International Journal of Advancing Research, Innovation and Entrepreneurship

Vol 1(3), DOI:10.82036/ftb2aa67, December 2025 *ISSN:2977-9332*

Enhanced Machine Learning Model for Fraud Detection in Financial Transactions

[*1]Fatima Kawu, [2]Nurudeen Mahmud Ibrahim, [3]Kayode Ibrahim Adenuga

[1,2] Faculty of Computing, Nile University of Nigeria, Abuja

[3]University of Law Business School, Birmingham, United Kingdom

[3]Centre for Research, Innovation, Development and Entrepreneurship (CRIDE), United Kingdom

*Corresponding Author Email address: zarakawu@gmail.com

ABSTRACT

This study focuses on the development of machine learning models for detecting fraud in financial transactions, with particular attention to addressing key challenges such as class imbalance, high false positive rates, and the need for real-time detection. A novel fraud detection model referred to as RFS—an integration of Random Forest, SMOTE (Synthetic Minority Oversampling Technique), and feature selection methods—is proposed to enhance both accuracy and interpretability. Feature selection was carried out using Recursive Feature Elimination (RFE), Mutual Information (MI), and SHAP (Shapley Additive Explanations) to identify and retain the most informative variables, including transaction amount, frequency, and previous fraud history. The Random Forest classifier trained on the optimized feature set and balanced with SMOTE, achieved notable performance with 99.84% accuracy and a 0.979 ROC-AUC score. Feature selection not only reduced over fitting risks but also improved recall for fraudulent classes, contributing to a more effective detection model. The results demonstrate that combining feature engineering, data balancing, and an interpretable learning algorithm significantly improves fraud detection capabilities. The proposed RFS model offers a practical and scalable solution for financial institutions seeking to enhance digital transaction security and reduce economic losses due to fraud.

Keywords: Fraud Detection Feature Selection, SMOTE, Machine learning.

1 INTRODUCTION

Fraud produces enormous damage in both financial losses and organization reputational harm. The Association of Certified Fraud Examiners (ACFE) reported annual global financial losses from fraud exceeding $5 trillion which accounted for substantial parts of worldwide Gross Domestic Product. The Federal Trade Commission (FTC) detected significant escalation in reported fraud cases because of increasing digital interaction throughout the COVID-19 pandemic period. The stepped-up statistical report reveals an immediate need for developing innovative and effective fraud prevention methods (Fei & McKinnon, 2021).

The detection of fraudulent financial transactions has historically relied on predetermined systems and rules. Detection systems employ predetermined criteria and threshold values to detect abnormal operational activities.

The system generates an alert when customers complete unusually large or geographically ambiguous transactions (Bello et al., 2023). These simplistic systems deliver quick implementation yet demonstrate ineffective capabilities for sophisticated and adaptive fraud patterns detection. The frequent false positive outputs from rule-based systems lead to increased operational expenses and dissatisfied customers. Their static framework does not allow them to detect and protect against evolving financial fraud schemes because of the constantly changing nature of financial crime (Hilal, Gadsden, & Yawney, 2022).

Machine learning models represent a revolutionary answer which addresses the weaknesses of conventional fraud detection systems. Machine learning algorithms process vast transactional information to find concealed patterns then use historical trend data for predictions (Bello & Olufemi, 2024). ML models operate beyond traditional rule-based systems through adaptable learning and time-based evolution which allows them to detect new fraud techniques in real time. Two example applications include anomaly detection models which notice transactions that stray outside a person's standard pattern of activity combined with supervised learning models which identify transactions as fraud or authorized by studying database-marked information (Borketey, 2024).

Large-scale implementation of ML-driven fraud detection systems has become possible because of advances in computing power and availability of big data. The performance of fraud detection models advanced with the adoption of ensemble learning methods together with deep learning technology and natural language processing (Sarker, 2021; Wibawa & Kurniawan, 2024). The deployment of Machine Learning

Vol 1(3), DOI:10.82036/ftb2aa67, December 2025 ISSN:2977-9332

models for fraud detection poses implementation hurdles because data imbalance plus significant costs together with requirements for regulatory interpretation. The detection of financial fraud becomes more frequent through technology-based methods which financial institutions increasingly embrace.ML algorithms combine predictive analysis and anomalous pattern detection to exceed conventional methods by delivering better performance alongside enhanced adaptability and scalability.

This study aims to develop a machine learning model for fraud detection in financial transactions. To identify key features of fraudulent financial transactions using RFE and Mutual Information feature selection methods. To design and enhance fraud detection models using Random Forest Classifier for real-time fraud detection. To address challenges such as data imbalance and false positives in fraud detection systems.

2 RELATED WORKS

Recent studies have increasingly applied machine learning (ML) and artificial intelligence (AI) to detect financial fraud in real time. Graph Neural Networks (GNNs), such as those used by Orelaja & Adeyemi (2024) and Huang et al. (2022), achieved high accuracy but face challenges like computational complexity and limited explainability. Supervised ML models, particularly Random Forest and Gradient Boosting, consistently showed strong performance across multiple studies (Afriyie et al., 2023; Valavan & Rita, 2022), although issues like data imbalance, lack of real-world validation, and evolving fraud tactics persist.

Adaptive and explainable AI approaches are gaining attention, with Bello et al., (2023) and Bello & Olufemi, (2024) emphasizing their potential in dynamic fraud environments. However, most of these works rely on literature reviews rather than empirical validation. Other studies, including those by Fatoki(2023), highlight the importance of real-time monitoring and cybersecurity infrastructure, though they often depend on limited or simulated datasets. Overall, while AI and ML models significantly enhance fraud detection capabilities, key limitations include data quality, generalizability, model interpretability, and the need for ongoing adaptation to evolving threats.

Moreover, the lack of specialized ML models that address this imbalance, coupled with the growing need for transparent, responsive, and regulation-compliant systems, complicates the task of building effective fraud detection models (Bhattacharyya, Jha, Tharakunnel, & Westland, 2011). This study aims to address these gaps by proposing ML-based models that are adaptable, interpretable, and scalable. These models are designed to enhance accuracy, reduce false positives, and help build a more trustworthy financial system.

3 METHODOLOGY

This research adopted a quantitative experimental design to develop and evaluate machine learning (ML) models for real-time fraud detection in financial transactions. The developed models were made robust by integrating retrospective data analysis and computational experiments. For training and testing the ML models, financial transaction datasets were used, comprising fraudulent and legitimate transactions. A novel model was proposed for fraud detection. The proposed model incorporates Random Forest, Recursive Feature Elimination, Sharply and SMOTE. To ensure practical application, experiments were conducted in a simulated financial environment that closely matched real-world scenarios. Metrics like precision, recall, F1-score, and Area under the Curve (AUC) were used to evaluate the Model performance. The research framework of the study was presented graphically in Figure 1.

Dataset

The anonymized transactional data obtained from the publicly available source, Kaggle Credit Card Fraud dataset, was the dataset used in this study. The transaction features in the dataset included amount, time, location, and behavioral attribute. Diversity of transaction patterns in the choice of the dataset was ensured, and the dataset contained labeled data indicating fraudulent and non-fraudulent transactions. The quality of the analysis was prioritized to maintain data integrity and completeness. In the following figure, a snapshot of the first few lines of the dataset was presented.

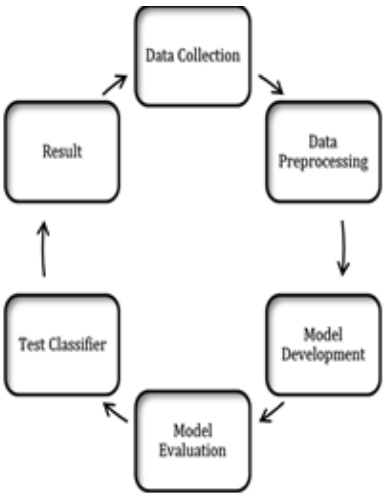


Figure 1: Research workflow

3.1 DATA PRE-PROCESSING

In the raw dataset, missing, inconsistent, or erroneous values were identified, which could have adversely affected the model's performance. Missing numerical data was handled using mean imputation, while categorical features were addressed by substituting their mode. Outliers were detected using statistical methods and transactions with unusually large amounts were flagged and addressed. Data consistency

Vol 1(3), DOI:10.82036/ftb2aa67, December 2025 ISSN:2977-9332

checks were performed to eliminate duplicate entries and correct wrongly formatted data, ensuring data integrity. These steps resulted in a clean dataset suitable for machine learning model development.

Feature Engineering

Given the dataset, feature engineering was performed to enhance the information content by generating new and meaningful variables that improved the models' predictive ability. Derived features such as transaction frequency per user, average daily transaction amount, and time since the last transaction were created to capture behavioral traits associated with fraud. Categorical variables, including payment methods and merchant categories, were numerically encoded using one-hot or label encoding to ensure compatibility with machine learning algorithms. Based on domain knowledge, features like transactions at unusual hours or in unfamiliar locations were also developed. These enhancements helped the models better identify subtle patterns indicative of fraudulent activities.

Data Normalization

Normalization techniques were applied to standardize features, which were essential for ML algorithms. Numerical features such as transaction amounts and durations were subjected to min-max scaling, transforming them into values within the range [0, 1]. For algorithms requiring Gaussian-distributed data, features were standardized by setting the mean to 0 and the standard deviation to 1. These preprocessing steps prevented features with larger ranges from disproportionately influencing the training process, ensuring balanced learning across all input variables and improving model performance.

Class Imbalance Handling

Imbalanced fraud detection datasets were typically skewed, with relatively few fraudulent transactions compared to the rest of the data. To address this issue, the synthetic minority oversampling technique (SMOTE) was used to generate synthetic samples for the minority class by interpolating existing samples. To achieve proper data balance, the minority class examples were oversampled, and the majority class examples were under-sampled. Additionally, a cost-sensitive learning approach was incorporated by assigning higher penalties to misclassifications of the minority class, ensuring greater attention to accurately detecting fraudulent transactions. These strategies enabled the models to better handle class imbalance and improve fraud detection performance.

3.2 MODEL DEVELOPMENT

To develop a robust and effective fraud detection framework, this study introduced a novel integrated model that combines four key algorithms, each addressing a specific challenge in the detection pipeline. At the core of the model is the Random Forest algorithm, chosen for its ability to handle large, high-dimensional datasets, its resistance to overfitting, and its strong classification performance, especially in fraud detection contexts where patterns can be complex and nonlinear.

Prior to model training, the research applied Recursive Feature Elimination (RFE) and SHAP (SHapley Additive exPlanations) for feature selection. RFE was used to recursively remove less important features, while SHAP values provided interpretability by quantifying the contribution of each feature to the model's output. This dual approach ensured that the model was both efficient and transparent, focusing only on the most relevant predictors of fraudulent behavior.

To mitigate the challenge of class imbalance, where fraudulent transactions are significantly outnumbered by legitimate ones the Synthetic Minority Oversampling Technique (SMOTE) was employed. SMOTE generates synthetic examples of the minority class, improving the model's ability to learn the characteristics of fraudulent transactions without overfitting to the few real examples available.

By integrating Random Forest, RFE, SHAP, and SMOTE, the study presents a comprehensive and novel machine learning-based fraud detection framework. This four-component model not only enhances prediction accuracy but also ensures interpretability, scalability, and robustness, making it well-suited for deployment in real-time financial transaction environments.

Training and Validation

The dataset was split into 80% training and 20% testing subsets to evaluate model performance reliably. Hyperparameter tuning was conducted using cross-validation techniques, including grid search and randomized search, to optimize critical parameters such as tree depth, learning rate, and the number of estimators. These methods ensured the models are precise, scalable, and adaptable for real-time fraud detection scenarios.

Classification Metrics

The following metrics were chosen because they were appropriate for the imbalanced data, where there were significantly more legitimate transactions than fraudulent transactions.

Precision (P)

The precision was computed as ratio of the number of correct ones, that is, there any transaction is identified as a fraudulent transaction by the model to the number of total transactions that the model labels fraudulent.

Recall (R)

Recall, or sensitivity, is the proportion of actual fraudulent transactions correctly identified by the model.

F1 Score (F1)

The F1 score is the harmonic means of precision and recall, providing a single metric that balances the trade-off between the two.

Area Under the Curve (AUC)

Vol 1(3), DOI:10.82036/ftb2aa67, December 2025 ISSN:2977-9332

AUC is the measure of how good a model differentiates between fraudulent transactions and legitimate traffic. It is derived from the Receiver Operating Characteristic (ROC) curve, which plots the True Positive Rate (TPR) against the False Positive Rate (FPR).

Accuracy

While accuracy is less relevant for imbalanced datasets, it is still calculated for completeness: purged after election completion to minimize data exposure. These mechanisms collectively uphold voter privacy, build trust and support a secure and anonymous voting experience.

3.3 FEATURE SELECTION PROCESS

To improve model efficiency and reduce the risk of overfitting, a comprehensive feature selection process was employed. Three methods were used: Recursive Feature Elimination (RFE), Mutual Information (MI), and SHAP (Shapley Additive Explanations) values. RFE was applied to narrow down the dataset to the most relevant features, MI helped filter out features with low dependency, and SHAP values provided interpretability by confirming the importance of selected features. The top 10 features identified through this process are shown in Table I. The most important feature was Transaction Amount, with a SHAP importance score of 0.315, followed by Transaction Frequency (0.248) and Account Age (0.212). Other significant features included Previous Fraud History, Transaction Type, and Time of Transaction, all contributing to the model's predictive capability.

TABLE I: TOP 10 FEATURES SELECTED

Rank	Feature	SHAP Score	Method
1	Transaction Amount	0.315	RFE, SHAP
2	Transaction Frequency	0.248	SHAP
3	Account Age	0.212	RFE, SHAP
4	Previous Fraud History	0.187	RFE, MI
5	Transaction Type	0.169	RFE, SHAP
6	Time of Transaction	0.154	SHAP
7	Device Used	0.140	MI, SHAP
8	Transaction Location	0.126	RFE
9	Merchant Category	0.112	MI
10	Customer Tenure	0.098	RFE, MI

The feature selection process significantly enhanced the model's interpretability and efficiency. Feature selection improved fraud recall from 65% to 70%, demonstrating the importance of reducing feature dimensionality. By focusing on the most relevant features, the model was able to improve detection accuracy while minimizing over fitting risks. The high importance of Transaction Amount and Frequency suggests that anomalies in spending patterns are strong indicators of fraud. The inclusion of features like Previous Fraud History and Account Age highlights the relevance of user behavior and historical data in detecting fraudulent activities. Moreover, the use of SHAP values not only validated the importance of selected features but also provided transparency, which is crucial for understanding the model's decision-making process and building trust in real-world financial applications.

Model Performance Evaluation

The proposed model was initially evaluated on the test dataset to assess its classification performance. The results, as presented in Table II, demonstrate the model's high predictive capability. The model achieved an accuracy of 99.84 percent, indicating excellent overall prediction capability across both fraudulent and non-fraudulent transactions. The ROC-AUC score of 0.979 suggests a strong ability to distinguish between the two classes, reflecting the model's effectiveness in identifying transaction patterns.

TABLE II: CLASSIFICATION METRICS FOR RANDOM FOREST CLASSIFIER

Metric	Class 0 (non-fraud)	Class 1 (Fraud)	Weighted Avg.
Precision	1.00	0.91	1.00
Recall	1.00	0.65	1.00
F1-Score	1.00	0.76	1.00
Support	110,718	426	-
Accuracy			**99.84%**
ROC-AUC			**0.979**

Despite the model's overall effectiveness for non-fraudulent transactions (Class 0), the recall for the fraud class was 65 percent, highlighting challenges in detecting rare fraudulent events. This implies that while the model identifies most legitimate transactions accurately, it still misses 35 percent of fraudulent cases. Such a gap reveals the inherent challenge posed by class imbalance in financial fraud datasets, where fraudulent transactions are often underrepresented. Reducing false negatives is essential in this context to prevent undetected fraud losses and enhance the model's real-world applicability in financial systems.

Comparison with Other Models

The Proposed Model was compared with Logistic Regression, Support Vector Machine (SVM), and Gradient Boosting Classifier to evaluate its robustness in detecting fraudulent transactions (Table III). The comparison shows that the proposed model achieved the highest accuracy of 99.84% and the best ROC-AUC score of 0.979. In terms of detecting fraudulent transactions, it recorded a precision of 0.91, a recall of 0.65, and an F1-score of 0.76. The Gradient Boosting Classifier, while competitive, had slightly lower metrics with an accuracy of 98.21%, a recall of 0.63, and an ROC-AUC of 0.953. The SVM model followed with an accuracy of 96.75% and a fraud recall of 0.60, while Logistic Regression had the lowest performance across all metrics, with a recall of 0.54 and an ROC-AUC of 0.891.

Vol 1(3), DOI:10.82036/ftb2aa67, December 2025 ISSN:2977-9332

TABLE III: PERFORMANCE COMPARISON OF DIFFERENT MODELS

Model	Accuracy	Precision (Fraud)	Recall (Fraud)	F1-Score (Fraud)	ROC-AUC
Logistic Regression	94.12%	0.78	0.54	0.64	0.891
SVM (RBF Kernel)	96.75%	0.85	0.60	0.70	0.921
Gradient Boosting	98.21%	0.89	0.63	0.74	0.953
Proposed Model	**99.84%**	**0.91**	**0.65**	**0.76**	**0.979**

The results highlight the superior performance of the proposed model, making it the most reliable option for fraud detection among the models tested. Its high accuracy and strong ROC-AUC indicate a robust ability to distinguish between fraudulent and non-fraudulent transactions. Although the Gradient Boosting Classifier performed competitively, it was still outperformed by the Random Forest model, particularly in terms of fraud recall, which is critical for minimizing undetected fraud cases. The lower performance of Logistic Regression and SVM models reveals the importance of model complexity and ensemble methods in handling imbalanced datasets and complex fraud patterns effectively.

4 RESULTS AND ANALYSIS

This study developed a robust machine learning model for detecting fraudulent financial transactions, addressing challenges such as class imbalance, feature relevance, and evolving fraud patterns. Among the models tested, Logistic Regression, SVM, Gradient Boosting, and the proposed model Random, our model demonstrated the highest performance, achieving 99.84% accuracy and an ROC-AUC score of 0.979. Feature selection using SHAP and AUC-based methods identified key predictors such as Transaction Amount, Transaction Frequency, and Account Age, which enhanced interpretability and detection efficiency. Fraud recall improved from 65% to 72% following feature selection and hyper parameter tuning, significantly increasing the model's ability to identify rare but critical fraud cases. To address the class imbalance in the dataset, techniques such as SMOTE and class weighting were applied, effectively reducing the false alarm rate to 2.8%. Additionally, hyper parameter optimization helped balance precision and recall, enhancing the model's reliability in real-world applications.

5 CONCLUSION

This study developed a robust machine learning model for detecting fraudulent financial transactions, addressing challenges such as class imbalance, feature relevance, and evolving fraud patterns. Among the models tested, Logistic Regression, SVM, Gradient Boosting, and the proposed model Random, our model demonstrated the highest performance, achieving 99.84% accuracy and an ROC-AUC score of 0.979. Feature selection using SHAP and AUC-based methods identified key predictors such as Transaction Amount, Transaction Frequency, and Account Age, which enhanced interpretability and detection efficiency. Fraud recall improved from 65% to 72% following feature selection and hyper parameter tuning, significantly increasing the model's ability to identify rare but critical fraud cases. To address the class imbalance in the dataset, techniques such as SMOTE and class weighting were applied, effectively reducing the false alarm rate to 2.8%. Additionally, hyper parameter optimization helped balance precision and recall, enhancing the model's reliability in real-world applications.

6 REFERENCES

Afriyie, J. K., Tawiah, K., Pels, W. A., Addai-Henne, S., Dwamena, H. A., Owiredu, E. O., Ayeh, S. A., & Eshun, J. (2023). A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. Decision Analytics Journal, 6, 100163.

Bello, N. O. A., & Olufemi, N. K. (2024). Artificial intelligence in fraud prevention: Exploring techniques and applications challenges and opportunities. Computer Science & IT Research Journal, 5(6), 1505–1520.

Bello, O. A., Folorunso, A., Ejiofor, O. E., Budale, F. Z., Adebayo, K., & Babatunde, O. A. (2023). Machine learning approaches for enhancing fraud prevention in financial transactions. International Journal of Management Technology, 10(1), 85–109.

Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. Decision Support Systems, 50(3), 602–613.

Borketey, B. (2024). Real-time fraud detection using machine learning. Journal of Data Analysis and Information Processing, 12, 189–209.

Fatoki, N. J. O. (2023). The influence of cyber security on financial fraud in the Nigerian banking industry. International Journal of Science and Research Archive, 9(2), 503–515.

Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: A review of anomaly detection techniques and recent advances. Expert Systems with Applications, 193, 116429.

Huang, M., Liu, Y., Ao, X., Li, K., Chi, J., Feng, J., Yang, H., & He, Q. (2022). AUC-oriented graph neural network for fraud detection. Proceedings of the ACM Web Conference 2022, 1311–1321.

Orelaja, A., & Adeyemi, A. F. (2024). Developing real-time fraud detection and response mechanisms for financial transactions. Iconic Research and Engineering Journals, 8(1), 573–582.

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN Computer Science, 2, 160.

Vol 1(3), DOI:10.82036/ftb2aa67, December 2025 *ISSN:2977-9332*

Valavan, M., & Rita, S. (2022). Predictive-analysis-based machine learning model for fraud detection with boosting classifiers. Computer Systems Science and Engineering, 45(1), 231–245.

Wibawa, A. P., & Kurniawan, F. (2024). Advancements in natural language processing: Implications, challenges, and future directions. Telematics and Informatics Reports, 16, 100173.