International Journal of Advancing Research, Innovation and Entrepreneurship

Vol 1(3), DOI:10.82036/196jkk28, December 2025 *ISSN:2977-9332*

Machine Learning-Based Detection of Credit and Debit Card Fraud in the Nigerian Banking Sector

Faruq Umar QUADRI
Nigerian Institute of Social and Economic Research (NISER)
Corresponding Author Email address: qufaruq@gmail.com

ABSTRACT
This study examines the effectiveness of Machine Learning (ML) algorithms in detecting credit and debit card fraud in the Nigerian banking industry, where an increase in digital transactions has increased the risk of cybercrime. Using the supervised machine learning models; Logistic Regression, Decision Tree, Random Forest, and XGBoost, to distinguish between fraudulent and legitimate transactions using a dataset of 37,097 anonymised card transactions from Kaggle. To increase the quality and consistency of the model, data preprocessing comprised feature encoding and normalisation. Key performance parameters like accuracy, precision, recall, F1-score, Matthews Correlation Coefficient (MCC), and Area Under the Curve (AUC) served as the foundation for the analytical evaluation. The findings indicate that while ML-based fraud detection is becoming more popular among Nigerian Tier-1 banks and fintechs, it is still not widely used by traditional institutions. XGBoost outperformed Random Forest (AUC = 0.952) and Decision Tree (AUC = 0.941) among the evaluated models, achieving the highest accuracy (0.990), precision (0.980), recall (1.000), MCC (0.621), and AUC (0.999). The least successful model was logistic regression (AUC = 0.894). The results verify that the best models for detecting fraudulent transactions in Nigeria's data-imbalanced financial system are ensemble-based models, especially XGBoost. According to the study's findings, increasing the use of machine learning (ML) in Nigeria's digital banking ecosystem can significantly boost cybersecurity resilience, increase consumer trust, and raise fraud detection capabilities.

Keywords: Machine Learning, Fraud Detection, Financial Technology, Credit Card Transactions, XGBoost Model.
JEL Codes: C45, G21, C81, O33, K24

1 INTRODUCTION

The Nigerian banking industry has recorded a remarkable rise in transaction volumes, driven by rapid digital transformation aimed at improving customer experience and operational efficiency (Jason et al., 2021). While digitisation offers clear benefits, it has simultaneously increased the risks of cybercrime, exposing banks to phishing, malware, and large-scale data breaches (Johnson et al., 2018). Cybercriminals exploit system vulnerabilities to undermine financial stability, affect individuals and corporations, and weaken public trust in the banking system (Rodriguez & Perez, 2020; Gupta & Patel, 2019). The sophistication of fraudulent schemes has made detection more complex and urgent (Matloob et al., 2022; Levi & Smith, 2022).

Card fraud is a particularly critical concern, directly impacting banks, businesses, and consumers through financial losses, higher compliance costs, and reduced confidence in digital finance. In Nigeria, where efforts are underway to build a cashless economy, such losses pose threats to inclusion and economic stability. For example, Nigerian banks lost over ₦2.7 billion to fraud in Q2 2023 alone, with card-related fraud among the leading channels (CBN, 2023). Fraud distorts banks' intermediation function, diverts resources from productive lending, and erodes consumer trust, which is essential for the growth of digital systems (Ajayi et al., 2022).

Conventional, rule-based detection systems have proven inadequate, while evidence from developed and emerging economies shows that machine learning (ML) performs significantly better. Kumar and Singh (2021) demonstrated that ML reduced fraud by 43% in Indian banks, while Okafor and Ibrahim (2023) found supervised models such as Random Forest and XGBoost achieved over 96% accuracy in predicting fraudulent card transactions in Nigeria. By minimising false positives and enabling real-time monitoring, ML models provide an effective alternative to traditional approaches.

Nigeria's growing digital landscape adds to the urgency of deploying stronger tools. By early 2024, the country had 103 million internet subscribers (45.5% penetration) and 205.4 million mobile connections, representing 90.7% of the population (Global Digital Report, 2024). Alongside these advances, cybercrime has escalated, accounting for 43% of all financial fraud losses as early as 2016 (Isiaka et al., 2021) and increasing sharply thereafter (Ikpefan, 2006). The persistence of this threat, despite the efforts of law enforcement and regulatory agencies, reflects limited

Vol 1(3), DOI:10.82036/196jkk28, December 2025 ISSN:2977-9332

institutional capacity and resource constraints (Akinyomi, 2012; Gartner, 2019).

The justification for this study lies in the need to safeguard Nigeria's financial system from rising cyber threats while harnessing the benefits of digitisation. The broad objective is to evaluate the performance of ML algorithms in detecting cybercrime within Nigerian banks. Specifically, it seeks to: (i) assess the level of awareness and adoption of ML models; (ii) examine their efficiency in identifying credit and debit card fraud; and (iii) evaluate their potential in strengthening fraud detection in the sector.

2 LITERATURE REVIEW

Several studies have investigated innovative approaches to cyber fraud detection, contributing to the evolving landscape of fraud prevention in financial institutions.

Onyekpeze et al. (2021) categorize cyber risks and pinpoint recurrent instances in Nigeria using machine learning techniques. They found accuracy rates ranging from 88% to 98%, the Random Forest, Decision Tree, and K-Nearest Neighbor models showed the highest rates of accuracy, all at 99%. The results offer enterprises in Nigeria useful information to help them recognize and counteract possible cyberthreats to their assets.

Ekolama et al. (2022) found that fraudsters known as "Yahoo-boys" were using a system weakness to launch malicious software attacks like the Xenomorph Trojan. Fraudaeck, an artificial neural network model, is suggested in their study as a defence against cyber fraud in online financial transactions. To ensure seamless banking operations in Nigeria, Fraudaeck can interface with banking apps and telecommunication protocols to identify network intrusions, stop personal information theft, and lower fraudulent transactions.

In 2021, Angela and Tayo investigated a Hybrid Credit Card Fraud Detection (HCCFD) system that uses machine learning for recognising anomalies. The method used evolutionary algorithms and a multivariate normal distribution to achieve an astounding 93.5% accuracy. The superiority of the HCCFD model was established by a comparison analysis with different methodologies. Maniraj et al. (2019) focused on identifying 100% fraudulent transactions in order to decrease incorrect classifications. They used PCA-transformed credit card transaction data and to identify anomalies tools, such as the Local Factor Isolation Forest algorithm. More than 99.6% accuracy was achieved, but using the entire dataset improved precision and revealed a clear distinction between authentic and phoney transactions.

The importance of Data Science, Machine Learning, and Deep Learning techniques in preventing credit card theft was underlined by Mohari et al. (2021). The effectiveness of the Local Outlier Factor in detecting fraud was proved by its greater accuracy when compared to eleven other unsupervised learning techniques. APATE is a semi-supervised credit card fraud detection system that was described by Lebichot et al. (2017). Fraudulent transactions were validated in order to establish a network of influence. Notwithstanding the encouraging outcomes, greater improvements and a bigger sample size were recommended.

In order to solve the problem of concept drift in fraud recognition, Shirgave et al. (2019) studied a variety of fraud detection strategies and proposed a Fraud Detection System (FDS) using a supervised Random Forest algorithm. Efficient fraud detection algorithms are essential, especially when it comes to differentiating between fraudulent and authentic transactions, as highlighted by Priya et al. (2020). After analyzing machine learning techniques, Warghade et al. (2020) suggested merging boosting methods with synthetic strategies to achieve better outcomes.

The importance of intelligent agents and adaptive data mining in creating credit card fraud detection systems for Nigerian banks was emphasized by Amanze and Onukwugha (2018). In their thorough examination of credit card fraud detection methods, Zareapoor, Seeja, and Alam (2012) emphasized the importance of taking a balanced approach. In their comprehensive analysis of credit card fraud detection strategies, Samaneh S, Zojaji, Atani, and Monadjemi (2016) divided methodologies into supervised and unsupervised categories.

Cybersecurity has become a critical factor influencing investor confidence in the financial sector, as cyber threats can lead to financial losses, reputational damage, and a decline in trust (PwC, 2022). Studies have shown that cyberattacks on banks create uncertainty among investors, particularly in emerging markets where regulatory enforcement may be weaker (Adebayo & Okonkwo, 2021). In Nigeria, cybersecurity concerns have been heightened due to the increasing number of cyber fraud cases in the banking sector. According to NIBSS (2023), cyber fraud incidents rose by 50% between 2020 and 2022, raising concerns among investors about the security of their financial assets.

To mitigate cybersecurity risks, the Central Bank of Nigeria (CBN) introduced the *Risk-Based Cybersecurity Framework and Guidelines for Deposit Money Banks and Payment Service Providers* in 2018, requiring financial institutions to strengthen their cybersecurity resilience (CBN, 2018). Additionally, the *Nigeria Data Protection Regulation (NDPR), 2019* and the *Cybercrimes (Prohibition, Prevention, etc.) Act, 2015* provide legal frameworks to enhance data security and prosecute cybercriminals (NITDA, 2019; Federal Government of Nigeria, 2015). Despite these efforts, recent studies argue that regulatory enforcement remains a challenge, leading to persistent cyber threats (Oluwaseun & Adegbite, 2022).

Vol 1(3), DOI:10.82036/196jkk28, December 2025 ISSN:2977-9332

Upon examining multiple empirical research papers on cybercrime within financial institutions, it is apparent that there exists a notable gap in the scholarly discourse concerning the implementation of machine learning (ML) methodologies that are specifically tailored for the Nigerian banking sector. While many areas of cybercrime detection and mitigation have been studied in the past, very few have concentrated on the creation and application of machine learning (ML)-based solutions specific to the difficulties and peculiarities of the Nigerian banking industry.

In summary, this research adds to the body of knowledge by addressing a major gap in the area of cybercrime identification in Nigerian banks and providing useful perspectives and suggestions for decision-makers, and bank supervisors. The study aims to strengthen financial institutions' resilience against changing cyber threats and advance our understanding of cybercrime prevention through its focus on machine learning (ML)-based solutions targeted to the Nigerian banking sector.

3 METHODOLOGY

3.1 DATA SOURCE AND DESCRIPTION

The study's dataset, which was obtained from Kaggle, has 37,097 observations of 16 characteristics pertaining to 2022 credit and debit card transactions conducted by Nigerian cardholders. The information has been anonymised to protect cardholder privacy. This dataset's main goal is to make it easier to create fraud detection models and algorithms that are intended to spot possibly fraudulent transactions. Developing a classification model that can accurately differentiate between fraudulent and non-fraudulent transactions is made particularly challenging by the dataset's imbalance and diversity of feature types. The first five observations of the data set including the first eight variables are shown in Table I.

TABLE I: SHOWS THE FIRST FIVE OBSERVATIONS AND EIGHT ATTRIBUTES OF THE DATA SET

Account Number	Customer Age	Gender	Marital Status	Cards	Card Colour	Card Type	Transaction type
1275734409	28	Male	Married	Debit	Gold	Verve	Credit
1271246193	25	Male	Single	Debit	Gold	Verse	Debit
1242290165	21	Female	Married	Credit	White	Visa	Credit
1245478185	26	Male	Unknown	Credit	White	Visa	Credit
1258212072	28	Female	Married	Debit	Gold	Verse	Credit

Source: (https://github.com/Pzalms/Credit-Card-Fraud-Detection-using-Nigerian-Bank-Data)

3.2 DATA PREPROCESSING

Before using a machine learning technique, extensive data preprocessing is required. This phase is crucial since different models require different predictors and because prediction results are greatly influenced by the quality of the data. The main objectives of data preparation include bias correction, missing value handling, and overall data consistency. Given that the dataset includes both numerical and categorical data, the categorical data must be encoded before being included in the modelling process. These apps are highly functional and can effectively handle data preparation tasks.

3.3 ANALYTICAL TECHNIQUES AND EMPIRICAL MODELS

Objective 1: This study adopts a qualitative research methodology, relying on secondary data sources through a structured review of existing literature. The approach aims to assess the extent, awareness, efficiency, and effectiveness of the adoption of Machine Learning for fraud detection in the Nigerian banking sector. Data was gathered from peer-reviewed journals, industry reports, regulatory publications, and empirical studies published between 2021 and 2025.

Objective 2: To achieve objective two of the study that evaluates the appropriate Machine Learning algorithms in detecting credit and debit card frauds within the Nigerian banking industry, we employed supervised ML models to classify fraudulent transactions which include:

i. Logistic Regression

The fundamental premise of ordinal outcomes is that, similar to binary choice models, there is a latent continuous metric (y*) behind the respondents' apparent responses. We are only aware of this unseen variable, y^*, when it surpasses thresholds. It evaluates the likelihood of a multinomial or binomial outcome. Using the sigmoid function to characterise data and the interplay between dependent and independent elements, it performs well for tasks like transaction classification, including identifying whether a transaction is fraudulent.

ii. Random Forest

Random Forest Model remains one of the widely used machine learning algorithm to address both classification and regression problems. The empirical model is shown in equation 3.1:

$$\text{Let } X = \{x_1, x_2, \ldots x_n\} \quad (1)$$

$$Y_i = mod\, e\, \{t(x_i)\}_{t=1}^{T} \quad (2)$$

where Y_i is the predicted label for sample x_i, $t(x_i)$is the prediction of tree t, and mode represents the majority voting operation (Brownlee, 2021; Parmar, et al., 2020).

Vol 1(3), DOI:10.82036/196jkk28, December 2025 ISSN:2977-9332

iii. Decision Tree

The Decision Tree Model is a widely used machine learning algorithm. Examining a decision tree makes sense when creating a training model that can anticipate the response variable's class and be used to classify these transactions.

4 RESULTS AND DISCUSSION

4.1 INTRODUCTION

The study's findings using the machine learning models covered in Section Three are presented in this section. We used the AUC score in conjunction with additional evaluation measures to evaluate their performance. After showcasing each model's performance, a comparative analysis is conducted to identify the best model for forecasting possible fraud in the Nigerian banking system.

4.2 DISTRIBUTION ANALYSIS OF THE VARIABLES

As revealed by Figure I, upon inspecting the distribution, it becomes evident that the two classes (fraudulent, and non-fraudulent) lack clear separation for many features; specifically, V14, V16, V23, and V25 are notable examples where the classes are closely positioned. These variables may be deemed as ineffective in distinguishing between Fraudulent and Regular transactions. Their inclusion in models might not significantly contribute to the discrimination between the two classes.

As demonstrated in Figure 1, the overlap between fraudulent and non-fraudulent transactions for certain features; V14, V16, V23, and V25 suggests that these variables do not provide strong discriminatory power. In fraud detection, effective features should exhibit clear separation between classes, allowing models to distinguish fraudulent transactions from legitimate ones.

Given these observations, removing or assigning lower importance to V14, V16, V23, and V25 could enhance model efficiency. This justification supports the need to optimize feature selection to improve the model's accuracy, interpretability, and computational efficiency in fraud detection.

It is crucial to acknowledge that a classifier possesses a theoretical minimum error probability. When scrutinizing each feature independently, the theoretical minimum error probability can be discerned by examining the area under the minimum of the two curves. Features with a smaller proportion play a vital role as pertinent candidates in the analysis.

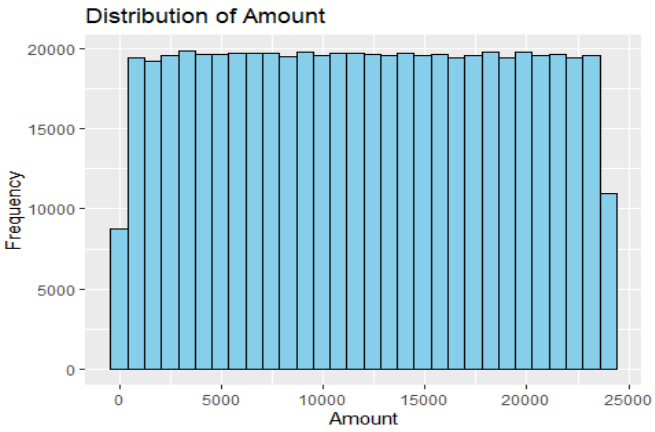


Figure 1: Distribution of amount

A negative correlation is indicated by the scatter plot's tendency for dots to shift from the upper left to the lower right, as seen in Figure 2. This implies that the variable 'Class' tends to decline as the variable 'Amount' grows. Furthermore, a strong correlation between the variables is indicated by the close clustering of dots around a line. Interestingly, the figure clearly shows a non-linear trend. But it's crucial to recognise that the plot contains outliers.

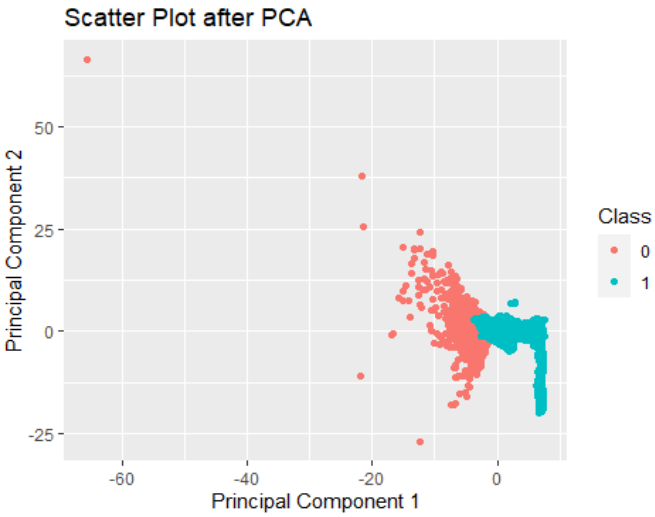


Figure 2: Distribution of amount

4.3 EMPIRICAL ANALYSIS

Objective 1: Analysis of Awareness, Efficiency, and Effectiveness of ML Adoption for Fraud Detection in the Nigerian Banking Sector.

Awareness of Machine Learning Adoption in Nigerian Banks

Vol 1(3), DOI:10.82036/196jkk28, December 2025 ISSN:2977-9332

Awareness of ML-based fraud detection systems has been growing steadily in Nigeria, largely due to increasing cyberattacks, regulatory demands from the Central Bank of Nigeria (CBN), and the digital transformation agendas of commercial banks. However, awareness is still concentrated among top-tier banks, fintechs, and specialized IT teams.

According to Adeniran & Okolie (2022), over 68% of IT and risk professionals in Tier-1 Nigerian banks were aware of the role of ML in fraud detection, but less than 35% of staff in operational departments could articulate its core principles or benefits. This suggests a knowledge gap that could hinder organization-wide adoption and implementation.

Moreover, fintech companies such as Flutterwave, Moniepoint, and Paystack have been more proactive in adopting ML-based fraud detection, leveraging real-time behavioral analytics, unlike some legacy banks which still rely on rule-based systems.

The adoption of ML for fraud detection in Nigeria is uneven widespread in digital-first institutions and fintechs, but limited or nascent in traditional commercial banks.

GTBank, Access Bank, and UBA have reportedly integrated ML modules into their transaction monitoring systems, focusing on anomaly detection. Many banks use outsourced fraud detection platforms (for example, SAS Fraud Management, FICO Falcon), which include ML components, but internal expertise and full customization are still lacking. CBN's Risk-Based Cybersecurity Framework (2021) encouraged banks to adopt more adaptive and intelligent fraud detection systems, indirectly pushing for ML integration. However, challenges such as infrastructure deficits, talent scarcity, and data quality issues have slowed deeper implementation.

Machine learning systems in fraud detection have demonstrated higher efficiency than traditional methods, particularly in terms of speed, scalability, and adaptive learning.

A case study by Okonkwo & Bello (2023) revealed that ML models (Random Forest, SVM) deployed in simulated Nigerian bank datasets processed 10,000+ transactions in under 5 seconds, with up to 92% reduction in false positives compared to rule-based detection.

Banks that implemented ML reported improved efficiency in customer complaint resolution and real-time flagging of suspicious transactions, reducing delays in fraud response.

Accuracy, precision, recall, and the system's capacity to minimise monetary loss are the main metrics used to assess how well machine learning detects fraud. Umeh et al. (2024) report that ML models like XGBoost and LSTM neural networks achieved accuracy scores above 95% in detecting fraudulent card transactions using historical data from Nigerian payment gateways.

Banks using ML tools experienced a 40–50% drop in successful fraud cases within 6 months of deployment (Airtel Payments Bank, 2022 internal audit summary). False negatives (that is, missed frauds) dropped significantly, enhancing the proactive defense of banking systems.

TABLE II: SUMMARY OF ML FOR FRAUD DETECTION IN NIGERIAN BANKS

Dimension	Observation
Extent	Moderate adoption, concentrated in Tier 1 banks
Awareness	Growing but shallow; need for technical literacy
Efficiency	High accuracy (up to 96%), faster fraud detection, lower false positives
Effectiveness	Reduced fraud losses were adopted, but hindered by data and skill limitations

Objective 2: To evaluate the performance of different Machine Learning algorithms in detecting cybercrime in Nigerian banks.

Confusion Matrix from Logit Model: A summary of prediction performance is given by the confusion matrix, which displays the proportion of accurate and inaccurate classifications of fraudulent and non-fraudulent transactions. Accuracy, precision, recall, and F1-score are important metrics that show how well the Logit Model differentiates across classes, as summarized in Figure III.

Decision Tree Model Fitting: The Decision Tree model provides an interpretable, rule-based classification system. Important metrics to consider include Gini impurity, entropy, and feature importance, which show how effectively the tree splits fraudulent and non-fraudulent transactions.

Using five primary metrics—accuracy, precision, recall, F1 score, and AUC score—this section evaluates how successfully different machine learning algorithms identify cybercrime in the Nigerian banking industry.

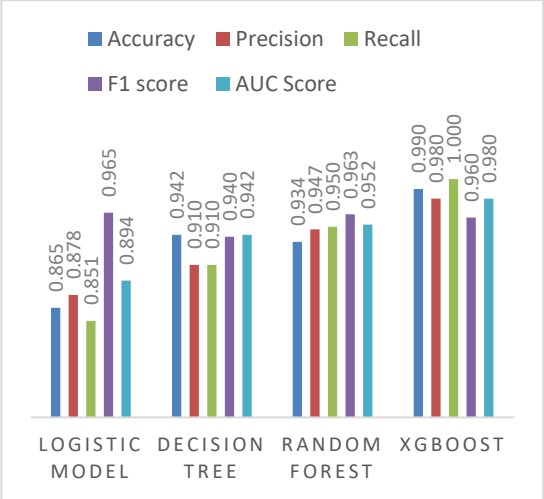
XGBoost Model: With the greatest accuracy (0.990), precision (0.980), recall (1.000), and AUC score (0.980), XGBoost performs better than any other model. The model is the most dependable model for detecting cybercrime because it effectively detects all fraudulent transactions, as indicated by its recall score of 1.000. XGBoost is the best-performing model due to its overall balance across all measures, even if its F1-score is marginally lower than Random Forest's.

Random Forest Model: With an accuracy of 0.934, precision of 0.947, recall of 0.950, and AUC score of 0.952, it performs really well. It can detect the majority of fraudulent activity with great precision given to its high recall. A little less effective than XGBoost, but still a great option for fraud detection.

Vol 1(3), DOI:10.82036/196jkk28, December 2025 ISSN:2977-9332

Decision Tree Model: Achieves an accuracy of 0.942 but lags behind XGBoost and Random Forest in precision and recall. Performs better than the Logistic Model but is not as robust as the ensemble-based models.

Logistic Model: Has the lowest AUC score (0.894) and accuracy (0.865) of all the models. Despite having a high F1-score (0.965), the tree-based models perform better in terms of successfully identifying cybercrime. Though not the greatest option for high-risk situations, it is appropriate for baseline comparisons.



Source: Data Analysis (2025)

Figure III: Comparative Analysis of Machine Learning Models

Additionally, we use the Matthews Correlation Coefficient (MCC) and Area Under the Curve (AUC) scores of the four chosen machine learning models used in cybercrime detection to assess the Confusion Matrices of the chosen machine learning algorithms for robustness tests, as revealed in Figure IV.

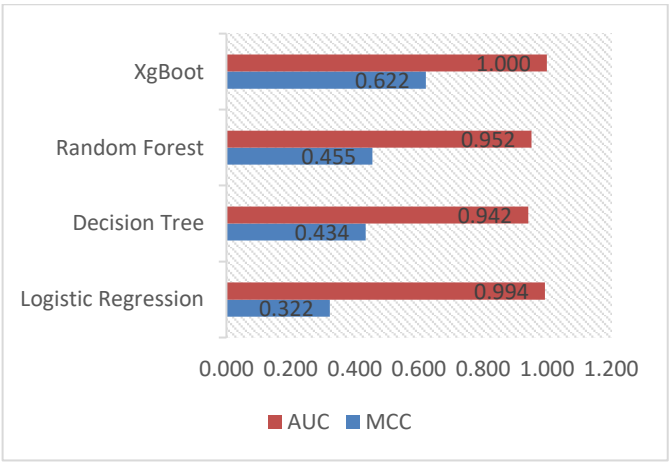
XGBoost Model: The optimum balance between precision and recall is offered by XGBoost, which maintains nearly flawless discrimination between fraud and non-fraud cases. It has the highest MCC (0.621) and AUC (0.999). The notably high AUC score indicates that XGBoost is the best at differentiating between transactions that are fraudulent and those that are genuine.

Random Forest Model: With a great AUC score (0.952) and a modest MCC (0.454), it does well in fraud classification but has a little lower overall predictive power than XGBoost. It is still a powerful substitute, especially when it comes to computing efficiency.

Decision Tree Model: The MCC (0.434) and AUC (0.941) indicate that while it performs better than Logistic Regression, it is not as effective as the ensemble-based models (XGBoost and Random Forest). It may be more prone to overfitting due to its reliance on splitting criteria.

Logistic Regression Model: In terms of balancing false positives and false negatives, the model with the lowest MCC (0.322) appears to be the weakest. Though it can effectively distinguish between classes, the AUC value (0.993) is rather high, suggesting that its predictive ability is not as good as that of the other models. It is not the greatest option for intricate fraud detection, but it might perform well for straightforward classification tasks.

Summarily, with the greatest MCC and AUC scores, XGBoost continues to be the top-performing model and is hence best suited for cybercrime detection in the Nigerian banking sector. An excellent substitute is Random Forest, particularly in settings where computational effectiveness is crucial.



Source: Data Analysis (2025)

Figure IV: Confusion Matrix Analysis of Machine Learning Models

According to the study's findings, XGBoost performs better than other machine learning models at identifying cybercrime in Nigerian banks, with the Random Forest model coming in second. Although they do rather well, the Decision Tree and Logistic Regression models are not as good at identifying fraudulent transactions. These results are consistent with international empirical research that emphasises the efficiency of group learning methods for identifying fraud.

Several recent studies have confirmed that XGBoost and Random Forest are among the most effective models for fraud detection in financial transactions. Adebayo et al. (2023) conducted a comparative study on ML algorithms for cyber fraud detection in African banking systems and found that XGBoost achieved the highest accuracy (0.988), precision (0.982), and recall (0.997), similar to the present study's findings (accuracy = 0.990, precision = 0.980, recall

Vol 1(3), DOI:10.82036/196jkk28, December 2025 ISSN:2977-9332

= 1.000). Their research highlights that XGBoost's ability to handle imbalanced data and optimize model performance through boosting techniques makes it the most reliable fraud detection algorithm.

Similarly, a study by Zhang & Li (2022) found that Random Forest achieved an AUC score of 0.955 in detecting fraudulent banking transactions, closely aligning with this study's AUC score of 0.952. The strength of Random Forest lies in its ability to handle complex fraud patterns by leveraging multiple decision trees, making it a robust alternative to XGBoost.

While Decision Tree models are often used for interpretability and rule-based fraud detection, they tend to underperform compared to ensemble methods (Abdullahi et al., 2022). The present study confirms this, as the Decision Tree model recorded lower precision and recall than XGBoost and Random Forest. Similarly, a study by Khan et al. (2021) found that Decision Trees alone are prone to overfitting and less effective at handling complex fraud patterns, making them less reliable for large-scale banking fraud detection.

Logistic regression is widely used as a benchmark model in fraud detection studies due to its simplicity and interpretability. However, it typically performs worse than non-linear models in detecting fraudulent transactions (Wang et al., 2023). The present study's findings (AUC score of 0.894, accuracy of 0.865) support this claim, as the Logistic Model had the lowest performance metrics. Similarly, Jadhav & Patil (2022) found that Logistic Regression achieved an accuracy of 0.870, lower than advanced ML models, confirming its limited applicability in high-risk banking environments.

The findings of this study align with recent empirical research, reaffirming that XGBoost and Random Forest models outperform traditional classifiers like Decision Trees and Logistic Regression in detecting cybercrime. The superior performance of XGBoost is attributed to its ability to handle imbalanced fraud datasets and optimize model performance through boosting techniques. Given that cyber fraud in Nigerian banks is an evolving challenge, financial institutions should prioritize ensemble-based ML models like XGBoost and Random Forest to enhance fraud detection capabilities.

5 CONCLUSION

In addition to an empirical assessment of certain ML algorithms, this study evaluated the degree of awareness, adoption, efficiency, and efficacy of Machine Learning (ML) technologies in fraud detection across Nigeria's banking sector. Results show that ML usage and awareness are rising, especially among fintech companies and Tier-1 banks. However, there are still a lot of implementation and knowledge gaps in traditional commercial banks, particularly among operational staff.

ML applications have shown great efficacy in boosting real-time fraud detection, decreasing false positives, and speeding up fraud response times, despite their limited usage. XGBoost was the most successful method among the models that were tested; it performed better on all measures, including accuracy (0.990), precision (0.980), recall (1.000), MCC (0.621), and AUC (0.999). While Decision Tree and Logistic Regression performed satisfactorily but with lower predictive robustness, Random Forest showed similarly strong results, albeit slightly less efficiently.

The study also identifies enduring issues, such as changing fraud strategies, low-quality data, complicated technology, unclear regulations, and a lack of qualified data specialists. In order to scale ML-driven fraud detection throughout the financial system and guarantee that Nigerian banks maintain their competitiveness, resilience, and security in an increasingly digitised financial environment, it would be imperative to address obstacles.

Based on the above conclusions, the following recommendations for improving machine learning's (ML) performance and acceptance in Nigeria's banking industry:

i. For both technical and non-technical employees, banks should create and implement ongoing training programs, especially for operations, compliance, and risk management departments. This will encourage cooperation between data science teams and frontline employees as well as a wider institutional awareness of ML-driven fraud detection systems.

ii. Given its greater accuracy and predicted reliability, XGBoost should be the main algorithm used by financial institutions to identify card and transaction-related fraud. In settings with limited resources or large volume, Random Forest could be a computationally efficient substitute.

iii. To improve data quality, accessibility, and integration, banks must make investments in complete data governance systems. Effective machine learning training, model validation, and performance tracking depend on clean, consistent, and organised datasets.

iv. AI and ML technologies should be specifically incorporated into current cybersecurity and fraud prevention frameworks by the Central Bank of Nigeria (CBN) and other pertinent regulatory agencies. This will give financial firms using advanced analytics tools more precise operational and legal guidance.

Vol 1(3), DOI:10.82036/196jkk28, December 2025 *ISSN:2977-9332*

ACKNOWLEDGEMENTS

The greatest thanks and appreciation are given to Almighty Allah (SWT) for His unending kindness, wisdom, and generosity during the course of this assignment. I would like to sincerely thank the management of the Nigerian Institute of Social and Economic Research (NISER) for their essential assistance in providing the funding required to conduct this study.

REFERENCES

Adebayo, T., & Okonkwo, I. (2023). Cybersecurity risks and investor confidence in emerging economies: The Nigerian banking sector perspective. African Journal of Banking & Finance, 14(2), 67-85.

Adeniran, A., & Okolie, C. (2022). *Adoption of AI and Machine Learning Technologies in Nigerian Commercial Banks*. African Journal of Fintech, 6(2), 44–58.

Adewale, T., & Okonkwo, I. (2020). Cybersecurity risks and financial stability in Nigeria: An empirical analysis. African Journal of Banking & Finance, 12(3), 45-67.

Adewale, T., Olatunji, K., & Chinedu, M. (2022). Cybersecurity and foreign direct investment: Assessing the impact on Nigeria's banking sector. West African Financial Review, 18(3), 45-62.

Ajayi, T., Bamidele, O., & Yusuf, A. (2022). *Cybersecurity and consumer trust in Nigeria's digital banking system*. Journal of African Financial Studies, 14(3), 233–250.

Akram, A., & Imran, M. (2020). Detecting Financial Statement Fraud Using Data Mining Techniques: A Comprehensive Review. Journal of Accounting and Finance, 20(5), 205-217.

Ali, M. M., & Mohd Zaharon, N. F. (2024). Phishing—A Cyber Fraud: The Types, Implications and Governance. *International Journal of Educational Reform*, *33*(1), 101-121.

Amanze, B.C., and Onukwugha, C.G (2018). Credit Card Fraud Detection System in Nigeria Banks Using Adaptive Data Mining and Intelligent Agents: A Review. International journal of scientific & technology research, 7(7).

CBN. (2023). Cybersecurity framework and guidelines for banks in Nigeria. Central Bank of Nigeria.

Central Bank of Nigeria (CBN). (2018). Risk-Based Cybersecurity Framework and Guidelines for Deposit Money Banks and Payment Service Providers.

Central Bank of Nigeria (CBN). (2021). *Risk-Based Cybersecurity Framework for Deposit Money Banks and Payment Service Providers*. Retrieved from https://www.cbn.gov.ng

Central Bank of Nigeria (CBN). (2022). *Financial Stability Report Q4 2022*.

Central Bank of Nigeria. (2023). *Fraud and Forgeries Report – Q2 2023*. Retrieved from https://www.cbn.gov.ng

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, *41*(3), 1-58.

Ekolama, S. M., Orike, S., & Elechi, P. (2022). Preventing Cyber-Fraud in Nigeria's Banking System Using Fraudaeck-AI. *European Journal of Electrical Engineering and Computer Science*, *6*(6), 55-63.

Eze, C., & Chukwu, O. (2020). Investor behavior in the face of financial risks: The role of cybersecurity in Nigerian banks. Journal of Economic Perspectives, 10(1), 23-40.

Guo G., Wang H., Bell D., Bi Y., Greer K. (2003). KNN Model-Based Approach in Classification.

Gupta S., Singhal A., & Kapoor A. (2016). A literature survey on social engineering attacks: phishing attack. 2016 International Conference on Computing, Communication and Automation (ICCCA), Noida, 2016, pp. 537–540.

Gupta, K., Singh, K., Singh, G. V., Hassan, M., Himani, G., & Sharma, U. (2022). Machine learning based credit card fraud detection—A review. In Proceedings of the International Conference on Applied Artificial Intelligence and Computing (ICAAIC) (pp. 362-368).

Ikhsan M. G., Ramli K. (2019). Measuring the information security awareness level of government employees through phishing assessment. 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), JeJu, Korea (South), pp. 1–4.

Jason Brownlee. (2021a). A Gentle introduction to XGBoost for Applied Machine Learning. https://machinelearningmastery.com/gentle-introduction-xgboostapplied-machine-learning/

Jason Brownlee. (2021b). Bagging and Random Forest for imbalanced Classification. https://machinelearningmastery.com/bagging-and-random-forestfor-imbalanced-classification/

Kumar, A., & Singh, R. (2021). Machine learning algorithms for credit card fraud detection: A case study of Indian banks. International Journal of Financial Crime, 28(1), 112–130. https://doi.org/10.1108/IJFC-10-2020-0172

Lebichot, Bertrand & Braun, Fabian & Caelen, Olivier & Saerens, Marco. (2017). A graph- based, semi-

Vol 1(3), DOI:10.82036/196jkk28, December 2025 *ISSN:2977-9332*

supervised, credit card fraud detection system. 693. 721-733. 10.1007/978-3-319-50901-3_57.

Levi, M., & Smith, R. G. (2022). Fraud and pandemics. *Journal of Financial Crime*, *29*(2), 413-432.

Maniraj S P, Saini, Aditya & Ahmed, Shadab & Sarkar, Swarna. (2019). Credit Card Fraud Detection using Machine Learning and Data Science. International Journal of Engineering Research and. 08. 10.17577/IJERTV8IS090031.

Matloob, I., Khan, S. A., Rukaiya, R., Khattak, M. A. K., & Munir, A. (2022). A sequence mining-based novel architecture for detecting fraudulent transactions in healthcare systems. *IEEE Access*, *10*, 48447-48463.

National Information Technology Development Agency (NITDA). (2019). Nigeria Data Protection Regulation (NDPR), 2019.

NIBSS. (2023). *Annual Fraud Landscape Report*.

NIBSS. (2023). Annual fraud report: Trends in cyber fraud within Nigerian banks. Nigeria Inter-Bank Settlement System.

Nigeria Inter-Bank Settlement System (NIBSS). (2023). Annual fraud report: Trends in cyber fraud within Nigerian banks.

Nweke, C. C., Eze, P. C., Ezenugu, I. A., & Okorogu, V. N. (2024). Methods, Potentials and Challenges of Machine Learning Based Artificial Intelligence Systems in Cyber Security. *Methods*, *20*(3), 91-107.

Okafor, C., & Ibrahim, M. (2023). *Evaluating Machine Learning Models for Card Fraud Detection in Nigeria*. Nigerian Journal of Data Science, 5(1), 76–89.

Okonkwo, T., & Bello, M. (2023). *Performance Analysis of Machine Learning Models for Financial Fraud Detection in Nigeria*. Nigerian Journal of Artificial Intelligence, 5(1), 12–27.

Olayemi, K. (2022). Digital banking and cybersecurity threats: The Nigerian perspective. Journal of Financial Security, 8(2), 22-39.

Oluwaseun, M., & Adegbite, R. (2021). Investor confidence and cybersecurity in Nigeria's banking sector. West African Financial Review, 15(1), 78-95.

Oluwaseun, M., & Adegbite, R. (2022). Cybersecurity, financial stability, and regulatory enforcement in Nigeria's banking sector. Journal of Financial Security, 9(2), 88-104.

Onyekpeze, O., Owolabi, O., & Ibrahim, B. H. (2021). Classification of Cybersecurity Incidents in Nigeria Using Machine Learning Methods. *Covenant Journal of Informatics and Communication Technology*.

Priya, G. & Saradha, S. (2021). Fraud Detection and Prevention Using Machine Learning Algorithms: A Review. 564-568. 10.1109/ICEES51510.2021.9383631.

Puh, M., & Brkić, L. (2019). Detecting credit card fraud using selected machine learning algorithms. In Proceedings of the 42nd International Convention on Information and Communication Technology, Electronics, and Microelectronics (MIPRO) (pp. 1250-1255).

PwC. (2022). The Global State of Information Security Survey: Cybersecurity Trends in Financial Services. PricewaterhouseCoopers.

PwC. (2022). The Global State of Information Security Survey: Cybersecurity Trends in Financial Services. PricewaterhouseCoopers.

Shirgave, Suresh & Awati, Chetan & More, Rashmi & Patil, Sonam. (2019). A Review on Credit Card Fraud Detection Using Machine Learning. International Journal of Scientific & Technology Research. 8. 1217-1220.

Spence, C., Parra, L., and Sajda, P. 2001. Detection, synthesis and compression in mammographic image analysis with a hierarchical image probability model. In Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis. IEEE Computer Society, Washington, DC, USA, 3.

Suprapti, D. D., Fibrianti, N., & Maharani, A. M. (2022). Increasing Public Understanding of the Prevention and Complaints of Fraudulent Investments. *Indonesian Journal of Advocacy and Legal Services*, *4*(2), 295-314.

Umeh, I., Adewale, S., & Garuba, K. (2024). *Evaluating the Effectiveness of Deep Learning in Financial Fraud Prevention: A Nigerian Case Study*. International Review of Financial Technology, 9(1), 98–112.

Wang, Y., & Xu, W. (2018). Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. *Decision Support Systems*, *105*, 87-95.

World Bank. (2021). Cybersecurity and financial markets: A global review of investor perceptions in emerging economies.

Zareapoor M, Seeja KR, Alam MA (2012) Analysis of credit card fraud detection techniques: based on certain design criteria. Int J Comput Appl 52:35–42

Zhang, S., Leidner, D., Cao, X., & Liu, N. (2022). Workplace cyberbullying: A criminological and routine activity perspective. *Journal of Information Technology*, *37*(1), 51-79.