Volume 2 No 6, 2026
Journal of Global Research in Multidisciplinary Studies (JGRMS)
Research Paper
Available online at https://saanvipublications.com/journals/index.php/jgrms/index

A Predictive Model for Fraud Detection in Digital Payment Transactions using ML Algorithms

Dr. Neetu Sikarwar
Department of Electronice Engineering
Institute of Engineering, Jiwaji University
Gwalior, India
Neetusik1@gmail.com

***Abstract*—The rapid evolution of online payment technologies has greatly simplified access to capital. But instead, the possibility of fraudulent transactions is on the rise. Fraud of a more complex kind may occasionally go undetected by conventional rule-based systems. Combining LightGBM and Random Forest (RF), two machine learning approaches, this study presents a robust fraud detection tool to combat fraudulent financial transactions. The method consists of applying data preprocessing methods which include handling missing data, deleting duplicate data, data normalization, applying one hot encoder method, and applying data balancing using SMOTE technique. Moreover, exploratory data analysis includes using correlation matrix and analyzing transactions in order to discover any fraud patterns. In addition, dividing the processed dataset into a test and train dataset for implementing the model. The LightGBM model surpasses the other strategies with a 98.80% accuracy (ACC), 99.37% precision (PRE), 99.37% recall (REC), 98.80% F1-Score (F1), and an AUC-ROC of 0.99%.**

Keywords*—*Digital Payment Fraud Detection, Machine Learning, Financial Transaction Security, Credit Card Fraud Detection, Predictive Analytics.

I. INTRODUCTION

The rapid growth of digital tools has changed how people around the world conduct financial transactions. The widespread use of safe, quick, and easy online payment methods, including internet banking, mobile wallets, and credit/debit cards, has made these systems necessary [1][2]. With the increasing adoption of e-commerce platforms and digital banking services, the volume of online transactions has increased significantly[3][4][5].

Online payment fraud has increased with this fast growth, which is a major problem for consumers, businesses, and banks alike [6]. Digital payment systems lose customer confidence and suffer huge financial losses due to fraudulent activities such identity theft, unauthorized transactions, and account takeovers [7][8]. Handling massive amounts of transaction data and ever-changing fraud trends can be a real challenge for traditional fraud detection approaches that depend on predetermined criteria and manual verification [9][10]. Threats to individuals, businesses, and the entire financial system are growing in complexity and variety as a result of the rapid digitization of the financial industry [11]. Because of their inherent limitations, existing fraud detection technologies are finding it more and harder to properly respond to emerging frauds in this environment[12].

One revolutionary method for tackling this issue is machine learning, or ML [13][14]. A characteristic of ML models is their capacity to sift through mountains of transaction data in search of outliers and, with time, train for exponentially improved ACC [15][16]. Highly effective classification methods help these programs more accurately tell the difference between real and fake deals. This article suggests a method for detecting fraudulent transactions by utilizing supervised learning algorithms trained on datasets containing real financial transactions. These algorithms include LR, RF, and XGBoost. Finding important characteristics that can aid the system in efficiently detecting anomalies such as transaction amount, duration, frequency, IP address, and device information.

The rise in digital payment platforms and online financial transactions is increasing the potential for fraud, resulting in huge monetary losses[17][18]. Fraud detection techniques available to date do not work well for large and unbalanced datasets. An intelligent system for detecting fraudulent transactions quickly and accurately, with few false alarms, is definitely necessary. This paper addresses the problem of fraud in financial transaction systems by implementing ML algorithms. The key contributions of this study are discussed below:

- Using ML frameworks with models such as LightGBM and RF, digital payment fraud can be easily detected.
- Utilized efficient preprocessing methods to enhance model performance, such as data cleansing, feature modification, and class balancing.
- Performed exploratory data analysis to identify important fraud-related transaction patterns.
- Evaluated the proposed models using standard performance metrics for fraud classification.
- Verified the efficacy of the suggested framework by comparing it to preexisting ML to DL methods.

A. Justification And Novelty

Intelligent fraud detection algorithms should be included in digital payment systems due to the fast increasing number of fraudulent transactions, which is the reason for this research. Conventional methods sometimes find it challenging to detect fraud due to imbalanced transaction data sets, thus decreasing the efficiency of detection. In order to address such problems, an innovative algorithm based on the machine learning technique with the implementation of LightGBM and

Random Forest methods has been suggested. The novelty of the suggested model consists of implementing the methods of data balancing and comparative analysis in the process of fraud classification.

B. Organization of the Paper

The structured of the paper is in following way: Section II includes a literature review and discusses approaches for detecting fraud in online payment systems. Section III details the technique that has been suggested, which includes procedures for preprocessing data, balancing SMOTE, and utilizing LightGBM and RF models. Section IV discusses the experimental results and draws comparisons between the proposed models. The paper concludes in Section V, which also proposes future research directions for better fraud detection systems.

II. Literature Review

Recent studies have explored various ML and DL approaches for fraud detection in digital payment systems. Ambedkar et al. (2026) proposed a real-time bank transaction fraud detection framework that integrates Apache Kafka for high-throughput data streaming with ML techniques for intelligent fraud classification. Experimental evaluation demonstrates that the proposed ensemble-based framework achieves 96.8% ACC, 97.2% REC, and an average end-to-end detection latency of 42 milliseconds under normal streaming conditions, while maintaining stable performance during high transaction bursts and simulated broker failures [19].

R et al. (2025) developed a fraud detection system using a ML approach. The proposed system effectively identifies rogue transactions by pursuing transaction behaviour and user activity. Experimental results show that the ACC of the model based on logistics regression is 97%. Comparative analysis with decision trees, random forests, and Naive Bayesian models checks for excellent performance of logistics regression related to ACC and recirculation measurements [20].

Mandlik et al. (2025) introduced an AI-powered strategy to identify fraudulent UPI transactions in real-time. Method employs ML and DL methods, including anomaly detection and supervised learning, to detect suspicious activities. Testing system on simulated UPI transaction data, achieved an F1-score of 0.92 and a false positive rate of 0.03, outperforming existing methods [21].

Prayitno et al. (2024) focuses on detecting digital payment system fraud using the SVM algorithm with a RBF kernel. One of the main problems that this study aims to solve is the dataset's severe class imbalance, which accounts for only 0.17 percent of all transactions. The dataset was balanced, and the model's capacity to identify fraudulent patterns was enhanced using the SMOTE approach. With a REC of 75.51% and a PRE of 86.23%, the SVM model attained an ACC of 99.93%, according to the results [22].

Dahiphale et al. (2024) uses LLMs to improve the ACC of scam classification and creates a digital assistant to help human reviewers spot and prevent fraud. The findings show that LLMs can enhance current ML models, leading to more reliable scam reviews with higher efficiency, ACC, quality, and consistency; this, in turn, can make online payments safer. Found that the Gemini Ultra model could accurately classify scams with a 93.33% success rate when tested on carefully selected transaction data [23].

R et al. (2023 performed an investigation into real-time credit card fraud techniques and detection tools. The credit card user dataset has been used to classify legitimate and fraudulent transactions using a variety of ML approaches, including SVM, LR, FSVM, and RF for fraudulent transaction identification. With a success percentage of 98.61%, the FS VM surpassed the other algorithms in the comparison [24]. The paper provides an overview of the most recent research on Financial Fraud in Banking by summarizing the offered models, datasets, and findings, as well as limitations and future work in the field (Table I).

A. Gap Analysis

In recent years, most studies related to the detection of frauds in digital payment systems have considered high efficiency in terms of ACC. Despite the efforts made towards enhancing the ACC levels, some of the current methods have been facing problems like the issue of imbalanced classes, high false-positive rate, and lack of scalability when dealing with large transaction volumes. Additionally, some of the models consume considerable amounts of computational power and are less accurate in detecting frauds in real time. Another aspect that has not been explored sufficiently in the literature is the comparison between advanced ensemble models and deep learning algorithms.

TABLE I. Comparative Analysis of Existing Fraud Detection Models in Digital Payment Systems

Author	Proposed Work	Results	Key Findings	Limitations & Future Work
Ambedkar et al. (2026)	Built a platform for detecting fraud in real-time using ensemble ML and Apache Kafka.	Achieved 96.8% accuracy, 97.2% recall, and 42 ms detection latency.	Ensemble learning with streaming architecture improves real-time fraud detection efficiency.	Requires scalability testing on larger real-world banking datasets and advanced cyberattack scenarios.
R et al. (2025)	Proposed a fraud detection system using Logistic Regression and comparative ML models.	Logistic Regression achieved 97% accuracy.	Logistic Regression performed better than DT, RF, and Naïve Bayes for transaction classification.	Future work can focus on deep learning and hybrid models for complex fraud patterns.
Mandlik et al. (2025)	Introduced an AI-driven real-time UPI fraud detection system using anomaly detection and supervised learning.	Achieved F1-score of 0.92 and false positive rate of 0.03.	AI-based approaches effectively identify suspicious UPI activities in real time.	Performance can be enhanced using larger datasets and advanced transformer-based models.
Prayitno et al. (2024)	Implemented SVM with RBF kernel and SMOTE for fraud detection in digital payments.	Achieved 99.93% accuracy, 86.23% precision, and 75.51% recall.	SMOTE effectively addressed class imbalance issues in fraud datasets.	Recall performance can be improved to reduce missed fraud transactions.

Dahiphale et al. (2024)	Used Large Language Models (LLMs) for scam classification and fraud review assistance.	Gemini Ultra achieved 93.33% accuracy.	LLMs improve scam analysis quality and reviewer support in digital payment systems.	Future work should focus on real-time deployment and computational optimization of LLMs.
R et al. (2023)	Compared FSVM, RF, LR, and SVM for credit card fraud detection.	FSVM achieved the highest accuracy of 98.61%.	Fuzzy-based SVM outperformed traditional ML algorithms in fraud classification.	Future work can integrate deep learning and adaptive learning techniques for evolving fraud patterns.

III. Research Methodology

The suggested approach of detecting financial transaction fraud is shown in Fig. 1. In the beginning, the dataset containing credit card transactions goes through a series of preprocessing procedures, including dealing with missing values, eliminating duplicate entries, scaling the data, one-hot encoding, and data balancing. The next stage is to split the processed dataset in half, so that each half may be used for testing and training. At last, apply RF and LightGBM models to forecast fraudulent transactions, and measure their efficacy with ROC, F1, REC, ACC, and PRE.

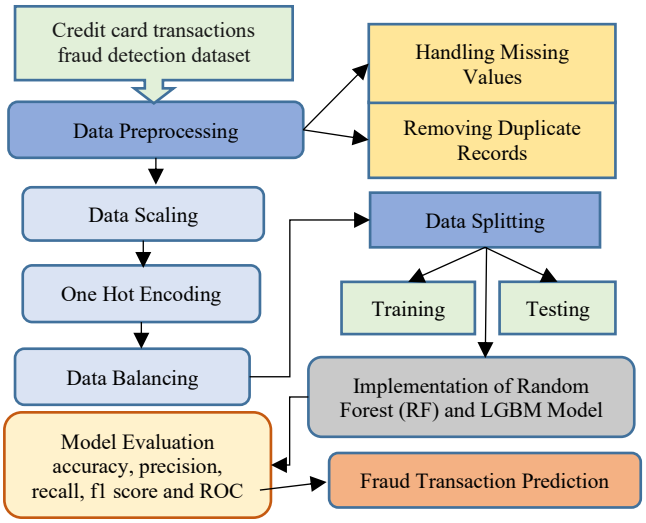

Fig. 1. Proposed flowchart for Fraud Detection

The following section lays out the suggested methodology's steps in great detail:

A. Data Gathering and Analysis

This data set was made public at Kaggle and included simulated credit card transactions that occurred in the western United States of America between January 1, 2020, and December 31, 2020. The transactions included both valid and fraudulent ones. All told, it comprises 1000 consumers' use of their credit cards to buy things from 800 different stores. The data set contains information regarding every purchase, such as the identity of the buyer, the name of the seller, the sort of purchase, and any signs of fraud. There is a total of 23, and there are 555,719 rows of observations. Qualitative data makes up twelve of these variables. The following are examples of data visualizations: bar graphs, heatmaps, and distribution charts:

Fig. 2 shows the correlation matrix of the chosen transaction features for fraud detection. A stronger negative link is indicated by numbers closer to -1 on the heat map, while numbers closer to 1 indicate a stronger positive association. The positive correlation of merch_lat and Latitude and of merch_long and Longitude is very high. The other features demonstrate low correlation values, which means there is very little multicollinearity and hence the models trained for fraud detection.

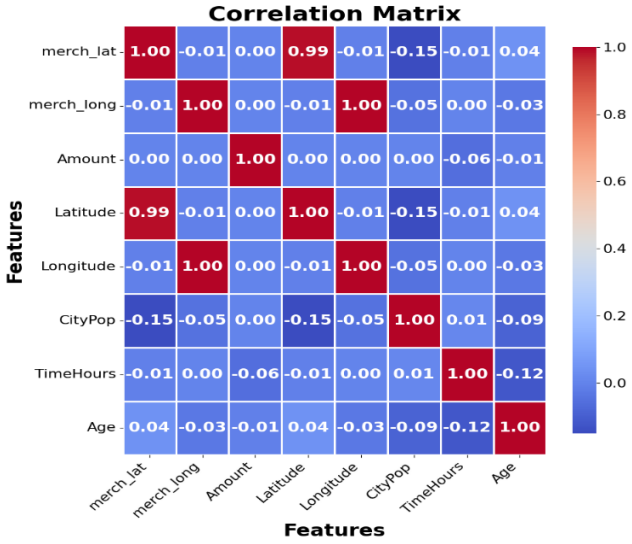

Fig. 2. Correlation Matrix of Financial Transaction Features

Fig. 3. Comparison of Fraud and Non-Fraud Transaction Amounts

Fig. 3 shows the boxplot comparison of amounts from valid and fraudulent transactions. The volume and variability of fraud transactions tend to be higher than those of non-fraud transactions. According to the data, the quantity of a transaction is a key indicator of financial fraud.

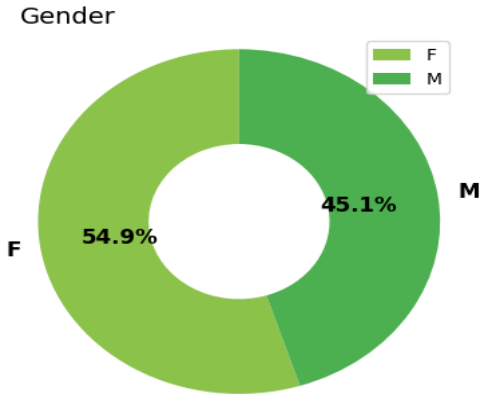

Fig. 4. Gender Distribution of Transaction Records

Gender distribution of transaction records is shown in Fig. 4. Females are involved in 54.9% of the transactions, and males are involved in 45.1% of the transactions. The distribution suggests that the sample included a fairly equal number of males and females in the financial transaction dataset.

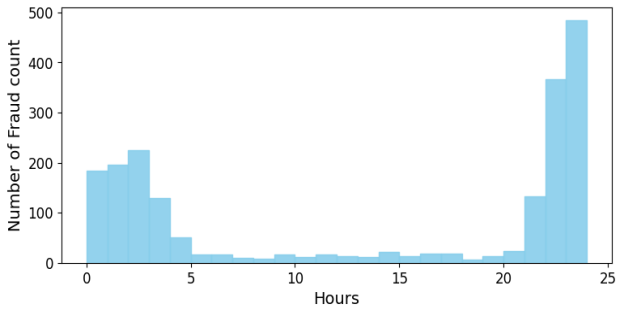

Fig. 5. Distribution of Fraud Transactions Across Different Hours

The figures in the bar chart below reveal the distribution of fraud across the day (Fig. 5). A higher number of fraudulent transactions are found during the late night (22-24 hours) and fewer during the daytime. This pattern suggests that transaction time is a key factor to consider when analyzing fraud.

B. *Data Pre-Processing*

The primary step in preparing datasets for analysis and simulation of financial fraud is to perform the required preparation. Standardizing data scaling, dealing with missing data, and eliminating duplicate information are all part of these procedures. These steps can help clean and organize the dataset, which in turn can allow for more precise analysis and more efficient modeling in the fight against financial fraud.

- **Addressing Missing Data:** The procedure for handling missing values include completing datasets that have gaps due to missing or incomplete data.
- **Eliminating Duplicate Records:** Eliminating duplicate records helps keep data accurate and reliable by making sure dataset only contains unique data points and avoiding bias that can be introduced by such duplications. This method includes going over each record and keeping only the unique ones after comparing them to all the others.
- **Data Scaling:** Data scaling is an important part of preparing data that aims to standardize and make comparable all numerical features. Standardization and min-max scaling are two popular approaches to this. The purpose of standardization is to create a feature X with a uniform distribution of values, where the mean is 0 and the standard deviation is 1. The transformation allows for more accurate comparisons of feature X with other features in the dataset, as demonstrated in Equation (1).

$$X_{scaled} = \frac{X-\mu}{\sigma} \quad (1)$$

In this context, Xscaled represents the modified feature, The initial characteristic is symbolized by X, the average is represented by μ, and the dispersion is symbolized by σ.

- **One Hot Encoding:** Feature vectors containing categorical variables can be handled via One Hot Encoding. Make these groups machine-learning-algorithm compatible by transforming them into binary vectors, then check if they improve the model's ACC.

C. *Data balancing using SMOTE*

Data preprocessing utilizing the SMOTE technique should be initiated to address class imbalance. When used on a dataset, SMOTE creates synthetic samples for underrepresented classes to help even out the distribution of those classes. This strategy brings the number of majority-involved cases down to the same level as the minority-involved cases for this minority group, or to a somewhat lower level.

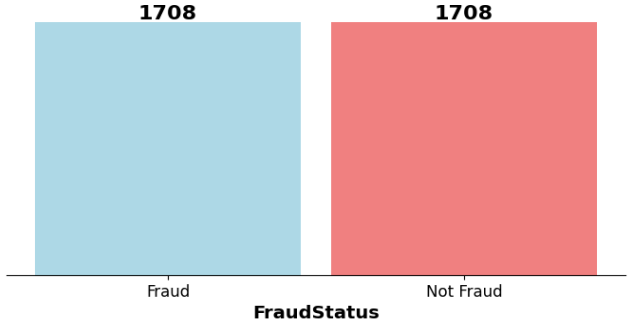

Fig. 6. Class distribution after SMOTE of Dataset

Fig. 6 displays the dataset's partitioning between fraudulent and legitimate transactions. Both classes contain 1,708 transactions, indicating that the dataset is balanced and suitable for training and evaluating an effective fraud detection model.

D. *Data Splitting*

A perfect 80:20 split between the dataset's training and testing sets ensures that each set contains the same number of classes as the original.

E. *Implementation of the Proposed Model*

This section explains the ML models:

1) *LightGBM Model*

LightGBM is a state-of-the-art ensemble ML approach that is an extension of gradient boosting. It is intended to create multiple decision trees one after another to increase the prediction ACC and to reduce the computation time. In the leaf-wise tree growth strategy, LightGBM has improved at reducing loss, and boosts model performance on large and imbalanced data sets like financial transaction fraud data. The LightGBM model's prediction function is shown in Equation (2):

$$\hat{y}_i = \sum_{k=1}^{k} f_k(x_i), f_k \epsilon F \quad (2)$$

In this context, $\hat{y}_i$ stands for the anticipated result, f_k for the kth decision tree, x_i for the input feature vector, and K for the overall count of trees. Here is the definition of the objective function of LightGBM, Equation (3):

$$Obj = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{K=1}^{K} \Omega(f_k) \quad (3)$$

The loss function $l(y_i, \hat{y}_i)$ measures the prediction error, and the regularization term $\Omega(f_k)$ prevents overfitting.

2) *Random Forest Model*

The RF approach is a form of ensemble learning that improves classification or regression performance by building and combining many decision trees. The trees are guaranteed

to be diverse by training them with a randomly selected subset of features at each split and a randomly selected portion of training data (bagging). Using the Gini impurity as a function, and can detect the quality of the split in each node. Use equation (4) to find the Gini impurity at node N.

$$g(N) = \sum_{i \neq j} P(w_i)P(w_j) \quad (4)$$

Where P(w_i) is the amount of the population that falls into class i. Equation (5) can be used to calculate the entropy in a node N.

$$H(N) = \sum_{i=1}^{i=d} P(w_i) log_2(Pw_i) \quad (5)$$

In which P(w_i) is the proportion of the population given the label i and d is the number of classes. At a node with an equal amount of each class, the entropy is at its maximum. The following Equation (6) can be used to compute the information gain that results from a split:

$$\Delta \mathrm{I(N)} = \mathrm{I(N)} - P_L * \mathrm{I}(N_L) - P_R * \mathrm{I}(N_L) \quad (6)$$

With I(N) representing node N's Gini or Shannon entropy, P_L and P_R denoting the population proportions that go to N's left and right children, respectively, following the split. This is where N_R and N_L come in, these are N's left and right children, respectively.

F. Evaluation Metrics

A handful of critical performance metrics activate the efficacy of the proposed architecture [25]. TP, FP, TN, and FN can be defined using this matrix. Important performance indicators like REC, ACC, PRE, and F1 are computed in the following way:

1) Accuracy

The ratio of a model's accurate forecasts to its total number of predictions, including both fraudulent and non-fraudulent predictions, is called the ACC in forecasting fraudulent conduct ACC. Equation (7) is used for the calculation -

$$Accuracy = \frac{TP+TN}{TP+Fp+TN+FN} \quad (7)$$

2) Precision

The ACC ratio is determined by dividing the total number of transactions (TP + FP) by the proportion of fraudulently predicted transactions. For the computation, Equation (8) is utilized-

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \quad (8)$$

3) Recall

The ratio of "correctly identified fraud transactions" (REC) to the total number of fraudulent transactions (TP) is the key metric to consider. Here is Equation (9):

$$Recall = \frac{\mathrm{TP}}{TP+FN} \quad (9)$$

4) F1 score

It's a great tool for maintaining a healthy equilibrium between PRE and REC since it's the complementary midpoint of PRE and memory. Its bounds are the integers 0 and 1. The Equation (10) for it is:

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (10)$$

5) Receiver Operating Characteristic Curve (ROC)

The AUC-ROC indicates the model's ability to differentiate between valid and fraudulent transactions across different threshold zones. Values closer to 1 indicate near-perfect discrimination, and higher AUC-ROC values indicate superior model performance.

An FP is the sum of all erroneous positive forecasts, and an FN is the sum of all erroneous negative predictions. If all of predictions come true, call it a TP, and if all of predictions come true, call it a TN.

IV. Results and Discussion

The proposed ML models' ACC for detecting fraudulent credit card transactions, using the Credit Card Transactions Fraud Detection dataset, is compared in Table II. With a 98.80% ACC rate, 99.37% PRE, 99.37% REC, 98.80% F1, and an AUC-ROC of 0.99, LightGBM outperformed all other models in accurately detecting fraudulent transactions with few FP. RF was the top-performing model as well, with a 96.89% ACC and an AUC of 0.98. A model that routinely beats RF in detecting financial fraud, LightGBM is highly effective.

TABLE II. Results of the Proposed Models for Fraud Detection in Financial Transactions using Credit Card Transactions Fraud Detection Dataset

Matrix	Accuracy	Precision	Recall	F1-score	AUC-ROC
LightGBM	0.9880	0.9937	0.9937	0.9880	0.99
RF	0.9689	0.9769	0.9606	0.9687	0.98

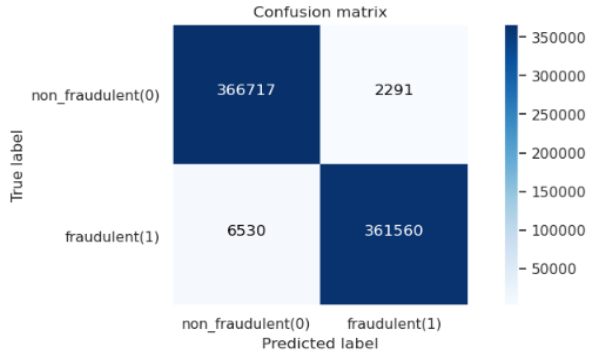

Fig. 7. Confusion matrix of the LGBM model

The LightGBM (LGBM) model, which can detect financial transaction fraud, uses a confusion matrix, as shown in Fig. 7. Proving its predictive prowess, the program accurately categorized 361,560 fraudulent transactions and 366,717 legitimate ones. On the other hand, 6,530 "new" transactions were detected as non-fraudulent, while 2,291 "new" transactions were genuinely legitimate but mistakenly labeled as fraudulent (FN).

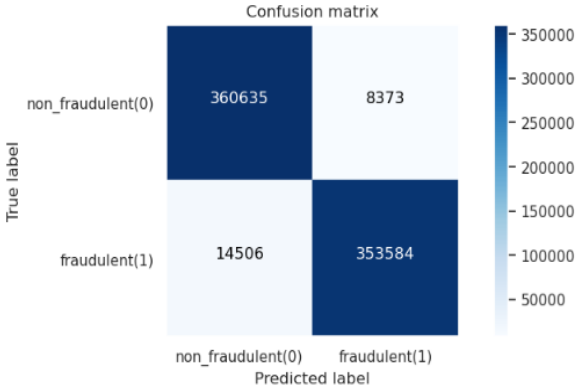

Fig. 8. Confusion Matrix of the RF Model

The RF model's confusion matrix is displayed in Fig. 8. They are able to accurately predict 360,635 non-fraudulent transactions and 353,584 fraudulent transactions. But, 8,373 transactions are incorrectly identified as fraud and 14,506 transactions are incorrectly identified as non-fraud. The RF model performed well in detecting fraud, according to the results, with a high number of predictions.

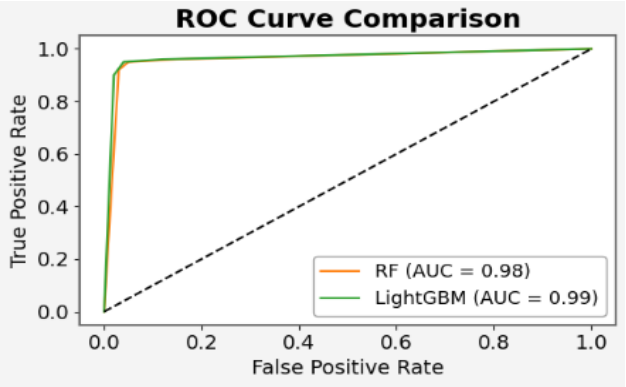

Fig. 9. ROC for the RF and LightGBM Model

The ROC curve comparison between fraud detection models, RF and LightGBM is shown in Fig. 9. ACC classification was better with the LightGBM model (AUC = 0.99 vs. 0.98 for the RF model). There was a big difference between real and fake trades in both models.

A. Comparative Analysis

Table III provides a comparison of the different ML and DL models put into use for detecting financial crime. The suggested LightGBM model outperformed all others with a total ACC of 98.80%, PRE of 99.37%, REC of 99.37%, and F1 of 98.80%. With an ACC of 96.49% and an F1 of 96.78%, the RF model was likewise very effective. When compared to other models, such as ResNet, CapsNet, CNN, GNN, and KNN, among others, the suggested models performed better in terms of fraud detection and classification.

TABLE III. COMPARISON OF DIFFERENT MACHINE LEARNING AND DEEP LEARNING MODELS FOR FINANCIAL FRAUD DETECTION

Model	ACC	Precision	Recall	F1-score
ResNet[26]	0.924	0.912	0.923	0.942
CapsNet[26]	0.912	0.870	0.915	0.920
CNN[27]	0.7610	0.7708	0.7719	0.8900
GNN[28]	0.7801	0.8163	0.8062	0.8136
KNN[29]	0.941	0. 965	0.9202	0. 942
LightGBM	0.9880	0.9937	0.9937	0.9880
RF	0.9689	0.9769	0.9606	0.9687

The suggested system for detecting fraudulent financial transactions relies on the LightGBM and RF models. In particular, the LightGBM model performed best due to its lightweight training, efficient processing of large-scale, imbalanced transaction data, high ACC and strong predictive performance. The RF also showed good classification performance, without extensive over-fitting and robustness. A trustworthy and safe digital payment fraud detection system might benefit from the suggested models since they were more accurate, precise, REC, and F1 than competing DL and ML models.

B. Limitations and Future Work

The suggested method of detecting fraud shows high results of classification, nevertheless, there are some disadvantages to consider. First of all, the research is based on just one dataset of transactions that might have an impact on the generalizability of the models within the real-life environment. Moreover, changes in fraud patterns and the increasing complexity of transactions could affect the adaptability of the suggested models in the future. The direction for further research is to combine methods of deep learning and hybrid ensembling with real-time systems.

V. CONCLUSION AND FUTURE STUDY

Online payment fraud has been identified as one of the leading frauds in the past few decades. This paper presented a robust and efficient fraud detection system for digital payments through LightGBM and RF. This study employs preprocessing steps on the raw dataset, such as filling in missing values, deleting duplicates, scaling data, one-hot encoding, and SMOTE data balancing, among others. Splitting the processed data into a training dataset and a testing dataset allowed for more precise model implementation. While other deep learning algorithms performed better than the LightGBM algorithm, it outperformed them all with a 98.80% F1-Score, 99.37% REC, 99.37% PRE, and an AUC-ROC value of 0.99. Digital payment fraud detection systems can be significantly improved in the future by including explainable AI approaches, real-time transaction monitoring, and DL.

REFERENCES

[1] Sonam, I. Mazhar, and A. Sulthana, "Online Payment Fraud Detection Using Machine Learning Techniques," *Iconic Res. Eng. Journals*, vol. 9, no. 6, Dec. 2025, doi: 10.64388/IREV9I6-1713148.

[2] 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)*, Boston, MA, USA: IEEE, 2026, pp. 01–06, April. doi: 10.1109/ISDFS69419.2026.11459006.

[3] A. K. Padhy, C. Medicherla, B. Vulugundam, C. Kulkarni, T. P. Patel, and S. Shivam, "Latency-Optimized Microservices Orchestration for Real-Time E-Commerce in Multi-Cloud Environments," in *2025 International Conference on Computer and Applications (ICCA)*, 2025, pp. 1–6. doi: 10.1109/ICCA66035.2025.11430930.

[4] T. P. Patel and P. Agarwal, "Dynamic Multi-Agent Reinforcement Learning for Competitive E-Commerce Price Optimization: Curriculum-Driven SAC, Equilibrium Analysis and Ethical Safeguards," in *2026 International Seminar on Intelligent Business and Edge-Computing Research (ISIBER)*, 2026, pp. 214-219,February. doi: 10.1109/ISIBER68248.2026.11470206.

[5] D. Patel, "Integrating Price Elasticity and Reinforcement Learning: A Data-Driven Framework for Strategic E-commerce Pricing," in *2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC)*, Houston, TX, USA: IEEE, 2026, pp. 1–6, February. doi: 10.1109/ICAIC67076.2026.11395747.

[6] M. E. Lokanan, "Predicting Mobile Money Transaction Fraud using Machine Learning Algorithms," *Appl. AI Lett.*, vol. 4, no. 2, Apr. 2023, doi: 10.1002/ail2.85.

[7] G. Ahmed, A. Ahmed, M. Ahmed, J. Latha, and P. Kumar, "Indian Banking Precision Marketing: A Comparative Analysis of Machine Learning Customer Segmentation Algorithms," in *2024 2nd International Conference on Cyber Resilience (ICCR)*, 2024, pp. 1–6. doi: 10.1109/ICCR61006.2024.10532917.

[8] C. Patel, "Generative AI for Personalized Marketing and Customer Experience in E-Commerce," *Int. J. Emerg. Res. Eng. Technol.*, vol. 7, no. 1, pp. 12–19, 2026, doi: 10.63282/3050-922X.IJERET-V7I1P103.

[9] S. Irfan, "AI-Driven Credit Scoring: A Study of Machine Learning Algorithms for Enhanced Risk Assessment," in *2026 IEEE International Conference on AI Engineering and Innovations (AIEI)*, IEEE, Mar. 2026, pp. 1–6. doi: 10.1109/AIEI69164.2026.11496824.

[10] S. A. Pushkala, "Financial Fraud Identification Using Graph Neural Network And LSTM With Autoencoder-Based Data Refinement," *J. Int. Cris. Risk Commun. Res.*, vol. 9, no. 1, pp. 198–213, Jan, 2026, doi: 10.63278/jicrcr.vi.3615.

[11] N. R. Panguluri and M. S. Jasti, "Fraud Detection in Digital payments Using Artificial Intelligence," *Int. J. Manag. IT Eng.*, vol. 14, no. 08, pp. 12–23, 2024.

[12] J. Mishra, S. Rai, B. Moharana, V. K. Singh, S. Dey, and T. Sarkar, "Revolutionizing Financial Fraud Detection in Healthcare Insurance Through Blockchain & AI-Driven Predictive Analysis," in *2025 International Conference on Technology Enabled Economic Changes (InTech)*, Tashkent, Uzbekistan: IEEE, 2025, pp. 1516–1522, February. doi: 10.1109/InTech64186.2025.11198190.

[13] L. S. Jattan and R. R. Prajapati, "Digital Payment Fraud Detection System Using Machine Learning," *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 07, no. 06, 2025, doi: 10.56726/IRJMETS79983.

[14] B. Mohan, V. R. Surasani, and R. Kumar, "Autonomous Data Stewardship: Multi-Agent AI for Real-Time Master Data Management in Financial Services," in *2026 IEEE 16th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, USA: IEEE, 2026, pp. 0689–0698, February. doi: 10.1109/CCWC67433.2026.11393832.

[15] P. Parida, "Language Translation Model by leveraging AI and its impact on banking expansion," *Int. J. Comput. Appl.*, vol. 187, no. 63, pp. 55–59, Dec. 2025, doi: 10.5120/ijca2025926048.

[16] 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, Nov. 2023.

[17] D. S. L. Anusha Nerella, Pratik Badri, Dr. P. Preethi and Sheeba, "Unified Finance : Balancing Mobile Wallets and Traditional Payment Methods," *J. Electr. Syst.*, vol. 17, no. 4, pp. 157–165, April, 2021.

[18] A. B. Chatterjee, "Designing Zero-Downtime, Cloud-Native Transaction Processing Architectures for 24×7 Global Payment Networks," *Int. J. Emerg. Trends Comput. Sci. Inf. Technol.*, vol. 7, no. 1, pp. 137–145, Feb, 2026, doi: 10.63282/3050-9246.IJETCSIT-V7I1P120.

[19] J. B. Ambedkar, K. H. Devi, D. M. Chandra, A. H. S. Vardhan, and C. B. V. Yashwanth, "Real-Time Bank Transaction Fraud Detection using Apache Kafka and Machine Learning," in *2026 International Conference on Smart Electronic Devices and Intelligent Systems (ICSEDIS)*, IEEE, Apr. 2026, pp. 1246–1253. doi: 10.1109/ICSEDIS68157.2026.11518536.

[20] R. R, S. C. H, and G. R, "Leveraging Machine Learning Techniques of Real Time Detection of UPI Fraud," in *2025 7th International Conference on Intelligent Sustainable Systems (ICISS)*, IEEE, Mar. 2025, pp. 1506–1510. doi: 10.1109/ICISS63372.2025.11076419.

[21] A. S. Mandlik, M. S. Ganeshpure, C. N. Kadadas, L. Korra, J. U. Kidav, and M. A. Lavadkar, "AI-Driven Real-Time Threat Detection for UPI Transactions," in *2025 International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA)*, IEEE, Aug. 2025, pp. 1–3. doi: 10.1109/ICAMIDA64673.2025.11209290.

[22] E. Prayitno, R. Kartadie, U. Lestari, and M. S. Bunga, "Implementation of AI-Based Support Vector Machine Algorithm for Fraud Detection in Digital Payment Systems," in *2024 7th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 2024, pp. 85–89. doi: 10.1109/ISRITI64779.2024.10963478.

[23] D. Dahiphale *et al.*, "Enhancing Trust and Safety in Digital Payments: An LLM-Powered Approach," in *2024 IEEE International Conference on Big Data (BigData)*, IEEE, Dec. 2024, pp. 4854–4863. doi: 10.1109/BigData62323.2024.10825105.

[24] Y. R. M. R, K. A, R. D, R. Reshma, D. R. Santhosh, and N. Mekala, "An Analytical Approach to Fraudulent Credit Card Transaction Detection using Various Machine Learning Algorithms," in *2023 Second International Conference on Electronics and Renewable Systems (ICEARS)*, 2023, pp. 1400–1404. doi: 10.1109/ICEARS56392.2023.10085157.

[25] S. Irfan, "Identification of Financial Fraud Transactions: A Cybersecurity via Machine Learning Methods," in *2026 IEEE International Conference on AI Engineering and Innovations (AIEI)*, NIT Jamshedpur, India: IEEE, 2026, pp. 1–6, May. doi: 10.1109/AIEI69164.2026.11497127.

[26] 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.

[27] R. Udayakumar, A. Joshi, S. S. Boomiga, and R. Sugumar, "Deep Fraud Net: A Deep Learning Approach for Cyber Security and Financial Fraud Detection and Classification," *J. Internet Serv. Inf. Secur.*, vol. 13, no. 4, pp. 138–157, 2023, doi: 10.58346/JISIS.2023.I4.010.

[28] F. Khaled Alarfaj and S. Shahzadi, "Enhancing Fraud Detection in Banking with Deep Learning: Graph Neural Networks and Autoencoders for Real-Time Credit Card Fraud Prevention," *IEEE Access*, vol. 13, no. August 2024, pp. 20633–20646, 2025, doi: 10.1109/ACCESS.2024.3466288.

[29] B. I. Hameed, M. A. Mohammed, and H. K. Yaseen, "Predicting Fraud: A Machine Learning Approach to Secure Transactions in Credit Card System," *J. Theor. Appl. Inf. Technol.*, vol. 103, no. 14, 2025.

