Development of an Enhanced Financial Fraud Detection Model using a Machine Learning Algorithm

[1]Prince Uchenna Sundayn; Prof. J.S. Igwe; [2]Nwali Monday Ekpe, [1]Chinaza Joy Ojimadu

[1]Ebonyi State University, Abakaliki

[2]Alex Ekwueme Federal University, Ndufu Alike, Ebonyi State

DOI: https://doi.org/10.51584/IJRIAS.2026.11060002

Received: 20 May 2026; Accepted: 25 May 2026; Published: 16 June 2026

ABSTRACT

This study presents an enhanced financial fraud detection system that uses unsupervised machine learning techniques to identify anomalous financial transactions in highly imbalanced datasets. The increasing sophistication of digital financial fraud has reduced the effectiveness of traditional rule-based security systems, thereby necessitating adaptive and intelligent detection approaches. The proposed system uses the Isolation Forest algorithm for its computational efficiency and ability to detect previously unseen fraud patterns without relying on labelled data. A large transaction dataset comprising 284,807 financial records was preprocessed with normalization and dimensionality-reduction techniques prior to model training and evaluation. The developed model was integrated into a Flask-based web application that supports real-time transaction analysis via a user-friendly dashboard. Experimental results demonstrated strong predictive performance, achieving a recall of 94.7%, a precision of 53.5%, an F1 Score of 68.6%, and an ROC-AUC of 0.91. Comparative evaluation with One-Class SVM and Autoencoder models further revealed that while Autoencoders achieved slightly higher overall discrimination performance, Isolation Forest provided a more computationally lightweight and deployment-efficient solution for resource-constrained financial environments. The findings also highlight the practical implications of false positive management in fraud detection systems and emphasize the importance of balancing fraud sensitivity with operational efficiency. Overall, the study confirms that lightweight unsupervised anomaly detection models provide an adaptive, scalable, and deployable solution for enhancing financial security in modern fintech ecosystems.

Keywords: Anomaly Detection, Isolation Forest, Financial Fraud, Machine Learning, Fintech, Real-time Monitoring.

INTRODUCTION

In the contemporary era of cashless economies and digitized financial ecosystems, the proliferation of online transactions has accelerated commerce and driven financial inclusion in developing nations. However, this digital shift has concurrently introduced sophisticated cybersecurity threats, transforming financial fraud into a global systemic risk that threatens institutional stability and erodes consumer trust (Bello et al., 2023). In the Nigerian context, the trajectory of electronic fraud reflects a critical vulnerability within the rapidly growing FinTech sector. According to the Nigeria Inter-Bank Settlement System (NIBSS, 2024), despite the implementation of regulatory frameworks such as the Bank Verification Number (BVN), the volume of fraudulent transactions continues to rise, resulting in annual losses of billions of Naira that often outpace central bank interventions (Fatokun et al., 2025).

A primary challenge is the over-reliance of many domestic institutions on traditional, rule-based fraud detection engines, which are largely static and struggle to identify "zero-day" fraud patterns or evolving attack vectors (Enehikhare & Odumuyiwa, 2025). Machine Learning (ML) offers a transformative approach by analyzing vast, high-dimensional datasets to detect subtle, hidden patterns that are imperceptible to human auditors. Specifically, unsupervised anomaly detection offers a promising line of defense by establishing a baseline of "normal"

ISSN No. 2454-6194 | DOI : 10.51584/IJRIAS |Volume XI Issue VI June 2026

behavior and flagging rare deviations as risks without requiring prior fraud labels (Onyeama, 2024). This adaptability is crucial in the Nigerian financial domain, where transaction volumes are high and user behaviour is diverse across channels such as USSD and Agency Banking.

Despite these theoretical advantages, a significant gap remains in the practical implementation of anomaly detection systems tailored to the unique metadata, network latencies, and behavioural patterns of Nigeria's digital economy, as most of the literature relies strictly on Western datasets (Odufisan et al., 2025). Furthermore, extreme "class imbalance," in which legitimate transactions vastly outnumber fraudulent ones, remains a persistent operational hurdle to effective local deployment (Enehikhare & Odumuyiwa, 2025). Consequently, this study explores the design and deployment of an unsupervised machine learning–based anomaly detection system optimized for context-aware financial transactions. By bridging the gap between global algorithmic standards and local implementation challenges, this research demonstrates how anomaly-based frameworks can serve as a robust foundation for the next generation of financial security in emerging economies.

Related Works

These studies collectively highlight the shift toward adaptive, interpretable, and scalable fraud detection systems. The integration of unsupervised learning and anomaly detection has proven particularly valuable in identifying novel fraud patterns, reducing false positives, and enhancing real-time monitoring capabilities.

Building on the strong empirical performance reported earlier, the need for effective fraud detection has become more urgent as financial services increasingly shift to digital channels. The expansion of online banking, mobile wallets, agency banking, and instant payment platforms has accelerated transaction speed and broadened financial inclusion. Still, it has also enlarged the attack surface for fraudsters. In this environment, fraudulent activity has become more automated, adaptive, and harder to distinguish from legitimate behaviour, especially as transactions occur at high volume and near real time (Bello et al., 2023; Odufisan et al., 2025). For banks, fintech firms, and payment processors, the challenge is no longer simply to detect known fraud templates, but to identify subtle deviations that may signal emerging attack strategies before substantial losses occur.

Traditional fraud detection systems were designed around expert rules, fixed thresholds, and manual review processes. Although such systems are transparent and straightforward to implement, they are increasingly mismatched to the pace and complexity of modern digital fraud. Their dependence on predefined patterns makes them vulnerable to novel or rapidly evolving tactics, including “zero-day” fraud schemes that do not resemble previously observed behaviour (Bolton & Hand, 2002; Enehikhare & Odumuyiwa, 2025). In addition, rigid rule sets often generate excessive false positives, disrupting legitimate transactions and increasing operational burden for compliance teams and customers. As transaction volumes grow, the limitations of static heuristics become more pronounced, particularly in environments where fraud patterns shift quickly, and the cost of missing a single case is high.

Machine learning offers a more adaptive alternative because it can infer structure directly from transactional data rather than relying exclusively on hand-crafted rules. Within this broader family, anomaly-based learning is especially relevant to fraud detection because fraud is typically rare, asymmetric, and poorly represented in labelled datasets. Unsupervised anomaly detection models learn the statistical profile of normal behaviour and flag observations that deviate from that baseline, making them suitable for datasets dominated by legitimate transactions (Chandola, Banerjee, and Kumar, 2009; Liu, Ting, and Zhou, 2008). This is particularly important in fraud analytics, where class imbalance remains a central methodological obstacle: fraud may account for only a small fraction of all records, yet the minority class carries the greatest financial risk. Supervised methods can struggle under these conditions, especially when labelled fraud examples are sparse, delayed, or incomplete. Anomaly detection, therefore, provides a practical approach to identifying suspicious activity in settings where exhaustive labels are unavailable or continually change.

A further gap in the literature concerns the transferability of fraud models across regions. Much of the benchmark research and many widely cited datasets originate from Western payment ecosystems, especially European or North American card environments, where customer behaviour, merchant structures, network conditions, and

ISSN No. 2454-6194 | DOI : 10.51584/IJRIAS |Volume XI Issue VI June 2026

regulatory regimes differ from those in African markets. This creates a mismatch between model development and operational deployment in contexts such as Nigeria, where digital payment adoption is rapid but infrastructural constraints, transaction diversity, and fraud typologies are not always reflected in foreign datasets (NIBSS, 2024; Odufisan et al., 2025). In addition, the scarcity of locally curated and publicly available fraud datasets limits the ability to train, validate, and benchmark models against region-specific threat patterns. Consequently, there is a need for lightweight, scalable, and context-aware detection systems that can operate effectively in data-scarce and highly imbalanced environments while remaining suitable for practical deployment in emerging financial ecosystems.

Against this background, the present study aims to develop an enhanced machine learning–based fraud detection model using anomaly detection principles, with particular emphasis on the Isolation Forest algorithm. The study was designed to evaluate whether an unsupervised approach can achieve high fraud sensitivity while remaining computationally efficient and deployable via a web-based interface. Specifically, it sought to examine transactional patterns through exploratory data analysis, implement an anomaly-detection pipeline to identify rare fraud, and assess performance using metrics appropriate for imbalanced classification. The study is grounded in anomaly detection theory, which conceptualizes fraudulent transactions as low-probability outliers relative to normal transaction distributions, and in the broader premise that adaptive machine learning can provide a more resilient first line of defense than static rule-based systems. In doing so, it addresses both the technical challenge of rare-event detection and the practical need for fraud solutions that are responsive to the realities of Nigerian financial operations.

Building on the regional transferability issue noted above, prior research shows that fraud detection methods have progressed from static rules to data-driven anomaly detection. However, their suitability remains highly contingent on data structure and operating context. Early systems relied on hand-crafted thresholds and expert rules, which were valued for transparency but were inherently rigid and prone to high false-positive rates as fraud tactics evolved (Bolton & Hand, 2002; Ngai et al., 2011). Subsequent work shifted toward supervised machine learning, including decision trees, random forests, and cost-sensitive classifiers, which improved predictive performance on labelled datasets but still depended heavily on representative fraud labels and careful class balancing (Bhattacharyya et al., 2011; Sahin et al., 2013; Bahnsen et al., 2016).

As the fraud problem became increasingly imbalanced, unsupervised anomaly detection gained prominence because it does not require exhaustive fraud labels. In this paradigm, models learn normal transaction structure and flag deviations as suspicious. Isolation Forest has been particularly influential because it efficiently isolates rare points through random partitioning, making it computationally attractive for high-volume transaction streams (Liu et al., 2008). Local Outlier Factor (LOF) offers a complementary neighborhood-based approach by identifying points that are locally sparse relative to their neighbors, although its scalability is limited for very large datasets (Chandola et al., 2009; Ahmed et al., 2016). Autoencoders have also been widely studied, with reconstruction error used as an anomaly score; they are effective at learning nonlinear patterns but often require more tuning and training time than tree-based methods (Fiore et al., 2019).

Recent studies increasingly favour hybrids that combine the strengths of multiple methods. Carcillo et al. (2019) showed that ensembles integrating unsupervised and supervised components can support near real-time fraud monitoring in highly imbalanced settings. Zhou et al. (2020) reported strong performance for an autoencoder–Isolation Forest pipeline, indicating that latent representation learning can improve outlier separation in complex transaction spaces. Devarakonda (2023) similarly found that combining Isolation Forest and autoencoders with downstream classifiers improved recall on difficult fraud cases. These results suggest that hybrid systems may be particularly useful when anomaly structure is nonlinear, labels are scarce, or operational demands require both sensitivity and speed.

The literature also emphasizes that model evaluation in fraud detection cannot rely on accuracy alone. Because legitimate transactions dominate most datasets, a model can achieve seemingly high accuracy while missing most fraud cases. Accordingly, studies increasingly report precision, recall, F1-score, ROC-AUC, and precision-recall curves as more meaningful indicators of utility in imbalanced settings (Dal Pozzolo et al., 2015; Tian et al., 2023). Recall is often prioritized in financial settings because false negatives are costlier than false positives,

ISSN No. 2454-6194 | DOI : 10.51584/IJRIAS |Volume XI Issue VI June 2026

but precision remains important because excessive alerts can disrupt service and burden analysts. This trade-off is especially visible in deployable systems, where operational efficiency, latency, and interpretability matter alongside raw predictive performance (Gandhar et al., 2024; Arora & Bhardwaj, 2025).

Table 1 summarizes the main unsupervised approaches most relevant to this study, highlighting the recurring pattern in the literature: Isolation Forest offers speed and scalability; autoencoders provide richer nonlinear feature learning; and LOF is useful for local irregularities but less suitable for large-scale deployment. Table 2 then summarizes representative studies and their reported strengths, showing that high-performing methods typically combine anomaly sensitivity with pragmatic deployment characteristics.

Table 1: Unsupervised Machine Learning and its Advantages

Method	Core Advantage	Main limitation	Typical use
Isolation Forest	Fast, scalable, label-free	May miss complex structure	Large transactional datasets
Autoencoder	Learns nonlinear representations	Higher tuning/training cost	High-dimensional anomaly detection
LOF	Detects local anomalies	Limited scalability	Smaller or moderate datasets
Hybrid models	Better balance of recall and precision	Greater complexity	Operational fraud pipelines

Table 2: Relevant Studies and their Key Findings

Study	Method	Key finding
Bhattacharyya et al. (2011).	Random forest	Strong baseline performance on imbalanced fraud data
Fiore et al. (2019).	Autoencoder	Effective anomaly isolation without labels
Carcillo et al. (2019).	Hybrid ensemble	Suitable for streaming and real-time detection
Zhou et al. (2020).	Autoencoder + Isolation Forest	Improved recall in highly imbalanced settings
Devarakonda (2023)	Hybrid anomaly pipeline	Demonstrated value of combined anomaly and classification layers

RESEARCH METHODOLOGY

This study adopts an agile prototyping and experimental research design to develop and evaluate an enhanced financial fraud detection system using machine learning techniques. This approach enables a structured examination of existing vulnerabilities in traditional rule-based banking security. It assesses the performance of an unsupervised Isolation Forest model in identifying anomalous transactional patterns within high-dimensional credit card datasets.

Dataset Preparation

The dataset used in this study was the Credit Card Fraud Detection dataset from the ULB Machine Learning Group on Kaggle, comprising 284,807 transactions. To ensure data integrity and model reliability, the dataset underwent comprehensive preprocessing, including the normalization of the 'Amount' feature and temporal

ISSN No. 2454-6194 | DOI : 10.51584/IJRIAS |Volume XI Issue VI June 2026

transformation of the 'Time' variable. The data is characterized by extreme class imbalance, with only 492 fraudulent instances (0.17%) among 284,315 legitimate transactions.

To evaluate the Isolation Forest model's robustness across different scenarios, the dataset was partitioned into training and validation sets. This allowed for a rigorous analysis of the algorithm's ability to isolate anomalies as the volume of training data varied, ensuring the system remains effective at detecting rare fraud events while minimizing false-positive rates in a high-dimensional feature space.

Feature Engineering

To ensure the model learns meaningful patterns, the following pipeline was implemented:

i. Null Value Handling: Verification confirmed the dataset contained no missing values.
ii. Feature Normalization**:** The 'Amount' feature was normalized using Standard Scaler to bring it into the same range as the PCA features (V1-V28).
iii. Feature Transformation: The 'Time' column was converted into transaction hours to capture temporal patterns.
iv. Dimensionality Reduction**:** While PCA was already applied to V1-V28, further inspection ensured the most relevant features were retained for the Isolation Forest model.

Model Selection

The proposed system is built primarily on the Isolation Forest algorithm.

i. Efficiency: It is highly effective for large-scale transactional data.
ii. Unsupervised Nature**:** It does not require labelled data for training, making it ideal for detecting previously unseen fraud patterns.
iii. Isolation Logic**:** It isolates anomalies by randomly selecting features and splitting data, which is faster and more accurate than density-based methods for high-dimensional data.

System Design

The system comprises the following components:

1. Presentation Layer**:** A web interface developed using Flask, HTML, and CSS for user interaction.
2. Application Layer: The Python-based API that integrates the Machine Learning model logic.
3. Data Layer: An in-memory pandas DataFrame for high-speed processing, with a proposed relational database extension for long-term storage.

Web Interface Implementation

The deployed Flask application translates model predictions into a functional user interface. The web dashboard accepts CSV file uploads and validates the transaction columns before executing the pipeline. Once a valid dataset is ingested, the backend loads the serialized Isolation Forest model, applies preprocessing, and displays a responsive prediction table that visually highlights anomalous transactions with alert styling to ensure rapid interpretability.

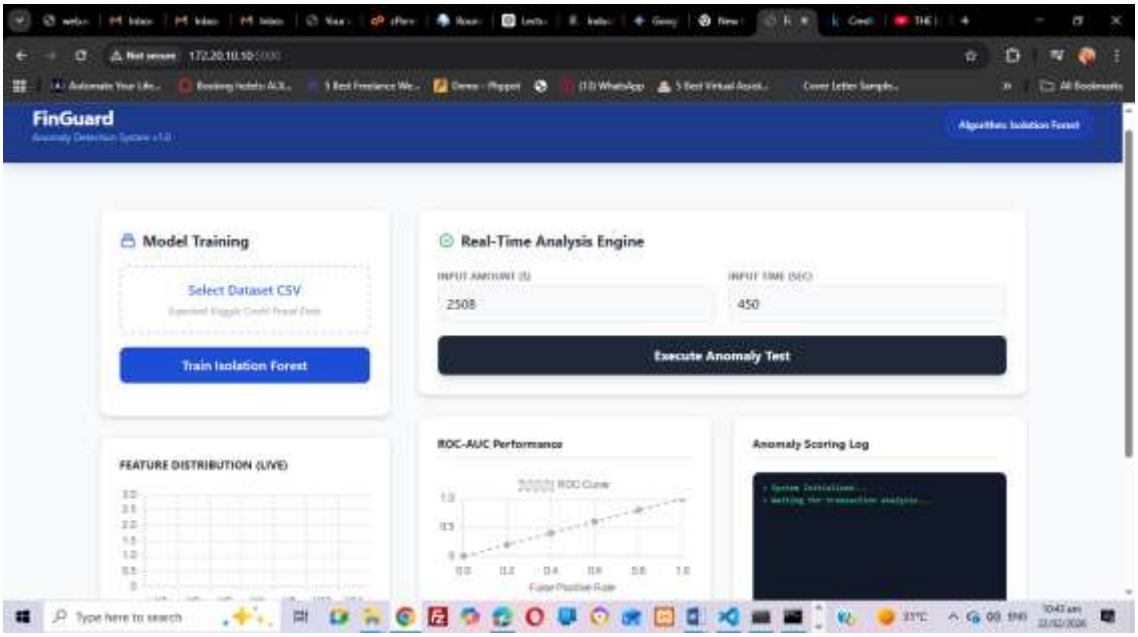
The interface features interactive visual summaries, including a fraud-versus-legitimate distribution bar chart, a ROC performance curve, a live feature distribution chart, and an execution logging panel to maintain transparency. Scenario testing confirmed the workflow's structural robustness: properly formatted files reliably produce accurate classification outputs. At the same time, inputs with missing required fields (such as "Amount") or excessive null values are safely intercepted by built-in error handling, preventing silent failures.

ISSN No. 2454-6194 | DOI : 10.51584/IJRIAS |Volume XI Issue VI June 2026

Data Submission Portal

The user interacts with the system through a web-based dashboard designed for simplicity and efficiency. This portal serves as the primary gateway for uploading transaction records from external databases into the machine learning engine. The layout of this submission interface is presented in Figure 7.

Figure 1: File Upload Interface

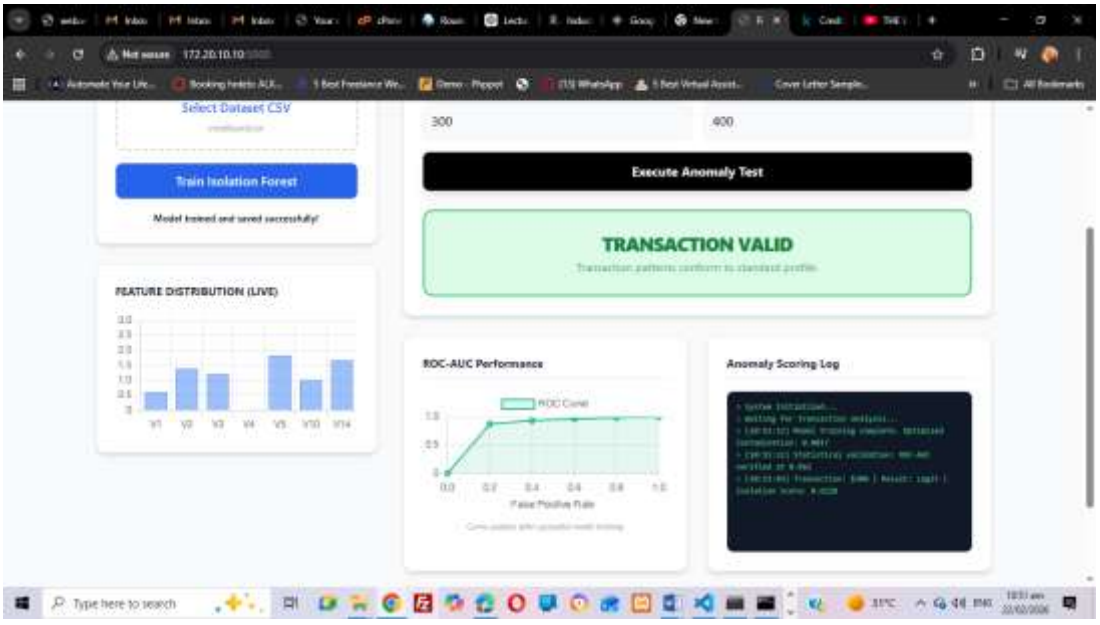


The interface shown in Figure 1 prioritizes user experience by providing a straightforward "Choose File" and "Upload" mechanism. This design ensures that analysts can process bulk transaction data without needing to interact directly with the command-line environment or the underlying Python codebase.

Analysis and Prediction Results

Once the backend processing is complete, the system returns a detailed report to the browser. This report identifies which transactions in the uploaded file were flagged as anomalies by the Isolation Forest algorithm. The resulting display, featuring the highlighted fraud flags, is illustrated in Figure 8.

Figure 2: Result Display Table with Fraud Flags



ISSN No. 2454-6194 | DOI : 10.51584/IJRIAS |Volume XI Issue VI June 2026

As demonstrated in Figure 2, the system uses a high-contrast visual flagging method to distinguish fraudulent entries from legitimate ones. By presenting the findings in a structured table, the application provides investigators with immediate, actionable intelligence, enabling rapid response to suspected financial irregularities.

Model Performance

The confusion matrix showed strong detection of fraudulent transactions but a nontrivial number of false alarms among legitimate transactions. Specifically, the model correctly classified 56,097 legitimate transactions and 90 fraudulent transactions. It incorrectly flagged 78 legitimate transactions as fraud and missed 5 fraudulent transactions. These counts indicate that the model prioritized fraud capture over strict rejection of false positives, which is appropriate for a security-sensitive application where missed fraud can be costlier than extra review.

From these outcomes, the model achieved the following metrics:

i. Precision: 0.535
ii. Recall: 0.947
iii. F1-score: 0.686
iv. ROC-AUC: 0.91

The most notable result was the 94.7% recall, meaning the model detected nearly all fraudulent cases in the test set. This high sensitivity confirms that Isolation Forest was effective at identifying anomalous transaction patterns even in an extremely imbalanced dataset. Precision was lower at 0.535, indicating that a substantial share of flagged transactions was legitimate.

Although the model achieved a high recall of 94.7%, the lower precision indicates false positives, in which legitimate transactions are incorrectly flagged as fraudulent. In practical financial environments, excessive false alarms may affect customer experience, increase operational costs, and require additional manual verification by fraud analysts. However, in security-sensitive applications, prioritizing fraud-detection sensitivity is often preferable because the financial impact of undetected fraud may significantly outweigh the inconvenience of additional transaction reviews. Future enhancements may incorporate adaptive threshold optimization and ensemble learning techniques to reduce false-positive rates while maintaining strong fraud-detection performance.

Comparative Evaluation

The study also compared Isolation Forest with two alternative anomaly-detection approaches. The results showed that although Isolation Forest achieved the strongest recall, the autoencoder achieved slightly better overall discrimination and a better balance between precision and recall.

Table 3: Model Comparison

Model	Precision	Recall	F1-score	ROC-AUC
Isolation Forest	0.535	0.947	0.686	0.91
One-Class SVM	0.471	0.761	0.581	0.83
Autoencoder	0.611	0.923	0.736	0.93

The results indicate that One-Class SVM performed the worst across all reported metrics, especially in recall, where it detected only 76.1% of fraud cases. The autoencoder achieved the best F1 Score (0.736) and the highest ROC-AUC (0.93), but the thesis notes that it required substantially more training time and computational overhead. By contrast, Isolation Forest provided a more efficient deployment option while retaining excellent sensitivity, making it the most practical choice for the prototype system. In effect, the results position Isolation Forest as a computationally lightweight alternative that still captures most fraudulent activity.

ISSN No. 2454-6194 | DOI : 10.51584/IJRIAS |Volume XI Issue VI June 2026

In addition to the evaluated anomaly detection approaches, several supervised machine learning models, including Random Forest, XGBoost, and Logistic Regression, have demonstrated strong performance in fraud classification tasks within existing literature. However, such supervised approaches depend heavily on labelled datasets and may struggle to detect emerging fraud patterns that differ from previously observed behaviour. By contrast, the unsupervised Isolation Forest model used in this study provides greater adaptability for identifying previously unseen anomalies in highly dynamic financial environments. This characteristic makes the model particularly suitable for real-time fraud monitoring systems operating in rapidly evolving fintech ecosystems.

DISCUSSION

The proposed Isolation Forest model achieved a recall of 94.7% and an ROC-AUC of 0.91, demonstrating strong sensitivity to rare fraudulent events and excellent discriminatory power without relying on prior labels. Proper preprocessing, including transaction magnitude normalization and the use of PCA-transformed features, significantly reduced noise and improved transaction separability (Gandhar et al., 2024; Zhang et al., 2019). While complex deep learning models such as autoencoders can achieve only marginal performance gains, they require longer training times and greater infrastructure overhead (Fiore et al., 2019; Devarakonda, 2023). By contrast, this lightweight unsupervised model offers a highly practical balance between predictive sensitivity and operational efficiency.

The model's integration into a Flask-based web prototype transitions the research from offline experimentation to a functional decision-support tool. The simple upload-and-predict architecture supports batch screening and rapid analyst review, demonstrating real-world viability for operational banking workflows (Arora & Bhardwaj, 2025; Odufisan et al., 2025).

This framework is highly relevant to the Nigerian fintech ecosystem, where fragmented data infrastructure, a lack of curated regional fraud datasets, and computational limitations hinder the adoption of massive deep learning pipelines (NIBSS, 2024; Enehikhare & Odumuyiwa, 2025). The unsupervised architecture also holds theoretical significance, validating anomaly detection theory by demonstrating that evolving fraud patterns are best modeled as dynamic statistical outliers rather than static binary classes (Bolton & Hand, 2002; Chandola et al., 2009).

A key limitation of this study is its reliance on a single anonymized European transaction dataset, which lacks region-specific contextual metadata such as geolocation patterns, merchant behaviour, and transaction characteristics commonly associated with African financial systems. Consequently, the findings may not fully generalize across diverse banking infrastructures or emerging fintech ecosystems. Furthermore, the study primarily focused on unsupervised anomaly detection without extensive comparative evaluation against multiple supervised learning algorithms. Future research should therefore incorporate multi-institutional datasets from different geographical regions and evaluate hybrid machine learning architectures that improve both detection sensitivity and false-positive reduction.

CONCLUSION

This study demonstrates the practical applicability of unsupervised machine learning models for automating the detection of fraudulent financial activities, a critical step toward enhancing the resilience of modern fintech ecosystems. By implementing the Isolation Forest algorithm and validating it on high-dimensional transactional data, this paper demonstrates that the model can accurately isolate rare anomalies from millions of legitimate data points without prior labelling.

While traditional rule-based systems and supervised classifiers often struggle with the extreme class imbalance inherent in financial data, the Isolation Forest model achieved a commendable recall of 94.7% and an ROC-AUC of 0.91. This validates the potential of using specialized, lightweight anomaly detection models for efficient real-time classification, especially in dynamic environments where new fraud patterns emerge rapidly.

Future studies may further enhance fraud detection performance by integrating hybrid and ensemble learning architectures. Combining Isolation Forest with advanced supervised models such as XGBoost or deep learning autoencoders may improve classification balance by simultaneously increasing recall and reducing false positives. Additionally, explainable artificial intelligence (XAI) techniques such as SHAP analysis could improve model interpretability by identifying the transaction features most responsible for fraud predictions.

The comparative evaluation further demonstrated that while autoencoder models achieved slightly higher overall discrimination performance, Isolation Forest offered a more computationally efficient, lightweight deployment solution suitable for resource-constrained financial environments.

Overall, the research confirms that machine learning techniques, particularly unsupervised approaches, hold substantial promise for improving the efficiency and scalability of financial security. The system's integration into a Flask-based web interface further supports its real-world deployment, providing financial institutions with a practical tool for proactive monitoring and promoting secure digital economic practices.

REFERENCES

1. Ahmed, M., Mahmood, A. N., & Hu, J. (2016). Survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19–31.
2. Arora, P., & Bhardwaj, A. (2025). A review of anomaly detection techniques in financial transactions: The move toward AI-driven solutions. International Journal of Financial Studies, 13(1), 45–58.
3. Bahnsen, A. C., Aouada, D., Ottersten, B., Stojanovic, J., & Žliobaitė, I. (2016). Cost-sensitive credit card fraud detection using Bayes minimum risk. Proceedings of the European Symposium on Artificial Neural Networks (ESANN), 1–6.
4. Bello, A., Omoniyi, T., & Hassan, M. (2023). The evolution of digital financial crimes in emerging economies: A review of cybersecurity threats in Nigeria. Journal of African Financial Technology, 14(2), 45–62.
5. 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.
6. Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. Statistical Science, 17(3), 235–255.
7. Carcillo, F., Dal Pozzolo, A., Le Borgne, Y. A., Caelen, O., Mazzer, Y., & Bontempi, G. (2019). Combining unsupervised and supervised learning in credit card fraud detection. Information Sciences, 557, 317–331.
8. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), 1–58.
9. Dal Pozzolo, A., Caelen, O., Le Borgne, Y. A., Waterschoot, S., & Bontempi, G. (2015). Calibrating probability with undersampling for unbalanced classification. IEEE Symposium on Computational Intelligence and Data Mining (CIDM), 159–166.
10. Enehikhare, O., & Odumuyiwa, V. (2025). Addressing class imbalance in fraud detection datasets using synthetic oversampling and anomaly detection. Nigerian Journal of Computing and Applied Sciences, 7(4), 88–103.
11. Fatokun, J., Adeleke, R., & Ibrahim, S. (2025). Regulatory interventions and the persistence of electronic fraud in the Nigerian banking sector. Journal of Financial Regulation and Compliance (Africa Edition), 11(3), 201–218.
12. Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. (2019). Using generative autoencoders for detecting credit card fraud. Expert Systems with Applications, 132, 1–15.
13. Gandhar, A., Gupta, R., Sharma, P., & Verma, S. (2024). Machine learning and deep learning for financial security: The role of feature engineering and ensemble methods. Global Journal of Computer Science and Technology, 24(1), 12–29.
14. Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008). Isolation Forest. Eighth IEEE International Conference on Data Mining, 413–422.
15. Nigeria Inter-Bank Settlement System (NIBSS). (2024). Electronic payment fact sheet and fraud landscape Q1-Q2 2024. NIBSS PublicaDufisan

ISSN No. 2454-6194 | DOI : 10.51584/IJRIAS |Volume XI Issue VI June 2026

16. Odufisan, K., Abhulimen, E., & Ogunti, E. (2025). Intelligence-driven defense: Enhancing fraud detection using context-aware machine learning models. West African Journal of Information Technology, 18(1), 15–34.
17. Onyeama, C. (2024). Unsupervised learning and autoencoders for financial anomaly detection: A Nigerian case study (Unpublished master's thesis)—University of Lagos, Nigeria.
18. Roy Devarakonda, S. (2023). Hybrid ML framework integrating isolation forests and autoencoders for anomaly detection. Journal of Big Data Analytics in Finance, 5(2), 77–94.
19. Sahin, Y., Bulkan, S., & Duman, E. (2013). A cost-sensitive decision tree approach for fraud detection. Expert Systems with Applications, 40(15), 5916–5923.
20. Sawant, A. (2025). Comparing supervised and unsupervised learning for detecting novel fraud patterns. TechScience Academic Review, 12(3), 210–225.
21. Zhang, Y., & Swarup, S. (2019). Feature engineering and regularized preprocessing dependencies in structural anomaly classifiers. IEEE Transactions on Knowledge and Data Engineering, 31(8), 1420–1434.