3 (10%)	4 (13.3%)	6 (20.0%)	12 (40.0%)	5 (16.7%)	3.4 (V)

Table 7 reflect strong support for electric vehicle (EV) promotion, affordability, and user satisfaction, despite some challenges. For the statement that public awareness campaigns about EV benefits should be improved, the majority (40%) agreed, followed by 23.3% who strongly agreed, while the lowest (6.7%) strongly disagreed. Regarding the importance of financial incentives for EV adoption, 33.3% strongly agreed the highest followed by equal proportions (20%) who were neutral or agreed, and the lowest (10%) strongly disagreed. On whether India is ready for a complete transition to electric mobility, the majority (30%) agreed, followed by 26.7% who were neutral, and the lowest (10%) strongly agreed, indicating a cautious optimism. Concerning the resale market for EVs, 36.7% agreed the majority followed by 20% neutral, and 10% strongly disagreed, suggesting room for improvement in resale confidence. For stricter government policies to phase out petrol/diesel vehicles, agreement and strong agreement were equal at 23.3%, while the majority (33.3%) remained neutral and the lowest (6.7%) strongly disagreed. On recommending EVs to others, 40% agreed and 30% strongly agreed the highest while the lowest (3.3%) strongly disagreed, showing overall user satisfaction. Regarding investment by automakers in affordable EVs, the majority (36.7%) strongly agreed, followed by 23.3% who agreed, and the lowest (6.7%) strongly disagreed. Finally, despite challenges, 40% agreed and 16.7% strongly agreed that they are satisfied with their decision to own an EV, while only 10% strongly disagreed, highlighting positive ownership experiences overall.

IV. CONCLUSION

The study reveals a generally positive perception and growing acceptance of electric vehicles (EVs) among respondents, despite certain concerns. Environmental benefits such as reduced air and noise pollution, lower fossil fuel dependency, and alignment with sustainability goals received strong support. However, skepticism remains regarding the government's role in EV promotion and policy enforcement. Key challenges identified include high initial costs, inadequate charging infrastructure, slow charging times, and uncertain resale value, although home charging is viewed positively. Most respondents cited fuel cost savings as the primary motivator for EV purchase, with many using their EVs daily and charging at home. While range anxiety and battery replacement costs remain concerns, maintenance costs and driving experience are generally viewed favorably. Mixed responses were noted about software updates and performance in extreme weather. Respondents strongly endorsed the need for increased government investment in charging infrastructure, public awareness, and affordability through subsidies. Overall, while barriers persist, user satisfaction and willingness to recommend EVs reflect a favorable shift in public attitude, signaling the potential for broader EV adoption in India with the right policy and infrastructural support.

RECOMMENDATION

- Government should continue and expand financial incentives (subsidies, tax breaks) to make EVs more accessible.
- Develop mechanisms or policies that can stabilize and improve the resale value of EVs, potentially through battery health standards or buy-back programs.

REFERENCES

[1] Bhalla, P., Ali, I. S., & Nazneen, A. (2018). A study of consumer perception and purchase intention of electric vehicles. *European Journal of Scientific Research*, *149*(4), 362-368.

[2] Beena. J. J., Rakesh S. (2020). Present and future trends for electric vehicles in India. Journal-case Studies, 3(1), -Special.

[3] Monika, Ms, and A. Mifzala. "A study on customer perception towards e-vehicles in Bangalore." *J Emer Tech Innov Res* 6 (2019): 87-92.

[4] Rajper, S. Z., & Albrecht, J. (2020). Prospects of electric vehicles in the developing countries: A literature review. *Sustainability*, *12*(5), 1906.

[5] Kumar, R., Jha, A., Damodaran, A., Bangwal, D., & Dwivedi, A. (2020). Addressing the challenges to electric vehicle adoption via sharing economy: An Indian perspective. *Management of Environmental Quality: An International Journal*, *32*(1), 82- 99.

[6] Dhar, S., Pathak, M., & Shukla, P. R. (2017). Electric vehicles and India's low carbon passenger transport: a long-term co-benefits assessment. *Journal of Cleaner Production*, *146*, 139-148.

[7] Preetha, P. K., & Poornachandran, P. (2019, February). Electric vehicle scenario in India: roadmap, challenges and opportunities. In *2019 IEEE international conference on electrical, computer and communication technologies (ICECCT)* (pp. 1-7). IEEE.

[8] Chidambaram, K., Ashok, B., Vignesh, R., Deepak, C., Ramesh, R., Narendhra, T. M., & Kavitha, C. (2023). Critical analysis on the implementation barriers and consumer perception toward future electric mobility. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, *237*(4), 622-654.

[9] Ankita Nagpal. (2020). Consumers' perception towards electric vehicles in India. Psychology and Education, 57(9), 4043-4050.

[10] Kishore, S., & Johnvieira, A. (2021). Shared electric mobility: a catalyst for EV adoption in India. *Contemporary issues in business, management, and society*, 125.

[11] Gupta, S., Bansal, R., Bankoti, N., Kar, S. K., Mishra, S. K., Kaur, P., & Harichandan, S. (2024). Factors affecting consumer's intention to use electric vehicles: Mediating role of awareness and knowledge. *Journal of Advanced Transportation*, *2024*(1), 5922430.

AD Publications
Sector-3, MP Nagar, Bikaner,
Rajasthan, India
www.adpublications.org, info@adpublications.org