ISSN : 2455-9679
SJIF Impact Factor : 6.190

IJRTSM

INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY SCIENCE & MANAGEMENT

"FRAUD IDENTIFICATION IN DIGITAL PAYMENT SYSTEMS BASED ON MACHINE LEARNING TECHNIQUES IN FINTECH"

Anath Bandhu Chatterjee [1]

[1] *Staff Software Engineer, PayPal Inc*

ABSTRACT

The fast increase of online payment systems has facilitated the growth in complexity and frequency of financial fraud, which makes safe processing of transactions a significant issue. The conventional processes cannot be effective in dealing with large-scale, skewed and dynamically moving data. This paper introduces a machine learning (ML) system to detect fraud in online payment platforms in FinTech industry. A representative sample of 200,000 records was used to build a predictive model that was effective based on a large-scale transactional dataset on Kaggle. Data preprocessing methods integrated into the methodology are IQR-based Winsorization to deal with outliers and label encoding to encode categorical variables. The addition of feature engineering with new features, such is featuring selection using Recursive Feature Elimination (RFE) and balancing adjustment features. Class imbalance was addressed using SMOTE, and data was split between training and test sets (80:20) with StandardScaler feature scaling. Extra Trees and XGBoost were used as ensemble models to perform classification. Compared to existing models, Extra Trees model had an accuracy of 98.02 which outperforms the XGBoost model with an accuracy of 96.81, which also showed better results on the experiment. These results illustration that ensemble learning techniques are effective in enhancing fraud detection performance and indicate the strength of the proposed framework for real-money management.

Keywords: Fraud Detection, Machine Learning, Digital Payment Systems, Synthetic Financial Datasets, Predictive Analytics

I. INTRODUCTION

Financial technology (FinTech) is a rapidly evolving area of digital payment systems, which provides the possibility to conduct quick, convenient and cashless payments [1]. As the internet has taken over online banking, mobile wallets, and online shopping, millions of online transactions are made every second worldwide [2]. The increasing adoption of digital payment methods has completely changed how people and companies handle financial transactions [3]. Online digital payment frauds are usually associated with unauthorized transactions, identity theft, or account information manipulation, with the aim of illegally obtaining financial benefits [4][5]. Scams, such as fraudulent transactions, account takeovers, and identity theft, are changing and becoming more complicated, undermining the utility of conventional protection systems [6]. Thus, development of sophisticated methods of detecting and preventing fraud on a real-time basis is direly needed [7].

Usually, the fraud detection techniques relying on predefined rules and carrying out manual checks become incapable when it comes to handling large amounts of transactional data processing and identifying novel patterns of frauds [8]. This has led to the intense demand for smart and responsive solutions that are able to learn and identify the concealed

ISSN : 2455-9679
SJIF Impact Factor : 6.190

trends of fraudulent activity in real-time [9][10].

In response to these challenges, machine learning approaches have become an effective tool to detect fraud in FinTech applications [11]. These methods have access to a huge amount of transactional data, which they exploit to find hidden patterns and anomalies in order to detect fraud with higher accuracy and even in real time [12][13]. In addition, algorithms like Random Forest, Extra Trees and XGBoost that can help model complex relationships that exist within financial data have demonstrated that they can increase detection levels. Furthermore, once fraud detection systems become empowered with machine learning processes, not only is it possible to enhance security and minimize financial risks, but also to build confidence in the modern FinTech digital payment context. The main key contributions of this work:

- Based on a real-life digital payment dataset and efficient preprocessing methods including outlier treatment, label encoding and feature scaling to improve data quality.
- Overcoming the issue of the imbalance in classes through use of SMOTE and, thus, improving ability of model to detect fraud cases within the minority group.
- Performed feature engineering and feature selection to identify and only keep the best features to identify fraud.
- Designed and tested several ML algorithms, including Extra Trees and XGBoost, to obtain more accurate prediction results.
- Provided far greater detection rates and accuracy of fraud, thus, creating a feasible solution to secure digital payment systems in the FinTech sector.

This paper is structured in the following way: Section II provides a thorough literature review. Section III describes proposed work. In Section IV, experimental setup and results of proposed model are provided. Lastly, the paper provides a conclusion and prospective work in Section V.0

II. Literature Review

The section evaluates the latest research contributions and methodology in the detection of frauds and more specifically, the detection systems based on ML and AI approaches which can optimize the performance of the detection, increase its scalability and more adequately address the conditions of a real-world application.

Recent studies emphasize key progress in fraud detection with using ML and AI, paying attention to the performance, scalability, and application to real-life. Sariat et al. (2025) propose a well-developed fraud detection system using ML that performs excellently (AUPRC 0.9998). Nevertheless, synthetic PaySim data usage limits the applicability of system to real-world and its testing in field [14]. Ingle et al. (2025) proposed a hybrid AI model which achieves AUC = 0.991 and F1 = 0.97. The questions about scalability and implementation of a real-time system are however not adequately discussed by the authors [15]. Bhutta and Mehmood (2025) outperform quantum models (F1 ≈ 1.000 vs SVM 0.9739) in detecting superior fraud, but are not feasible due to high computational cost (≈ 600×) [16]. Chakraborty et al. (2024) Chakraborty et al. (2024) introduce an integrated AML-blockchain system that leads to 95.8% performance and 27% robustness increase; however, scalability comments (1200 TPS, 3. 2s latency) must be proved in practice [17]. Deepa et al. (2024) use XGBoost on SMOTE and autoencoders, yet the duration of the dataset and the inability to generalize are constrained by the limited time of application and absence of real-time validation [18]. Rani and Mittal (2023) present a thorough literature review, but lack of quantitative measures and experimental replication restricts its analytical quality [19].

Nevertheless, in spite of the great improvements, there are still some research gaps in fraud detection systems. Most studies use syntactically produced or constrained datasets, which limit generalization and robustness in real-world scenarios. Real-time deployment and scalability tests are not fully validated. In addition, issues of data imbalance, concept drift, and appearance of new types of fraud are not adequately discussed in the literature. The fact that slight attention is paid to explainability, privacy, and security aspects also discourages the practical introduction. Other than

this, performance measures have been the center stage in most studies, whereas statistical validation has been barely touched, thereby leaving a gap between test outcomes and practice.

Table I provides a comparative overview of newer ML- and AI-based fraud detection methods, including the methods employed, datasets, main findings, and limitations. It provides a brief review of available methods, which helps identify gaps to be addressed in research and justify the usefulness of the proposed model.

Table 1: Comparative Analysis of Recent Machine Learning and Artificial Intelligence-Based Fraud Detection Technique

Ref.	Author(s) & Year	Method/Model Used	Dataset	Key Results	Limitations
[14]	Sariat et al. (2025)	XGBoost, Logistic Regression, Random Forest	PaySim (Synthetic)	AUPRC = 0.9998	Uses synthetic data; lacks real-world validation
[15]	Ingle et al. (2025)	Hybrid (ANN + LSTM + RF + XGBoost)	500K transactions	AUC = 0.991, F1 = 0.97	Scalability and deployment challenges not addressed
[16]	Bhutta & Mehmood (2025)	SVM, QSVM, VQC, Hybrid Quantum Models	Synthetic balanced dataset	F1 ≈ 1.000 (Quantum), 0.9739 (SVM)	High computational cost (~600×); simulated environment
[17]	Chakraborty et al. (2024)	AML + Blockchain	Real-world insurance data	95.8% performance, 27% robustness improvement	Needs real-world scalability validation
[18]	Deepa et al. (2024)	XGBoost + SMOTE + Autoencoders	Kaggle (2-day dataset)	~98% performance	Limited dataset; no real-time validation
[19]	Rani & Mittal (2023)	Review (AI-based methods)	Secondary data (2010–2023)	Identifies trends & challenges	No quantitative metrics or experimental validation

III. METHODOLOGY

The proposed framework (see in Fig. 1) is based on applying an ML pipeline to identify fraud in online payments. The initial step is preprocessing the dataset, including outlier treatment and label encoding. Feature engineering is performed, along with feature selection using RFE. After SMOTE is used to address a class imbalance, the data is divided into training and testing sets (80:20). StandardScaler is used for feature scaling, and classification is performed using the performance of ensemble models, including Extra Trees and XGBoost, which is evaluated using acc, prec, rec, and F1.

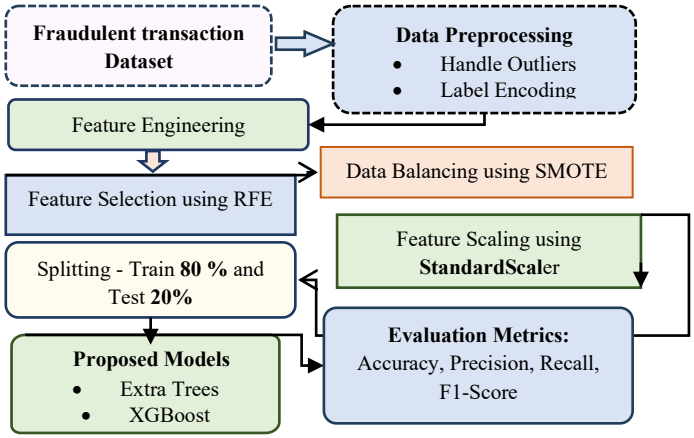

Fig. 1. Proposed Machine Learning-Based Fraud Detection Framework for Digital Payment System

ISSN : 2455-9679
SJIF Impact Factor : 6.190

A. Dataset Collection

Dataset collection is the process of gathering relevant data required to train and evaluate models. Millions of records with variables including transaction type, amount, and account balances were gathered from Kaggle repository to produce a massive fraudulent transaction dataset for this study. The initial data has around 6.3 million transaction records and 11 attributes, such as the type of transaction, transaction amount, account balances and fraud designations. A representative sample of 200,000 records was chosen for its size to develop and analyze an efficient model. This data is commonly used in other research on financial fraud detection to test ML models to determine the fraudulent nature of transactions.

B. Data Preprocessing and analysis

Data preparation is the process of cleaning and converting raw data into a format suitable for analysis. The following data preprocessing techniques applied on this dataset are:

1. Outlier Handling**:** The handling of outliers is critical to establishing that the extreme values used in model training or the estimation of statistical estimates are not misrepresented. Utilizing the Interquartile Range (IQR) as the basis for applying Winsorization to the balance attribute preserves the overall data integrity while simultaneously reducing the effect of extreme cases on the balance values [20]. Therefore, by preserving both the valid and extreme observations, the overall integrity of the dataset remains intact through the application of the Winsorization of the balance attribute by means of the IQR method.
2. Label Encoding: Categorical variables (e.g., type, nameOrig, and nameDest) were converted to numbers through label encoding so that the machine learning algorithms could process categorical variables while maintaining their significance.

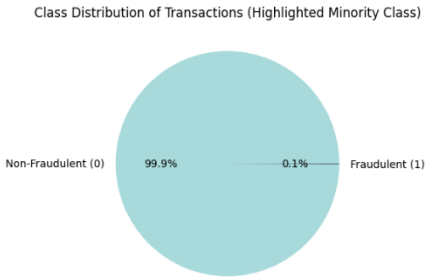

Fig.2. Class Distribution of Fraudulent and Non-Fraudulent Transactions

In addition, dataset represented in Fig. 2 is extremely biased, with 99.9% of the records belonging to non-fraudulent and only 0.1% being fraudulent. According to Fig. 3, there are notable differences in the distributions of amounts for fraudulent and non-fraudulent transactions.

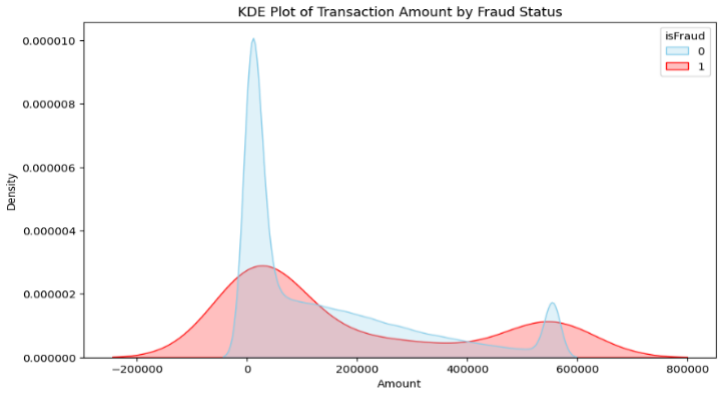

Fig. 3. Kernel Density Estimation (KDE) of Transaction Amount by Fraud Status

ISSN : 2455-9679
SJIF Impact Factor : 6.190

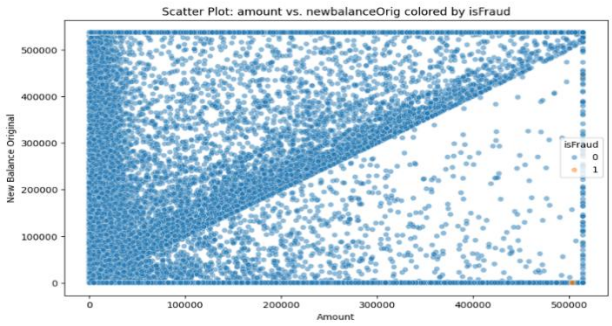

Fig. 4. Scatter Plot of Transaction Amount vs. New Balance Origin Colored by Fraud Status

There is also a clear deviation in how the transaction amount relates to the new balance between fraudulent and non-fraudulent transactions **(Fig. 4).** The deviated nature of the frauds versus non-frauds provides evidence of abnormal behavior and helps identify fraud for detection purposes.

C. Feature Engineering

To enhance model performance, developing new, important variables from existing data is known as feature engineering. In the current research report, the introduction of the two features - balance_change_orig and balance_change_dest - was intended to create new accounts and to help the model better understand how there are variations in account balances when a transaction occurs; as a result, a more accurate model can be created.

D. Feature Selection using RFE

The process of determining which variables are most crucial to model prediction is known as feature selection. A feature selection was utilized in RFE to identify those features that are most essential for detecting fraudulent behavior. The list of features that were identified as important includes step, amount, oldbalanceOrg, newbalanceOrig, oldbalanceDest, newbalanceDest, type_encoded, nameDest_encoded, balance_change_orig, and balance_change_dest, which hopefully improve the model's precision in identifying fraudulent transactions.

E. Data Balancing Using Synthetic Minority Over-Sampling Technique (SMOTE)

The problem of uneven class distribution is addressed by data balancing. In this study, SMOTE [21] was used to create synthetic samples in order to overcome problem of imbalance between classes and improve model's precision in identifying fraudulent transactions. Since fraudulent transactions are rare, SMOTE was applied to generate synthetic minority class data, guaranteeing that the model successfully picks up patterns from both classes.

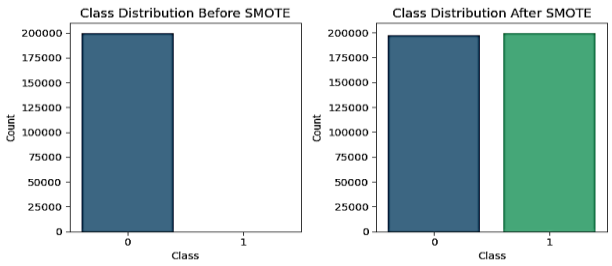

Fig. 5. Class Distribution Before and After Apply SMOTE for Data Balancing

Fig. 5 illustrates the distribution of classes before and after using SMOTE, which greatly increases minority class to provide a balanced dataset.

F. Feature Scaling:

The range of independent variables is standardized by feature scaling. Using StandardScaler, a dataset's mean is subtracted, and result is divided by standard deviation [22]. This procedure scales the data to have a standard deviation of 1 and centers it around 0. The StandardScaler formula is as follows, Equation (1).

ISSN : 2455-9679
SJIF Impact Factor : 6.190

$$z = \frac{x-\mu}{\sigma} \quad (1)$$

Where μ is the mean value of the feature, σ is the standard deviation, x is the original value, and z is the scaled value.

G. Splitting - Train 80 % and Test 20%

In the work, the dataset was split 80:20 between the training and testing halves. To make it reproducible, a random state of 42 was utilized. The split allows the model to train on training data and test its performance on unseen data, enabling credible predictions.

H. Proposed Machine Learning Model for Fraud Detection in Digital Payment Systems

Model building entails the decision and training of ML algorithms. The performance of the Extra Trees (ET) and XGBoost models in suggested fraud detection system is thoroughly examined in this section. It highlights the significance of critical steps in pre-processing data, data imbalance, feature engineering to improve the quality of prediction, and strength and resilience of ML in detecting fraud in a digital payment system.

1. Extra Trees Model

Extra Trees is an ensemble learning method that builds a large number of DT that are randomly chosen with regard to features and split points that ensure the reduction of overfitting and diversity [23]. The last prediction is derived by majority voting among all the trees. To optimize the model in this research, n_estimators = 50 and max depth = 6 were used to obtain better results.

2. Extreme Gradient Boosting (XGBoost) Model

An effective ensemble learning method called XGBoost creates DT one after the other to correct the mistakes of earlier models [24][25]. It also includes regularization to prevent overfitting and enhance generalization, which is effective in the detection of fraud. XGBoost was used in this study with parameters including estimates =50, max_depth =2 and learning_rate = 0.3 to enhance model performance and accurately capture complex patterns in transaction data.

I. Evaluation Metrics

The effectiveness of proposed ML models was measured in terms of important classification measures such as ROC-AUC, F1-score, recall, accuracy, precision, and confusion matrix [26]. These measures evaluate model with respect to correctness of fraud and non-fraud transactions (Equation 2–5).

$$Accuracy = \frac{TP+TN}{TP+Fp+TN+FN} \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$F1 = 2 * \frac{(precision*recall)}{precision+recall} \quad (5)$$

Accuracy denotes the general correctness, ***precision*** is percentage of the correct predictions of the frauds, ***recall*** is the capacity to identify the real frauds, and ***F1-score*** is a ratio between prec and rec. Furthermore, ROC-AUC is used to evaluate a model's ability to distinguish across classes, whereas the confusion matrix shows the detailed classification results in form of TP, TN, FP, and FN.

IV. Result Analysis

A. Experimental Setup

The experiments were performed with Python 3.10 in Google Colab and Jupyter Notebook using standard computing resources. Important packages such as scikit-learn, pandas, NumPy, Matplotlib, Seaborn, and imbalanced-learn were

ISSN : 2455-9679
SJIF Impact Factor : 6.190

used to preprocess, visualize, and model the data. The fraud detection framework was used and evaluated with ML models such as ET and XGBoost.

B. Performance Analysis of ML Models

Table II is a quantitative comparative analysis of Extra Trees and XGBoost in detecting fraud based on common classification measures, including F1, rec, acc, and prec. The ET model performs better than others, showing a well-balanced trade-off between FP and FN, with an acc of 98.02%, prec of 97.56%, rec of 0.9853, and an F1 of 98.04. Conversely, XGBoost performs somewhat worse overall, with an F1 of 0.9685, a rec of 0.9733, an acc of 96.81%, and a prec of 96.37%. These findings imply that Extra Trees classifier is more generalizable and robust at detecting fraud in the experimental environment.

Table 2 : Performance Comparison of Proposed Machine Learning Models for Fraud Detection

Models	Accuracy	Precision	Recall	F1-Score
Extra Trees	98.02	97.56	0.9853	98.04
XGBoost	96.81	96.37	0.9733	0.9685

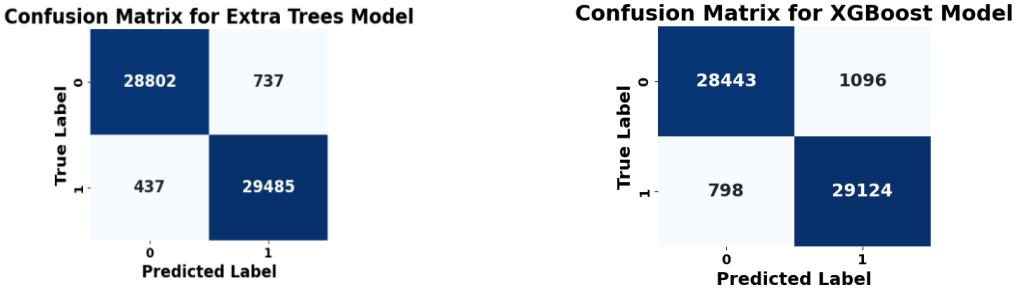

Fig. 6. Confusion Matrix Comparison of Extra Trees and XGBoost Models for Fraud Detection

Fig. 6 shows confusion matrices for Extra Trees and XGBoost classifiers, which reflect their performance in classifying the true TP, TN, FP and FN. The extra Trees model exhibited fewer false positives (737) and false negatives (437) than XGBoost (1096 and 798, respectively), indicating a better trade-off between sensitivity and specificity in fraud detection, as well as higher classification accuracy.

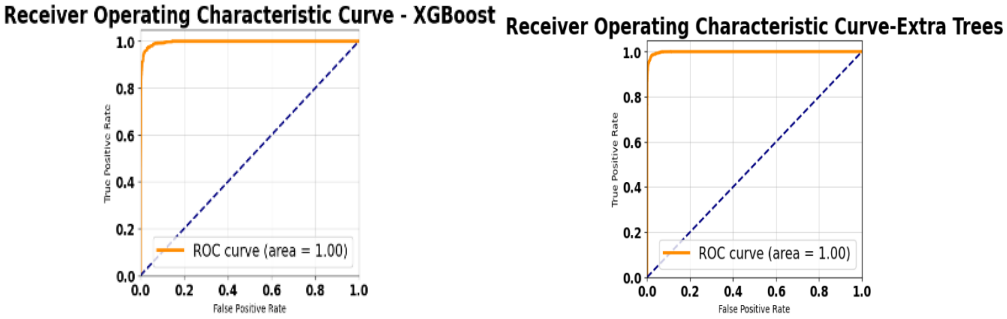

Fig. 7. ROC Curve Comparison of Extra Trees and XGBoost Models for Fraud Detection

The ROC curves for Extra Trees and XGBoost classifiers are shown in Fig. 7, illustrating their capacity to discern between genuine and fraudulent transactions. Strong classification performance is shown by AUC of 1.00 for both models. Good discriminatory performance in the dataset is shown by the curves' proximity to the upper-left corner, which shows high TPR and low FPR.

ISSN : 2455-9679
SJIF Impact Factor : 6.190

C. Comparative Analysis and Discussion

In this section, using key performance metrics, a comparison of suggested and existing fraud detection methods is provided. The most successful baseline models are ResNet and Logistic Regression (LR), with accuracy of 92.4% and 90.44%, respectively, while SVM shows very low values across all metrics. Conversely, the proposed models are more effective than the existing ones, with the Extra Trees classifier recording the best accuracy (98.02%), and F1-score (98.04%) and XGBoost having the second-best, albeit competitive results shows in Table III. This implies that the suggested ensemble-based approaches offer a higher predictive power and overall better classification to detect fraud.

Table 3: Comparison of Baseline and Proposed Machine Learning Models for Fraud Detection Performance Evaluation

References	Models	Accuracy	Precision	Recall	F1-Score
[27]	RestNet	92.4	91.2	92.3	94.2
[28]	LR	90.44	92.89	93.11	92.11
[29]	SVM	33.00	21.84	28.41	24.70
Proposed	Extra Trees	98.02	97.56	0.9853	98.04
Proposed	XGBoost	96.81	96.37	0.9733	0.9685

The presented work introduces an ML-powered fraud detection model that compares the existing and proposed models based on important evaluation metrics. Findings show that the proposed ensemble methods are more effective than traditional models, with the Extra Trees classifier being the one with the highest **accuracy of 98.02%**, thus, the most effective model in this work. This illustrates how effective ensemble learning methods are at enhancing fraud detection capability and classification performance.

D. Limitations and Future Work

Nevertheless, the research has ***several limitations***. It uses a specific dataset, which can limit its applicability to real-world conditions with varied fraud patterns. Little regard is given to real-time deployment and scalability. Also, other factors, including model interpretability, management of changing fraud behavior (concept drift) and more detailed statistical validation, are not yet well considered, which can affect practical applicability.

In the future, the work can be extended to implement a model for real-time fraud detection, thereby enabling its practical application in situ, i.e., in dynamic financial environments. Moreover, cutting-edge methods such as DL, ensemble techniques, and explainable AI (XAI) may enhance model transparency for users and decision-making. Furthermore, analyzing ways to address issues such as class disparity, idea drift, and fraud dynamism will improve the model even further. It could also be extended to distributed systems, blockchain, or federated learning to improve scalability, security, and privacy at the deployment and market levels.

V. CONCLUSIONS

This article presents an ML-driven fraud detection system that analyzes and juxtaposes various models. The Extra Trees model is the most accurate (98.02%), more precise (97.56%), has the highest recall (0.9853), and an F1-score (98.04%) as compared to XGBoost, whose measures are lower (96.81%). The overall findings indicate that the framework is effective at substantially increasing detection accuracy while maintaining an adequate balance between FP and FN. In this way, this solution may be a feasible and scalable approach for real-world financial systems, adding more certainty and trust to digital payment systems. Additional studies could use DL and live detection systems to improve fraud prevention.

REFERENCES

[1] Zoph, B., & Le, Q. V. (2017). Neural architecture search with reinforcement learning. In Proceedings of the International Conference on Learning Representations (ICLR). Toulon, France.

[2] S. A. Pushkala, "Deep Learning Techniques for Real-Time Fraud Detection in Financial Transactions," in 2025 International Conference on Computational Innovations and Sustainable Technologies (ICCIST), IEEE, Dec. 2025, pp. 1–7. doi: 10.1109/ICCIST67338.2025.11438826.

[3] C. Pate, "AI-Driven Recommendation Systems for Improving Online Customer Journey," Int. J. Curr. Eng. Technol., vol. 14, no. 6, 2024, doi: 10.14741/ijcet/v.14.6.18.

[4] P. Jeyachandran, A. Satya, V. Vardhan, P. Subramani, O. Goel, and S. P. Singh, "Leveraging Machine Learning for Real-Time Fraud Detection in Digital Payments," no. 6, pp. 70–94, 2024.

[5] H. P. Cyril, "Serialization of Telecom Provisioning Transactions in Distributed Systems," Int. J. Eng. Adv. Technol. Stud., vol. 15, no. 6, pp. 526–533, 2025, doi: 10.14741/ijcet/v.15.6.6.

[6] D. Shah, S. Khade, and S. Pawar, "Anomaly Detection in Time Series Data of Sensex and Nifty50 With Keras," in 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), IEEE, Mar. 2021, pp. 433–438. doi: 10.1109/ESCI50559.2021.9396979.

[7] S. Singamsetty, "Efficacy of Data Governance a Cutting Edge Approach to Ensuring Data Quality in Machine Learning for Banking Industry," in 2024 2nd International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), IEEE, Dec. 2024, pp. 1–7. doi: 10.1109/SCOPES64467.2024.10991944.

[8] B. Stojanović et al., "Follow the Trail: Machine Learning for Fraud Detection in Fintech Applications," Sensors, vol. 21, no. 5, p. 1594, Feb. 2021, doi: 10.3390/s21051594.

[9] F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan, and M. Ahmed, "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms," IEEE Access, vol. 10, pp. 39700–39715, 2022, doi: 10.1109/ACCESS.2022.3166891.

[10] S. Singamsetty, "AI-Based Data Governance: Empowering Trust and Compliance in Complex Data Ecosystems," Int. J. Comput. Math. Ideas, vol. 13, no. 03, pp. 1007–1017, 2021, doi: 10.70153/IJCMI/2021.13301.

[11] A. Nerella, P. Badri, K. Sundravadivelu, and R. Murugesan, "Navigating Regulatory Hurdles in AI-Driven Credit Card Approvals : Balancing Innovation and Compliance," J. Inf. Syst. Eng. Manag., vol. 8, no. 4, pp. 1–9, 2023.

[12] N. R. Barot, "Transparency-Driven Operational Intelligence: A New Data Governance Model for High-Risk Industrial Automation," J. Inf. Syst. Eng. Manag., vol. 10, no. 63s, pp. 1019–1028, Dec. 2025, doi: 10.52783/jisem.v10i63s.13975.

[13] K. Ramkumar, S. Kilaru, P. Preethi, A. Nerella, S. Kilaru, and G. G. Battu, "A Temporal Graph Neural Network Approach for Deep Fraud Detection in Real-Time Financial Transactions," in The 16th International IEEE Conference on Computing, Communication and Networking Technologies (ICCCNT), 2025.

[14] D. Patel, "Explainable Risk Decision Systems Using Artificial Intelligence Models for Payment Fraud Identification with Mitigation," in 2026 14th International Symposium on Digital Forensics and Security (ISDFS), IEEE, Mar. 2026, pp. 01–06. doi: 10.1109/ISDFS69419.2026.11459006.

[15] A. F. Sariat, I. J. Siddique, M. Hossain, M. M. Islam, and T. Rahman, "AI Driven Fraud Detection in Financial Ecosystems: A Hybrid Machine Learning Framework," in 2025 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2025, pp. 1–8. doi: 10.1109/ECCE64574.2025.11013808.

[16] P. S. Ingle, S. B. Mujumale, P. Pandhare, N. Wasatkar, P. P. Hujare, and Y. Sanap, "Artificial Intelligence-

Driven Fraud Detection in Digital Payment Systems: A Hybrid Machine Learning Approach," in 2025 IEEE Pune Section International Conference (PuneCon), IEEE, Dec. 2025, pp. 1–5. doi: 10.1109/PuneCon67554.2025.11378213.

[17] M. Bhutta and A. Mehmood, "Evaluating Classical and Quantum Machine Learning for Credit Card Fraud Detection: Performance and Economic Impact," in 2025 Cyber Awareness and Research Symposium (CARS), IEEE, Oct. 2025, pp. 1–9. doi: 10.1109/CARS67163.2025.11337509.

[18] A. Chakraborty, G. Singh, Meenakshi, V. Sirvastava, and S. A. Dhondiyal, "Blockchain-Enhanced Adversarial Machine Learning for Fraud Detection and Claims Automation in the Insurance Sector," in 2024 5th International Conference on Data Intelligence and Cognitive Informatics (ICDICI), IEEE, Nov. 2024, pp. 80–86. doi: 10.1109/ICDICI62993.2024.10810927.

[19] N. Deepa, A. Alkhayyat, J. Lande, S. Armoogum, and S. S. Bhoomika, "Fraud Detection using Enhanced Secure Machine Learning Algorithm for Wireless Communication," in 2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON), 2024, pp. 1–4. doi: 10.1109/NMITCON62075.2024.10699185.

[20] S. Rani and A. Mittal, "Securing Digital Payments a Comprehensive Analysis of AI Driven Fraud Detection with Real Time Transaction Monitoring and Anomaly Detection," in Proceedings of International Conference on Contemporary Computing and Informatics, IC3I 2023, 2023. doi: 10.1109/IC3I59117.2023.10397958.

[21] A. Selvaraj, B. Krothapalli, and V. P. Rambabu, "Data Governance in Retail and Insurance Integration Projects: Ensuring Quality and Compliance," J. Artif. Intell. Res., vol. 3, no. 1, 2023.

[22] B. S. Prashanth et al., "Prediction of bank transaction fraud using TabNet—an adaptive deep learning architecture," Int. Rev. Econ. Financ., vol. 106, p. 104916, Mar. 2026, doi: 10.1016/j.iref.2026.104916.

[23] J. M. H. Pinheiro et al., "The Impact of Feature Scaling in Machine Learning: Effects on Regression and Classification Tasks," IEEE Access, vol. 13, pp. 199903–199931, 2025, doi: 10.1109/ACCESS.2025.3635541.

[24] A. A. Compagnino et al., "An Introduction to Machine Learning Methods for Fraud Detection," Appl. Sci., vol. 15, no. 21, p. 11787, Nov. 2025, doi: 10.3390/app152111787.

[25] N. B. Chakka and S. S. Saheb, "Mobile payment fraud detection in UPIs through machine learning techniques: A systematic review," Multidiscip. Rev., vol. 9, no. 6, p. 2026280, Nov. 2025, doi: 10.31893/multirev.2026280.

[26] C. S. Pareek, "From Detection to Prevention: The Evolution of Fraud Testing Frameworks in Insurance Through AI," J. Artif. Intell. Mach. Learn. Data Sci., vol. 1, no. 2, pp. 1–8, May 2023, doi: 10.51219/JAIMLD/chandra-shekhar-pareek/401.

[27] R. Q. Majumder, "Data Driven-Machine Learning-Based Fraud Detection Models in FinTech Financial Transactions," IEEE, pp. 1–6, 2025.

[28] A. A. Almazroi and N. Ayub, "Online Payment Fraud Detection Model Using Machine Learning Techniques," IEEE Access, vol. 11, pp. 137188–137203, 2023, doi: 10.1109/ACCESS.2023.3339226.

[29] N. Abid, "Improving Accuracy and Efficiency of Online Payment Fraud Detection and Prevention with Machine Learning Models," Int. J. Innov. Sci. Res. Technol., vol. 9, no. 12, pp. 711–723, 2024.

[30] M. A. Alrasheedi, "Enhancing Fraud Detection in Credit Card Transactions: A Comparative Study of Machine Learning Models," Comput. Econ., Aug. 2025, doi: 10.1007/s10614-025-11071-3.