International Journal of Emerging Research in Engineering and Technology
Pearl Blue Research Group| Volume 6 Issue 4 PP 102-110, 2025
ISSN: 3050-922X | https://doi.org/10.63282/3050-922X.IJERET-V6I4P113

Orignal Article

Integrating Explainable AI in Financial Fraud Detection Systems for Enhanced Decision Transparency

Sandeep Gupta[1], Ruhul Quddus Majumder[2]
[1]SATI, Vidisha.
[2]Independent Researcher.

Received On: 10/09/2025 ***Revised On: 23/09/2025*** ***Accepted On: 16/10/2025*** ***Published On: 06/11/2025***

Abstract - *The high rate of development of digital financial services has increased the vulnerability of financial systems to fraud. Despite the effect of traditional machine learning models on fraud detection, it is often not flexible and transparent to the evolving financial landscape and regulatory requirements. This research paper is a description of a new way of detecting financial fraud and is more effective than the previous one. It uses the Kaggle credit card data that comprises over 284,000 transactions yet only 492 frauds. The data set was preprocessed in different ways, like label encoding, Min–Max normalization, PCA-based feature selection, SMOTE balancing, and other methods, and then the models were trained. Two significant models, Decision Tree and Multilayer Perceptron, were developed and compared with the existing models (GBM, ANN, LR, NB) on the measurements of accuracy (ACC), precision (PRE), recall (REC), and F1-score (F1). MLP performed optimally with the highest accuracy of 99.52% and the DT model performed optimally in recall and F1-score since it could effectively detect fraudulent cases. SHAP and LIME were the keys to explaining the determinants of the model processing by the use of explanatory AI technologies. Overall, the results revealed that the accuracy of predictions and reproducibility are the cornerstones of building effective fraud detection systems in the finance sector, where the real-life context is vital.*

Keywords - *Explainable AI, Fraud Detection, Financial Security, Artificial Intelligence, PCA, Machine Learning.*

1. Introduction

The rapid advancement of financial transactions as e-commerce and digital money transfer has revolutionized the financial systems of the modern world. Such innovations have dramatically helped in improving corporate management, reduced the cost of operation, and the productivity in general [1][2]. As more businesses and organizations begin to embrace electronic financial transactions, the transition to digital versions has dramatically changed the way financial transactions are carried out [3]. This online revolution has, however, placed financial systems in novel and developed dangers as well, especially, with cybercrimes and fraudulent endeavors [4]. Financial fraud threatens the stability of the world economy in addition to creating enormous losses for individuals, governments, and businesses [5][6]. Bank account fraud is among other types of financial fraud that is more challenging to detect and prevent. In comparison with CCF [7], where abnormal transactional behavior can be easily detected early, bank account fraud is more covert, as it can be conducted by illegal transfers of funds, account hijacking, or identity theft, resulting in the establishment of new fraudulent accounts [8][9].

The unapproved use of an account to carry out transactions without the knowledge of the bank authorities or the account's actual owner is known as credit card fraud (CCF). the necessity of taking the appropriate safety measures when conducting these transactions in order to prevent these frauds. Furthermore, the bank authorities need to use state-of-the-art technology to foresee these robberies and alert their customers beforehand. Financial institutions are now using AI to apply in a wide range of solutions, algorithmic trading, customer care and portfolio optimization [10]. The capacity of AI systems to assess massive amounts of organized and unstructured data has changed the nature of data-driven decision-making in banking.

Predictive models powered by AI are also able to identify complex, non-linear relationships that would otherwise go unnoticed by conventional statistical tools in data which results in more precise forecasting [11], more transparent decisions and better risk management. Also, the AI-based fraud detection methods are not transparent and are black-box. In such a serious application like bank fraud detection, there must be the requirement that the AI system be accurate and reliable [12]. The suggested based banking fraud detection. fraud detection technique is proposed to address the concern on the basis of the integrate Explainable AI (XAI) frameworks [13][14]. Most of the financial institutions today are operating centralized machine learning (ML) systems, where they train their own models using proprietary data to identify potentially fraudulent behavior. The financial industry has continued to favor this centralized approach because of its effectiveness in analyzing a collection of data and identifying the underlying patterns in the transactions.

1.1. Motivation and Contribution

This research is motivated by the rise and advancement of financial frauds in the digital age that is threatening organizations, financial institutions, and customers in a major way across the globe. The old systems based on rules are sometimes incapable of identifying intricate and dynamic

fraudulent trends resulting in huge losses. Therefore, there is a dire need to have a smart, data intensive solution that capable of detecting fraudulent transactions on the spot. Through its application of ML algorithms and useful feature selection algorithms such as PCA, this paper create a powerful and efficient fraud detection model that can work with big, unbalanced datasets and enhance the overall detection ACC using ensemble learning. This study contributes in a number of ways as enumerated below:

- An actual dataset of CCF was obtained on Kaggle to create the most practical and realistic estimation.
- The data has undergone an elaborate preprocessing pipeline such as label encoding, Min max normalization, PCA and SMOTE to handle the enhancement of the performance of fraud detection.
- Developed the suggested DT and MLP to take advantage of their complementary characteristics, leading to improved fraud detection ACC, stability, and generalization.
- Evaluated the model on a set of overall evaluation measures, such as ACC, PRE, REC, F1, and ROC-AUC to provide a complete and credible performance evaluation.
- The use of explainable AI methods, including LIME and SHAP, led to the interpretation of the model decisions and enhancement of the aspect of transparency.
- Provided a systematic workflow and methodology applicable to real-world financial fraud detection and other anomaly detection domains.

1.2. Justification And Novelty

The requirement for reliable fraud detection systems that can manage complicated and skewed financial transaction data served as the impetus for this project. The traditional approaches tend to rely on one model and limited number of features, which may lead to reduced detection rate and loss of real cases of fraud. The current solution to these issues is the use of PCA to extract features and SMOTE to balance classes. This innovation is in the fact that the models of DT and MLP are adopted that demonstrate high ACC, stability, and excellence in the real world. More so, the usage of LIME along with SHAP improves the interpretability thus, making the model more transparent and making it more amicable to use in financial practice.

1.3. Organization of the Paper

The paper on Fraud Detection is structured in the following manner: Section II provides the related work on financial fraud detection and ML, the dataset and suggested model are described in Section III, the experiment's results are presented in Section IV in terms of comparative performance analysis, and the study's conclusion is presented in Section V along with a summary of its key conclusions and recommendations for future research.

2. Literature Review

A critical assessment and examination of significant research papers performed on the issue of Financial Fraud Detection Systems were conducted to enlighten and improve the basis of this research. A summary of the recent works in this area is provided in Table I below which states the proposed models, data sets employed, main findings and the challenges faced.

Azad et al. (2025) utilized to optimize the model's parameters and enhance prediction performance through hyperparameter optimization. They were predicted using a soft voting ensemble method using three machine learning models: RF, KNN, and DT. Each model underwent independent training. With an outstanding ACC of 99.96% and an F1 of 99.84%, the ensemble model outperformed all individual classifiers. SHAP was created in an effort to increase the model's transparency by comprehending each feature's contribution and increasing the transparency of the decision-making process [15]. The predictions Sariat et al. (2025) Using a dataset of fictitious mobile money transactions based on PaySim, a simulator that collects actual financial data from an African country's mobile money provider. used and assessed RF, LR, and XGBoost, three popular ML algorithms. Three ML models were put through rigorous testing to see how successfully they identified fraud: (i) The RF Classifier proved to be the most dependable classifier, with an outstanding AUPRC of 0.9998 [16].

Kasoju and Vishwakarma (2024) trained on the artificial PaySim mobile money transaction dataset, takes advantage of the adaptive property of DRL to constantly train and refreeze its fraud detection models as transactional patterns vary. Simultaneously, the use of XAI ensures transparency and fosters trust by offering understandable insights into the decision-making process. With an ACC of 99.5% and an F1 of 99.0%, the suggested model performs better than conventional fraud detection methods [17]. The aim Dhasaratham et al. (2024) processed to perform data scaling and remove the duplicates. The data is sampled using SMOTE for solving data imbalance issues. The relevant features are extracted using Residual Network (ResNet) followed by the feature selection by ABIF method. The final EML based trained FFD model is used for detecting the frauds which resulted in the overall detection ACC 99.76%, PRE of 98.64%, REC of 98.02%, and f1 of 97.45% which are superior when compared to the existing FFD models namely RUS+XGB and LR-RF [18].

Rallapalli, Hegde and Thatikonda (2023) offer a novel method for identifying financial fraud by fusing the bio-inspired optimization strategy with a two-stage ESVM. To guarantee data balance, the dataset is initially preprocessed using the Bird Mating Optimization Algorithm (BMOA) in three phases. Lastly, throughout the fraud detection process, classification is done using Latent Variable SVM (LV-SVM) and Least Square SVM (LS-SVM). Additionally, a comparison study has been carried out between the proposed method and the existing methodologies. The suggested approach successfully identified fraudulent transactions with an ACC of 98% when compared to the most advanced techniques [19].

Maurya and Kumar (2022) require financial firms to regularly update and enhance their models. Credit card transaction fraud was identified using machine learning techniques, albeit real-time data processing might be challenging. To increase the model's effectiveness and ACC, blockchain technology and machine learning are being combined. Using the Ethereum dataset, the suggested approach use machine learning techniques to detect and prevent fraudulent transactions. For the given dataset, XGBoost has the greatest ACC (99.21%) of all classifiers [20]. Thus Islam et al. (2022) The ACC was displayed using XGBoost classifiers with SmOTE, GbM, Baseline logistic regression, Adaptive Synthetic Sampling Method (AdAsYn), LR, and LgBm. Finally, this research study reminds that 0.1731% of the dataset under study is fraud. The results and findings are supported by actual transactional data provided by a major European card processing company [21].

Research gaps: Even though ML and ensemble-based methods of fraud detection have made major progress, there are still gaps in research that have not been filled. The majority of the current literature is based on artificial or stagnant data, and thus it is not applicable to the dynamic real-life financial setting. Although these models like the RF, XGBoost and NNs have high ACC, they may not be very interpretable and may not be flexible to the changing fraud trends. Furthermore, few approaches effectively integrate real-time detection capabilities with explainable AI frameworks. Optimization techniques and hybrid models have shown promise, yet their computational complexity and scalability issues persist. Thus, intelligent, understandable, and adaptable fraud detection systems that can efficiently manage enormous volumes of real-time financial data are desperately needed.

Table 1: Recent Studies on Financial Fraud Detection Systems using Machine Learning Techniques

Author	Proposed Work	Results	Key Findings	Limitations & Future Work
Azad et al. (2025)	Addressed data imbalance using SMOTETomek, optimized model parameters with GridSearchCV and PSO, and combined DT, KNN, and RF using a soft voting ensemble.	Accuracy: 99.96%, F1-Score: 99.84%	Soft voting ensemble outperformed individual classifiers; SHAP used for explainability and transparency.	Could explore larger datasets and real-time fraud detection for scalability.
Sariat et al. (2025)	XGBoost, LR, and RF were used to create a fraud detection framework for mobile money transactions utilizing PaySim data.	AUPRC: 0.9998 (Random Forest best performer)	Feature engineering (error balance calculations) enhanced fraud detection accuracy.	Future work could involve real-time detection and deep learning integration.
Kasoju and Vishwakarma (2024)	Integrated Deep Reinforcement Learning (DRL) with Explainable AI (XAI) for adaptive fraud detection on PaySim dataset.	Accuracy: 99.5%, F1-Score: 99.0%	DRL improved adaptability; XAI increased interpretability and trust.	Could be extended for large-scale, real-time applications and multi-platform use.
Dhasaratham et al. (2024)	Combined ABIF with Ensemble ML (RF + AdaBoost) after SMOTE balancing and feature extraction via ResNet.	Accuracy: 99.76%, Precision: 98.64%, Recall: 98.02%, F1-Score: 97.45%	Outperformed RUS+XGB and LR-RF models; effective in feature selection and fraud detection.	Future work may involve testing on real-world datasets and optimizing computational efficiency.
Rallapalli, Hegde & Thatikonda (2023)	Proposed a two-stage Ensemble SVM (ESVM) optimized using Bird Mating Optimization Algorithm (BMOA) for balanced data classification.	Accuracy: 98%	BMOA improved data balancing and ESVM enhanced classification accuracy.	Could integrate other bio-inspired algorithms and deep learning approaches.
Maurya & Kumar (2022)	Combined Blockchain with ML for secure fraud detection using Ethereum dataset; compared multiple ML models.	Accuracy: 99.21% (XGBoost best performer)	Blockchain improved transaction security; XGBoost achieved highest accuracy.	Future work could focus on scalability and integration with live financial systems.
Islam et al. (2022)	Compared multiple ML models (Logistic Regression, XGBoost, GbM) using SMOTE and ADASYN on real-world European card transaction data.	Fraud represented 0.1731% of dataset; improved performance with resampling.	Sampling and optimization enhanced model accuracy on highly imbalanced data.	Could explore ensemble or deep models for improved precision and recall.

3. Research Methodology

The Financial Fraud Detection methodology proposed and depicted in Figure 1, starts with the CCFD dataset and then moves through the data preprocessing steps of the dataset by removing the unnecessary identifiers and applying label encoding. Min-Max scaling, where PCA is used for selection. To undo the effects of class imbalance, SMOTE produces artificial samples, and subsequently, the data is split into training and testing sets. Finally, DT and MLP models are trained, and their performance is measured by ACC, PRE, REC, F1, and ROC, allowing for obtaining trustworthy results for detecting fraudulent cases.

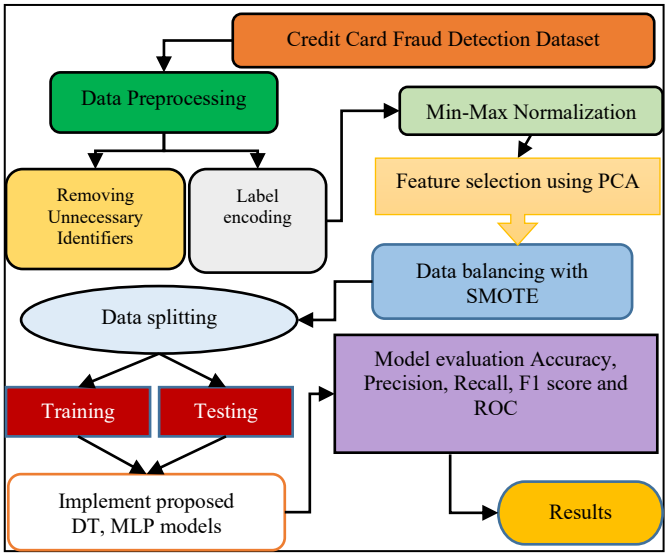

Fig 1: Proposed flowchart for Financial Fraud Detection Systems using machine learning

3.1. Data Gathering and Analysis

A publicly accessible CCF detection dataset on Kaggle provided the dataset used in this investigation. The dataset includes credit card transactions from European cardholders in September 2013. Only transactions that took place within the last two days are included. 492 of the 284,807 transactions were determined to be fraudulent. Every transaction was categorized as authentic or fraudulent. The type of card used, the transaction amount, and the transaction time are among the 31 attributes in the dataset. The following data visualization heatmaps were utilized to look at feature correlations, attack dispersion, etc:

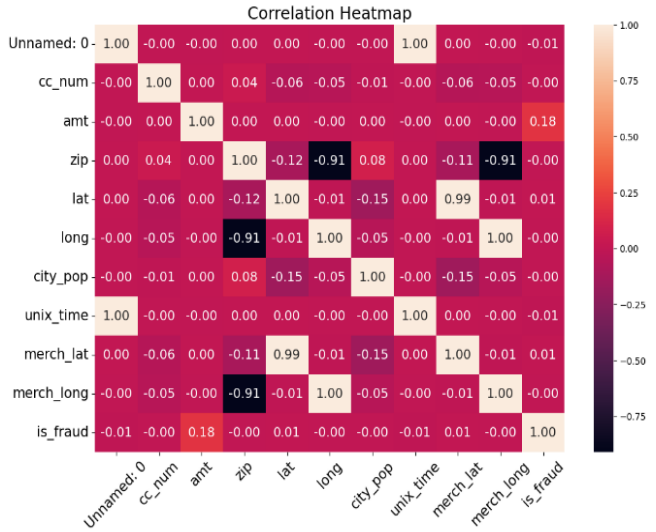

Fig 2: Correlation Matrix Heatmap for Financial Fraud Detection

The correlation heatmap depicted in Figure 2 gives us insight into the interrelation among the various features in the dataset. The majority of the features are almost independent of each other since their correlations are nearly zero. However, there are two robust correlations: the zip code is negatively correlated with latitude and merchant latitude, which is quite logical as the fields related to location are bound to be interdependent. The fraud indicator is not reliant on any particular variable to a great extent but does demonstrate a marginal positive correlation with the transaction amount.

3.2. Data Pre-Processing

The CCF Detection Dataset was utilized for data preparation, which included concatenation, data cleansing, and feature engineering. The preprocessing phase involved removing irrelevant identifiers, and performing data leveling and normalization. The summarization of the major preprocessing processes is as follows:

3.2.1. Removing Unnecessary Identifiers

Eliminating superfluous identifiers entails removing or changing both directly identifiable information, like names and addresses, and indirectly identifying information, such certain dates or geographic details. Various approaches can be applied, including eliminating personal information on a search page, de-identifying a dataset to conduct research, or removing unnecessary structure identifiers in a database.

3.2.2. Label Encoding

Label encoding is a data preparation technique that assigns a unique number to each category, converting categorical data into numerical data.

3.3. Min-Max Normalization

The records were normalized with the MinMax scaling technique to limit all the values of the features in the range 0 to 1. This was carried out to enhance classifier performance and reduce the impact of outliers. The mathematical expression below was used in normalizing (Equation 1):

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

The feature's initial value is represented by X, its minimal value is represented by X_{min}, its maximum value is represented by X_{max}, and $X'$is normalized.

3.4. Feature selection using PCA

The selection of the features was used to determine and retain the most important attributes that can help to detect fraud correctly and remove redundant or irrelevant features. This boosts the efficiency of the models, minimizes overfitting, and the performance of the predictors. According to the PCA characteristics, the majority of them have a conventional normal distribution with a standard deviation of one and a mean of zero. However, certain features include outliers or skewed distributions that could impact how well particular algorithms work. Figure 3 highlights the most important features found using PCA. The transaction amount is the strongest indicator of fraud, followed by transaction time and merchant location. In contrast, details like zip code and credit card number add little value. Consequently, significant clues of fraud can be the points at which funds were spent, the extent of the resources spent and the time frame.

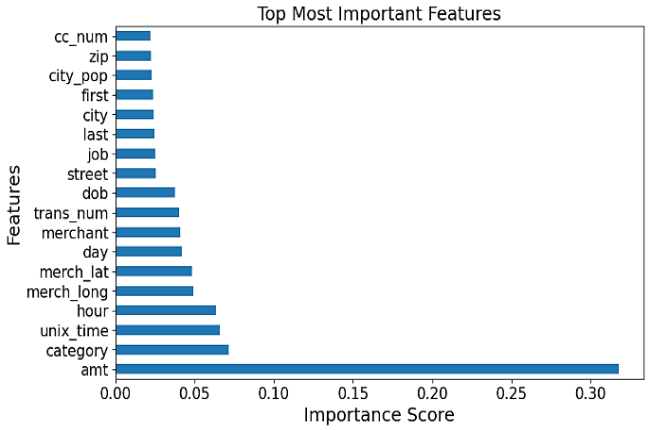

Fig 3: Bar graph of the PCA Features

3.5. Data Balancing using SMOTE

Data balancing is a preprocessing step used to resolve class imbalance in datasets. It ensures that each class has a comparable number of samples, minimizing model bias. Typical strategies include oversampling minority courses and under sampling majority classes in order to improve ACC, stability, and equity. SMOTE produces a more balanced dataset and improves model generalization by generating fresh instances for the minority class rather than replicating the existing data.

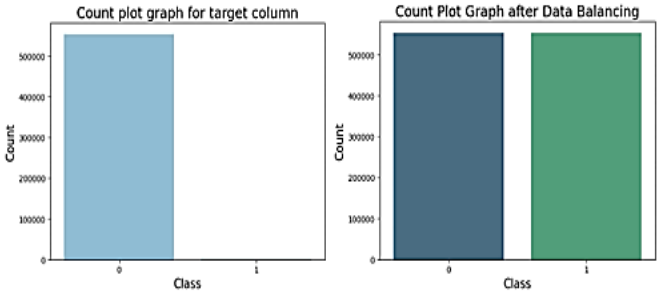

Fig 4: Bar Graph of Class Distribution before and After SMOTE

The SMOTE, a method for dataset balancing, caused the alteration in class distribution as depicted in Figure 4. The original data set was largely non-fraud cases that the model would have trouble learning if fraud were to be detected. With SMOTE in place, both classes are at 50% which helps the model significantly in detecting the fraud cases. This balancing factor is the most important for boosting ACC and eliminating bias during the training process.

3.6. Data Splitting

The original class distribution was preserved by splitting the data into training and testing subsets at an 80:20 ratio. The best intrusion detection model was determined by evaluating its performance on the testing set after predictive models were developed using the training data.

3.7. Financial Fraud Detection Using DT and MLP Models

The analysis in this part focuses on the Decision Tree (DT) and Multilayer Perceptron (MLP) models performance concerning financial fraud detection. Their classification metrics, which are the most important ones, are used to assess the models and finally to conclude about their practical usage.

3.7.1. Decision Tree (DT)

A DT is a supervised machine learning technique for classification and regression. In order to create a tree-like structure, it groups the data recursively based on feature values. Internal nodes indicate an attribute choice, branching communicates the decision's conclusion, and leaf nodes reflect a final prediction [22]. The measures used in the algorithm to pick splits include the Gini Index, Information Gain or Entropy and this is to ensure that the splits produced are the purest. Decision Trees are easy to understand, less data preprocessing is needed and they are able to capture a complex nonlinear relationship in data.

The splitting criterion in mathematical form is expressed in Information Gain (IG) in Equation (2).

$$IG(D, A) = Entropy(D) - \sum_{v \in Values(A)} \frac{|D_v|}{|D|} \times Entropy(D_v)$$

In which D is the dataset, A is the attribute to be used to split the dataset, D_v is the subset of D where attribute A has value v and Entropy (D) is a measure of the impurity of the data. The systematic application of this process enables the Decision Tree model to acquire an interpretable structure that

is effective at separating different classes; thus, useful in intrusion detection and other classification problems.

3.7.2. Multilayer Perceptron (MLP)

An MLP is a kind of feedforward artificial neural network that consists of an input layer, one or more hidden layers, and an output layer. Each neuron in a layer is connected to every other layer's neuron, and each connection's weight is changed during training to reduce prediction error [23]. The MLP looks for nonlinear relationships between input variables and output goals using nonlinear activation functions like Tanh, Sigmoid, and ReLU. It is trained using the backpropagation approach, and the gradient of the loss function is utilized to modify its weights. MLPs can be used to perform pattern recognition, regression and classification tasks with great efficiency.

Equation (3) is the mathematical expression of the neuron output in an MLP.

$$y = f(\sum_{i=1}^{n} w_i x_i + b)$$

The input features are x_i, where f() is the activation function, b is the bias term, and w_i is the weight. The MLP model is able to learn hierarchical feature representations on numerous layers making it an effective model in learning nonlinear patterns and can be used in more complex scenarios in intrusion detection as well as other predictive modeling tasks.

3.8. Evaluation Metrics

Validation ACC, PRE, REC, and F1 were used to evaluate the classification performance of the proposed model because they are optimal metrics to evaluate models on a balanced dataset. Their performance was also measured based on standard performance indicators after the training and evaluation of different ML models using baseline algorithms. These basic parameters are used to come up with complete performance measures, which are mathematically expressed in Equations (4) to (7).

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (4)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (6)$$

$$\text{F1} - \text{score} = 2 * \frac{(\text{precision} * \text{recall}}{(\text{precision} + \text{recall})} \quad (7)$$

ACC determines the fraction of right predictions; therefore, it is applicable to balanced data sets. The positive predicted value tells the number of predicted positives that are actually correct, and the sensitivity reveals the number of real positives that have been detected. PRE and REC have been combined into one measure called the F1-score, which, in the case of imbalanced classes, can be considered a balanced metric. ROC curves demonstrate the interdependent relation between true and false positives. Overall performance of classification can be summed up by AUC. The model is made easier to understand by applying XAI techniques like SHAP and LIME. SHAP use Shapley values to quantify the contribution of the input characteristics and provide a deeper understanding of the model decision-making process, while LIME uses local surrogate models to explain specific predictions.

4. Results and Discussion

The experimental setup and performance assessment of the suggested model for training and testing are covered in this section, with a focus on its computational effectiveness and efficiency. Google Colab and a desktop computer equipped with an Intel Core i7-5500U processor and 16 GB of RAM were used for the research. The findings indicate that the models are equally in fraudulent activity detection. And the Table II illustrate the MLP takes the lead over DT in ACC (99.79%) and PRE (99.16%), whereas DT has the upper hand in REC (99.40%) and F1 (99.78%). the MLP turns out to be a better predictor, whereas the DT model is more skilled in revealing fraudulent cases and managing PRE and REC all of which are pluses for explainable AI in situations where decision-making is critical.

Table 2: Classification results of the proposed model, Financial Fraud Detection Systems using CCFD Dataset

Measures	DT	MLP
Accuracy	98.79	99.52
Precision	99.16	99.59
Recall	99.40	98.44
F1-score	99.78	99.52

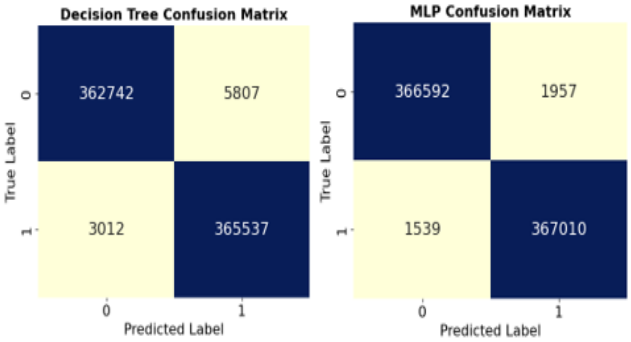

Fig 5: Confusion Matrix for Proposed DT Model

The confusion matrices of the Decision Tree and MLP models for fraud detection are depicted in Figure 5. The Decision Tree classifier was able to correctly identify all 362,742 verified non-fraud incidents and 365,537 fraud cases, but the MLP classifier was not only able to do so, but it also topped the Decision Tree by the correct classification of 366,592 non-fraud cases and 367,010 being fraud. Moreover, the MLP model had a lower count of misclassifications, making 1,539 wrongful predictions while the DT model was responsible for 3,012 errors. These results demonstrate the MLP model's superiority in identifying patterns in the data and, consequently, its greater fraud detection ACC.

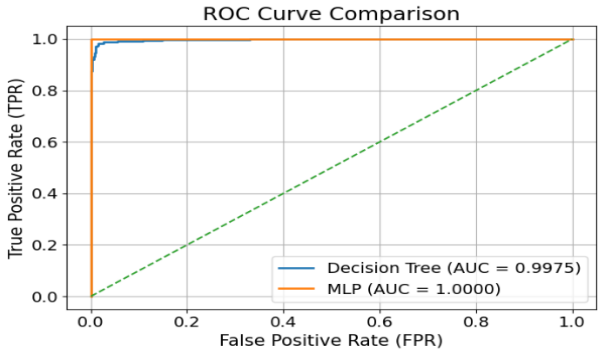

Fig 6: ROC Curve Comparison between Decision Tree and MLP

Figure 6 shows a comparison of ROC curves for the DT and MLP models. The MLP model excels among the two by receiving a perfect AUC score of 1.0000 which indicates it nearly faultlessly differentiates between fraud and non-fraud. The DT also scores high with an AUC of 0.9975 however MLP's curve still gets nearer to the ideal top-left corner implying there are lesser FP and overall prediction quality is better. This highlights the MLP model as the more reliable choice for fraud detection.

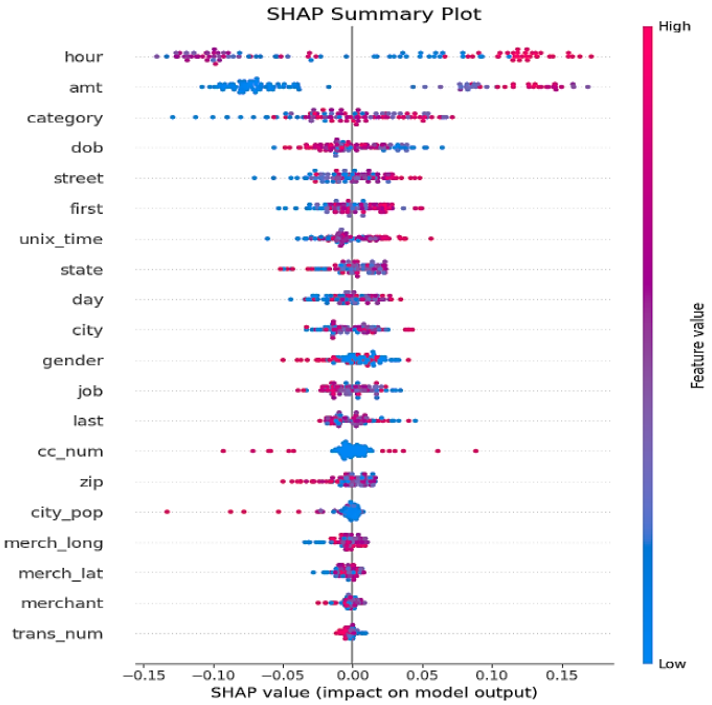

Fig 7: SHAP Summary Plot

The global average influence of various characteristics on the size of the model's output for a binary classification task (Class 0 and Class 1) is displayed in Figure 7's SHAP Summary Plot. The features' overall significance to the model is indicated by ranking them according to their average absolute SHAP value (x-axis). With the biggest average influence on the model's forecast magnitude for both classes and a marginally higher impact on Class 1 (red) predictions, the feature ratio_to_median_purchase_price is by far the most significant. The next most important features are distance_from_home and online_order, whilst used_pin_number and repeat_retailer have a negligible average influence on the model's output magnitude.

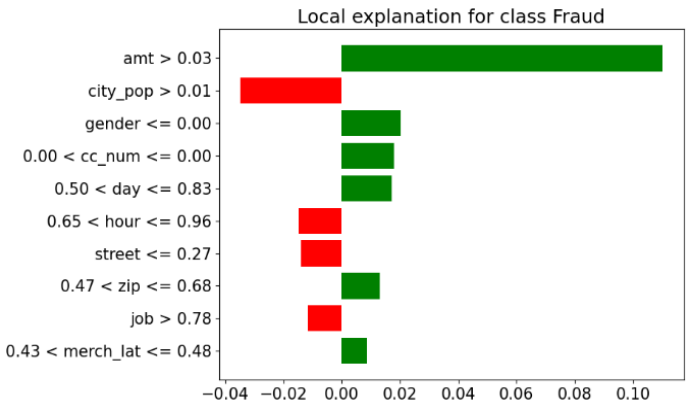

Fig 8: LIME Explanation for Fraud Prediction

The local explanation for fraud prediction is illustrated in Figure 8 and it shows the influence of certain factors on a particular choice. The transaction amount being over 0.03 is made clear and it is indicated that the fraud probability is considerably higher in this case. On the other hand, features

like city population and street information reduce the chance of it being classified as fraud. Other details such as time of transaction, zip range, and merchant location also play smaller roles.

4.1. Comparative Analysis

The performance of various ML models in the field of financial fraud detection is hinted at in Table III. The models GBM, ANN, LR, and NB exhibit inconsistencies in their findings on REC and F1, especially the actual fraud cases, despite their strengths in the ACC component. On the other hand, the results of the proposed DT and MLP models are much stronger and balanced. The ACC of the proposed DT reaches 98.79 and the PRE and the REC are almost perfect, whereas the MLP proposed is more accurate with 99.52% ACC and reliable REC rates. These findings suggest that the proposed models are unsuitable for real-world fraud detection applications since they are less predictable and reasonable.

Table 3: Comparison of Different ML and DL Models for Financial Fraud Detection Systems

Model	Accuracy	Precision	Recall	F1-score
GBM[24]	95	91	87	89
ANN[25]	96.14	95.74	96.55	95.50
LR [26]	95.87	89.11	60	71.71
NB[27]	98.1	84.9	6.83	12.65
Proposed DT	98.79	99.16	99.40	99.78
Proposed MLP	99.52	99.59	98.44	99.52

The suggested DT model, has the benefit of being easily interpretable, so that the paths of the decisions can be easily understood and the fraud can be easily detected. On the other hand, because the proposed MLP model can recognize and record nonlinear correlations in the data, it provides a stronger prediction capacity. These models are complementary to one another since they are interpretable and very accurate in classifications.

5. Conclusion and Future Study

The increase in credit card fraud, which causes yearly losses of billions of dollars, poses a major challenge to the security and trust of electronic payment systems. The CCF has been an immense deterrent against the security of the financial systems, hence why detection models that are both accurate and trustworthy are of paramount importance. This research made use of a real Kaggle dataset to assess different machine learning models including GBM, ANN, LR, NB as means of effectively identifying fraudulent transactions. Amongst all, the proposed MLP model had the highest performance with an ACC measure of 99.52% whereas the DT model was credited for its strong interpretability and thus it did score 99.78% on F1 being that it was able to detect actual fraud cases with reliability. In spite of these results being indicative of significant improvements over the previous models (GBM: 95%, ANN: 96.14%, LR: 95.87%, NB: 98.1%), there are still some issues like having a dataset of only two days, a lack of real-time evaluation and cross-domain validation. The overall findings indicate that the suggested models especially MLP possess numerous potentials in actual application within financial fraud detection systems without compromising transparency, strength, and dependability. Future studies might concentrate on the assessment of the models using bigger and real-time financial datasets and checking their versatility in various domains. The most sophisticated ML and DL methods like RF, XGBoost, LSTM, CNN, and hybrid ensemble models will be the ones to experiment with in order to significantly improve detection ACC, stability, and robustness in the stressful case of frauds.

References

[1] K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," 2021. doi: 10.1016/j.cosrev.2021.100402.

[2] R. Q. Majumder, "A Review of Anomaly Identification in Finance Frauds Using Machine Learning Systems," *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 101–110, Apr. 2025, doi: 10.48175/IJARSCT-25619.

[3] H. Kali, "Optimizing Credit Card Fraud Transactions Identification and Classification in Banking Industry Using Machine Learning Algorithms," *Int. J. Recent Technol. Sci. Manag.*, vol. 9, no. 11, pp. 85–96, 2024.

[4] A. Parupalli, "Business Intelligence in ERP ML-Based Comparative Study for Financial Forecasting," *ESP Int. J. Commun. Eng. Electron. Technol.*, vol. 2, no. 4, pp. 17–26, 2024, doi: 10.56472/25839217/IJCEET-V2I4P103.

[5] S. J. Wawge, "A Survey on the Identification of Credit Card Fraud Using Machine Learning with Precision, Performance, and Challenges," *Int. J. Innov. Sci. Res. Technol.*, vol. 10, no. 4, pp. 3345–3352, May 2025, doi: 10.38124/ijisrt/25apr1813.

[6] K. Lee and D. Choi, "An Artificial Intelligence Approach to Financial Fraud Detection under IoT Environment : A Survey and Implementation," *Secur. Commun. Networks*, 2018.

[7] V. Verma, "Deep Learning-Based Fraud Detection in Financial Transactions: A Case Study Using Real-Time Data Streams," *ESP J. Eng. Technol. Adv.*, vol. 3, no. 4, pp. 149–157, 2023, doi: 10.56472/25832646/JETA-V3I8P117.

[8] S. Farsi and M. Chowdhury, "EcomFraudEX: An Explainable Machine Learning Framework for Victim-Centric and Dual-Sided Fraud Incident Classification in E-Commerce," *ICST Trans. Scalable Inf. Syst.*, vol. 12, 2025, doi: 10.4108/eetsis.6789.

[9] S. Jagdish, M. Singh, and V. Yadav, "Credit Card Fraud Detection System: A Survey," *J. Xidian Univ.*, vol. 14, no. 5, pp. 5498 – 5505, May 2020, doi: 10.37896/jxu14.5/599.

[10] H. P. Kapadia, "API-Driven Banking: How COVID-19 Remote Work Boosted Open Banking and Fintech Integrations," *J. Emerg. Technol. Innov. Res.*, vol. 8, no. 10, pp. f514–f519, 2021.

[11] S. B. Shah, "Advancing Financial Security with Scalable AI: Explainable Machine Learning Models for

Transaction Fraud Detection," in *2025 4th International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, 2025, pp. 1–7. doi: 10.1109/ICDCECE65353.2025.11034838.

[12] B. R. Ande, "Federated Learning and Explainable AI for Decentralized Fraud Detection in Financial Systems," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 35s, pp. 48–56, 2025.

[13] N. Malali, "Exploring Artificial Intelligence Models for Early Warning Systems with Systemic Risk Analysis in Finance," in *2025 International Conference on Advanced Computing Technologies (ICoACT)*, IEEE, Mar. 2025, pp. 1–6. doi: 10.1109/ICoACT63339.2025.11005357.

[14] P. Hou *et al.*, "Technology and practice of intelligent governance for financial data security," *Chinese J. Netw. Inf. Secur.*, 2023, doi: 10.11959/j.issn.2096-109x.2023048.

[15] M. A. K. Azad, A. B. M. Y. Arafat, A. K. M. Masum, Y. Islam, M. M. Hassan, and D. M. Farid, "An Optimized Ensemble Learning Framework for Credit Card Fraud Detection with Explainable AI," in *2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN)*, 2025, pp. 1–6. doi: 10.1109/QPAIN66474.2025.11171906.

[16] 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.

[17] A. Kasoju and T. chary Vishwakarma, "Leveraging Explainable AI and Reinforcement Learning for Enhanced Transparency in Adaptive Fraud Detection," in *2024 IEEE 8th Conference on Energy Internet and Energy System Integration (EI2)*, 2024, pp. 103–108. doi: 10.1109/EI264398.2024.10991389.

[18] M. Dhasaratham, Z. A. Balassem, J. Bobba, R. Ayyadurai, and S. M. Sundaram, "Attention Based Isolation Forest Integrated Ensemble Machine Learning Algorithm for Financial Fraud Detection," in *2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)*, 2024, pp. 1–5. doi: 10.1109/IACIS61494.2024.10721649.

[19] S. Rallapalli, D. Hegde, and R. Thatikonda, "Feature Selection Based Ensemble Support Vector Machine for Financial Fraud Detection in IoT," in *2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques, EASCT 2023*, 2023. doi: 10.1109/EASCT59475.2023.10392566.

[20] A. Maurya and A. Kumar, "Credit card fraud detection system using machine learning technique," in *Proceedings - 2022 IEEE International Conference on Cybernetics and Computational Intelligence, CyberneticsCom 2022*, 2022. doi: 10.1109/CyberneticsCom55287.2022.9865466.

[21] M. B. Islam, C. Avornu, P. K. Shukla, and P. K. Shukla, "Cost Reduce: Credit Card Fraud Identification Using Machine Learning," in *7th International Conference on Communication and Electronics Systems, ICCES 2022 - Proceedings*, 2022. doi: 10.1109/ICCES54183.2022.9835811.

[22] S. Patil, V. Nemade, and P. K. Soni, "Predictive Modelling for Credit Card Fraud Detection Using Data Analytics," in *Procedia Computer Science*, 2018. doi: 10.1016/j.procs.2018.05.199.

[23] W. Priatna, H. D. Purnomo, A. Iriani, I. Sembiring, and T. Wellem, "Optimizing Multilayer Perceptron with Cost-Sensitive Learning for Addressing Class Imbalance in Credit Card Fraud Detection Wowon," *Decree Dir. Gen. High. Educ. Res. Technol.*, vol. 8, no. 158, pp. 2–9, 2024.

[24] A. Tomy and I. P. Ojo, "Explainable AI for credit card fraud detection: Bridging the gap between accuracy and interpretability," *World J. Adv. Res. Rev.*, vol. 25, no. 2, pp. 1246–1256, Feb. 2025, doi: 10.30574/wjarr.2025.25.2.0492.

[25] H. Hajiyev, E. Hajiyev, M. Avezov, S. Makhmudov, D. Abdukhalikova, and E. L. Lydia, "An Explainable AI-based Fraud Detection System Using Recursive Feature Elimination and Waterwheel Plant Optimization for Financial Transactions," *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 5, pp. 28114–28119, 2025, doi: 10.48084/etasr.13350.

[26] M. N. Alatawi, "Detection of fraud in IoT based credit card collected dataset using machine learning," *Mach. Learn. with Appl.*, vol. 19, Mar. 2025, doi: 10.1016/j.mlwa.2024.100603.

[27] E. Ileberi, Y. Sun, and Z. Wang, "A Machine Learning Based Credit Card Fraud Detection Using The GA Algorithm For Feature Selection," *J. Big Data*, vol. 9, no. 24, Dec. 2022, doi: 10.1186/s40537-022-00573-8.