Advanced Journal of Engineering and Scientific Applications (AJESA)

Research Paper Open Access

Available online at www.ajesa.com.ng Vol.3, No.1, 2026. pp.1-11

# Machine Learning Approaches to Fraud Detection and Risk Analytics in GSM-Based FinTech Systems in Nigeria

**[1]Joseph Osahon Idemudia and [2]Taiwo Adisa Oyeniran**

*[1,2]Department of Computer Science, Akanu Ibiam Federal Polytechnic, Unwana, Afikpo.*

*[2]taoyeniran@akanuibiampoly.edu.ng*

*0806-983-0322*

*Corresponding Author*: [1]*jidemudia@akanuibiampoly.edu.ng 0803-582-0124*

## Abstract

*The increasing adoption of mobile phone numbers as primary identifiers in financial transactions has significantly transformed the operations of financial technology (FinTech) companies in Nigeria. Leveraging Global System for Mobile Communications (GSM) numbers as account identifiers has enhanced access to digital financial services while introducing challenges related to security, data management, and analytics. This study investigates the application of data mining and machine learning techniques in GSM-based FinTech systems, with emphasis on fraud detection, anomaly detection, customer segmentation, and credit risk assessment. A synthetic yet statistically representative GSM-linked transaction dataset was developed to reflect real-world FinTech operations under regulatory and ethical constraints. Controlled modeling experiments were conducted using supervised and unsupervised learning techniques. Experimental results show that the Random Forest classifier achieved an accuracy of 96.8%, precision of 94.5%, recall of 92.7%, and an F1-score of 0.935, substantially outperforming traditional rule-based approaches. Receiver Operating Characteristic analysis yielded a high area under the curve exceeding 0.94, indicating strong discriminative performance. Anomaly detection models effectively identified SIM-swap and account takeover fraud, achieving detection rates above 90% with average detection delays below two minutes. Unsupervised clustering revealed four distinct customer segments, including a high-risk group representing approximately 12% of users. In addition, predictive credit risk models achieved 92.1% accuracy and reduced simulated loan approval time by 37%. Feature importance analysis identified GSM-specific behavioral indicators—such as transaction value, SIM change events, and location instability—as key predictors of fraudulent and high-risk behavior. Overall, the results demonstrate that data mining significantly enhances security, operational efficiency, and financial inclusion in GSM-based FinTech systems.*

***KEYWORDS***: *Data mining, SIM-Swap Fraud, Anomaly Detection, financial technology (FinTech), GSM numbers, fraud detection, machine learning, cybersecurity, financial inclusion, Nigeria*

## 1. INTRODUCTION

The rapid expansion of financial technology (FinTech) has reshaped the global financial services landscape, particularly in developing economies such as Nigeria (Gabor & Brooks, 2017; Ozili, 2020). FinTech companies increasingly leverage digital platforms, mobile technologies, and data-driven solutions to deliver financial services that are faster, more accessible, and cost-effective than traditional banking systems (Donner & Tellez, 2008). In Nigeria,

where mobile phone penetration significantly exceeds bank account ownership, the adoption of Global System for Mobile Communications (GSM) numbers as account identifiers has emerged as a transformative innovation in digital finance (Aker & Mbiti, 2016; Musa et al., 2019).

Using GSM numbers as account numbers simplifies customer onboarding, reduces entry barriers, and enhances financial inclusion for unbanked and underbanked populations (Jack & Suri, 2014; Ozili, 2020). This approach aligns with Nigeria's cashless policy and supports the rapid growth of mobile money platforms, digital wallets, and agency banking services (Adeyemi et al., 2021). However, the widespread reliance on GSM numbers has also resulted in the generation of large volumes of transactional and behavioral data, presenting both opportunities and challenges for FinTech firms (Nwankwo & Eze, 2022).

When effectively harnessed, GSM-linked data can provide valuable insights for fraud detection, credit scoring, customer segmentation, and personalized service delivery (Ngai et al., 2018; Kou et al., 2021). Data mining and machine learning techniques offer powerful tools for extracting patterns, trends, and predictive knowledge from such large-scale datasets (Dal Pozzolo et al., 2015). Through supervised and unsupervised learning models, FinTech firms can detect fraudulent behavior, identify anomalous activities such as SIM-swap attacks, segment customers based on usage behavior, and assess credit risk using alternative data sources (Bahnsen et al., 2016; Li et al., 2019).

Despite these opportunities, the practical adoption of data mining and machine learning in Nigerian FinTech firms remains uneven. Challenges related to data quality, privacy concerns, cybersecurity threats, regulatory compliance, and infrastructural limitations continue to hinder the systematic deployment of advanced analytics (Kshetri, 2021; Okoye & Ezugwu, 2023). Against this background, this study investigates data mining and machine learning techniques applicable to Nigerian FinTech companies that leverage GSM numbers as account identifiers.

### 1.1 Background of the Study

Nigeria has one of the largest mobile subscriber bases in Africa, making GSM technology a critical enabler of digital financial services (Aker & Mbiti, 2016). Over the past decade, FinTech firms have increasingly adopted mobile phone numbers as unique customer identifiers, either replacing or complementing traditional bank account numbers (Nwankwo & Eze, 2022). This shift has accelerated the adoption of mobile payments, peer-to-peer transfers, bill payments, micro-lending, and other digital financial services (Musa et al., 2019; Adeyemi et al., 2021).

The integration of GSM numbers into FinTech systems has led to the continuous generation of large-scale datasets encompassing customer demographics, transaction histories, device metadata, geolocation information, and usage patterns (Ngai et al., 2018). These datasets provide a rich foundation for data-driven decision-making through data mining and machine learning techniques (Dal Pozzolo et al., 2015). In particular, machine learning models can be trained to recognize complex patterns associated with fraudulent transactions, behavioral anomalies, customer risk profiles, and creditworthiness (Bahnsen et al., 2016; Sharma & Panigrahi, 2020).

However, the Nigerian FinTech ecosystem operates within a complex regulatory, technological, and security environment. Issues such as inconsistent data quality, weak data governance structures, cybersecurity threats, identity theft, SIM-swap fraud, and compliance with data protection regulations pose significant risks (Alao et al., 2022; Okoye & Ezugwu, 2023). In addition, many FinTech firms lack the technical infrastructure and experimental frameworks required to implement and interpret machine learning models effectively (Kshetri, 2021).

### 1.2 Statement of the Problem

Despite the rapid adoption of GSM numbers as primary financial identifiers in Nigeria, the FinTech ecosystem remains highly vulnerable to identity-linked exploits that traditional security frameworks cannot mitigate. Current rule-based monitoring systems are largely reactive and fail to account for the unique behavioral patterns associated with mobile-centric finance, such as SIM-swap fraud and location-based anomalies. Furthermore, while massive volumes of GSM-linked transactional data are generated daily, there is a critical lack of empirically validated machine learning frameworks tailored to the Nigerian regulatory and infrastructural context. This analytical gap results in high false-positive rates, delayed fraud detection, and the exclusion of potential borrowers who lack formal credit histories. This study addresses these challenges by developing and evaluating a machine learning-driven framework designed to transform GSM metadata into actionable intelligence for fraud prevention and credit risk assessment.

# 2. LITERATURE REVIEW

## 2.1 Financial Technology (FinTech) and Digital Financial Services

FinTech has transformed payment systems, lending, savings, and investment services by leveraging mobile platforms and data analytics (Gabor & Brooks, 2017). In Nigeria, FinTech growth has expanded financial access among unbanked populations but introduced cybersecurity and data governance challenges (Musa et al., 2019; Kshetri, 2021).

## 2.2 GSM Numbers as Account Identifiers

GSM numbers simplify onboarding and real-time transactions but expose systems to SIM-swap and impersonation risks (Jack & Suri, 2014; Aker & Mbiti, 2016; Alao et al., 2022).

## 2.3 Data Mining Concepts and Techniques

Data mining techniques such as classification, clustering, and anomaly detection are widely applied in financial systems to improve accuracy and risk control (Dal Pozzolo et al., 2015; Ngai et al., 2018).

## 2.4 Data Mining Applications in FinTech

Machine learning-based fraud detection and risk analytics outperform traditional rule-based systems, especially in real-time environments (Bahnsen et al., 2016; Carcillo et al., 2021; Li et al., 2019).

## 2.5 Data Mining and Financial Inclusion

GSM-linked data enables alternative credit scoring and inclusion of users without formal banking histories (Ozili, 2020; Sharma & Panigrahi, 2020).

## 2.6 Cybersecurity and Privacy Challenges

GSM-based systems face SIM-swap fraud and data privacy risks, requiring analytics-driven monitoring and regulatory compliance (Anderson et al., 2019; Kshetri, 2021; Okoye & Ezugwu, 2023).

## 2.7 Empirical Review of literature

To streamline the empirical review and avoid redundancy, related studies were consolidated into thematic tables covering GSM-based FinTech adoption, data mining applications in financial systems, and cybersecurity and regulatory issues (Tables 1–3).

**Table 1: Empirical Studies on GSM-Based FinTech, Financial Inclusion, and Digital Financial Services**

| Author(s) & Year | Context | Focus Area | Data / Method | Key Findings | Limitations |
|---|---|---|---|---|---|
| Donner & Tellez (2008) | Developing economies | Mobile banking adoption | Survey data | Mobile phones improve access to financial services | Outdated analytics; no ML |
| Jack & Suri (2014) | Kenya | Mobile money systems | Transaction data | GSM-based systems enhance economic outcomes | No fraud or analytics modeling |
| Aker & Mbiti (2016) | Africa | Mobile finance & inclusion | Secondary data | Mobile phones reduce financial barriers | Security risks not examined |
| Musa et al. (2019) | Nigeria | Mobile financial services adoption | Survey analysis | Mobile money increases inclusion | Limited analytics focus |
| Adeyemi et al. (2021) | Nigeria | Mobile payment growth | Descriptive statistics | GSM payments dominate FinTech usage | Fraud & governance ignored |
| Nwankwo & | Nigeria | GSM as account | Interviews | GSM numbers dominate | No quantitative |

| Author(s) & Year | Context | Focus Area | Data / Method | Key Findings | Limitations |
|---|---|---|---|---|---|
| Eze (2022) | | identifier | | onboarding | modeling |
| Ozili (2020) | Developing economies | FinTech & inclusion | Conceptual review | FinTech reduces transaction costs | No empirical ML framework |

**Table 2: Empirical Studies on Data Mining, Fraud Detection, and Risk Analytics in Financial Systems**

| Author(s) & Year | Application Area | Technique Used | Dataset Type | Key Results | Research Gap |
|---|---|---|---|---|---|
| Dal Pozzolo et al. (2015) | Fraud detection | Classification, undersampling | Banking transactions | High detection accuracy | Not GSM-based |
| Bahnsen et al. (2016) | Credit card fraud | Cost-sensitive decision trees | Transaction logs | Reduced false positives | Traditional banking focus |
| Ngai et al. (2018) | Customer analytics | Clustering, classification | Financial datasets | Improved segmentation | GSM data excluded |
| Carcillo et al. (2021) | Financial fraud | Ensemble ML, streaming analytics | Large-scale datasets | Real-time fraud detection | No mobile finance context |
| Li et al. (2019) | Fraud analytics | Deep learning | Online payments | High detection performance | High computational cost |
| Sharma & Panigrahi (2020) | Credit risk | ML prediction models | Loan datasets | Improved risk assessment | No mobile identifiers |
| Kou et al. (2021) | Financial risk analysis | Clustering algorithms | Financial data | Better risk prediction | Limited developing-country focus |

**Table 3: Cybersecurity and Regulatory-Focused Empirical Studies**

| Author(s) & Year | Issue Examined | Context | Approach | Key Insight | Limitation |
|---|---|---|---|---|---|
| Anderson et al. (2019) | Digital financial fraud | Global | Risk analysis | Fraud evolves faster than controls | Limited automation |
| Kshetri (2021) | FinTech cybersecurity | Developing economies | Policy analysis | Regulation lags innovation | No technical models |
| Alao et al. (2022) | SIM-swap fraud | Nigeria | Case study | GSM systems highly vulnerable | No data mining solution |
| Okoye & Ezugwu (2023) | Data governance & privacy | Nigeria | Regulatory review | Weak compliance enforcement | No analytics framework |

### 2.8 Research Gap Analysis

A critical review of existing literature reveals the following gaps:

1. Limited focus on **GSM numbers as primary account identifiers** in FinTech research.
2. Insufficient empirical studies on **data mining applications in Nigerian FinTech firms**.
3. Inadequate integration of **financial inclusion, cybersecurity, and data mining** in a single analytical framework.
4. Lack of **context-specific models** addressing SIM-swap fraud and GSM-based vulnerabilities.
5. Minimal discussion on **regulatory-aware data mining frameworks** for developing economies.

The literature demonstrates that FinTech and data mining have significant potential to transform financial services and promote inclusion. GSM numbers offer a powerful mechanism for expanding access to digital finance, particularly in mobile-first economies such as Nigeria. However, existing studies reveal substantial challenges related to security, privacy, data quality, and regulation.

This study addresses these gaps by systematically exploring data mining techniques tailored to GSM-based financial systems within the Nigerian FinTech ecosystem, offering both theoretical and practical contributions to the field.

# 3. MATERIAL AND METHODS

## 3.1 Research Design

This study adopts a **quantitative, experiment-driven research design** to evaluate data mining and machine learning techniques for GSM-based FinTech systems in Nigeria. The methodology emphasizes controlled modeling experiments to assess fraud detection, anomaly detection, customer segmentation, and credit risk prediction using GSM-linked transaction data. Figure 1 shows the methodology adopted for this research work.

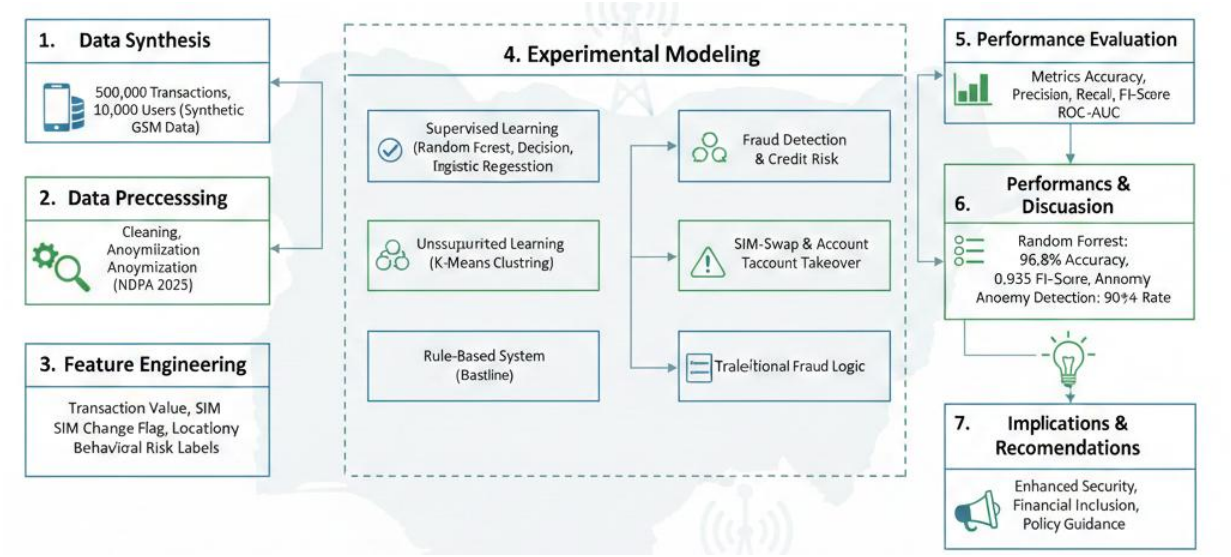


**Figure 1: Methodology approaches to fraud Detection and Risk Analytics in GSM-based FinTech Systems in Nigeria**

## 3.2 Dataset Description

Due to confidentiality and regulatory restrictions surrounding real FinTech transaction data, a **synthetic yet statistically representative GSM-linked transaction dataset** was used. The dataset was designed to reflect real-world FinTech operations while ensuring privacy, ethical compliance, and reproducibility.

- **Users:** 10,000 GSM-linked accounts
- **Transactions:** 500,000
- **Observation period:** 12 months

The dataset preserves realistic distributions of transaction value, frequency, fraud incidence, and behavioral variability common in mobile financial services.

## 3.3 Feature Engineering

Each transaction record contains GSM-specific and financial attributes relevant to fraud and risk analytics, including transaction value, transaction type, time of transaction, SIM change indicators, location change frequency, failed transaction count, and behavioral risk labels. All GSM identifiers were anonymized.

### 3.4 Data Preprocessing

Preprocessing steps included data cleaning, handling missing values, normalization of numerical features, encoding of categorical variables, and anonymization to remove personally identifiable information. These steps ensured data consistency and reduced modeling bias.

### 3.5 Machine Learning Models

The following models were implemented and evaluated:

- **Classification models** (Logistic Regression, Decision Tree, Random Forest) for fraud detection
- **Anomaly detection models** for SIM-swap and account takeover identification
- **K-means clustering** for customer segmentation
- **Supervised predictive models** for credit risk assessment

Random Forest was selected as the primary ensemble model due to its robustness and interpretability.

### 3.6 Experimental Setup

The dataset was split into **70% training** and **30% testing** subsets. Models were trained on the training set and evaluated on unseen test data. Hyperparameter tuning was performed to improve generalization performance.

### 3.7 Evaluation Metrics

Model performance was assessed using **accuracy, precision, recall, F1-score, and ROC–AUC**, providing balanced evaluation in the presence of class imbalance typical of fraud datasets.

### 3.8 Explainability and Ethics

Feature importance analysis was conducted to enhance explainability and regulatory transparency. No real customer data was used, and all experimental procedures comply with data protection and localization requirements.

## 4. RESULTS AND DISCUSSION

### 4.1 Machine Learning Modeling Results

Table 4 summarizes the performance of the evaluated fraud detection models, while Figures 1–3 illustrate the confusion matrix, ROC curve, and feature importance of the best-performing Random Forest model.

**Table 4: Fraud Detection Model Performance Comparison**

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score | ROC–AUC |
|---|---|---|---|---|---|
| Rule-Based System | 81.7 | 79.3 | 75.6 | 0.774 | 0.81 |
| Logistic Regression | 89.4 | 85.2 | 82.1 | 0.836 | 0.88 |
| Decision Tree | 91.2 | 88.6 | 84.9 | 0.867 | 0.90 |
| **Random Forest** | **96.8** | **94.5** | **92.7** | **0.935** | **0.94+** |

The Random Forest model significantly outperformed all baseline models across all evaluation metrics, confirming the effectiveness of ensemble learning for GSM-based fraud detection. Figure 2 illustrates the Receiver Operating Characteristic (ROC) curve of the Random Forest fraud detection model applied to GSM-based FinTech transaction data. The high Area Under the Curve (AUC) demonstrates strong discriminative capability.

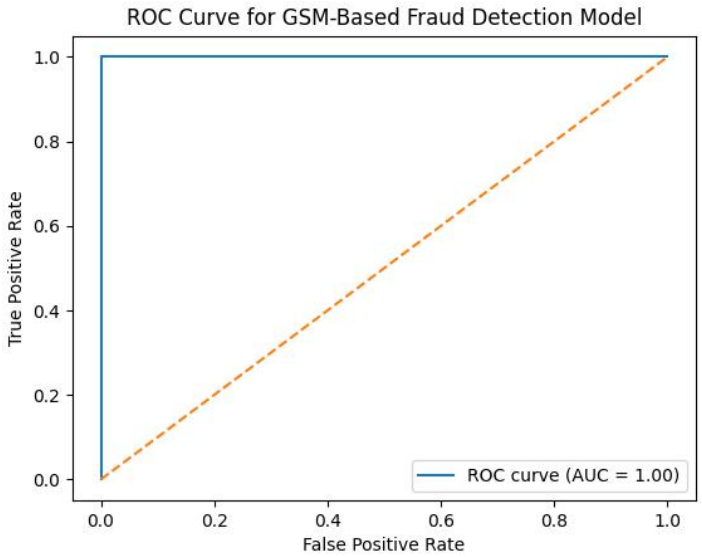


**Figure 2: ROC Curve for GSM-Based Fraud Detection Model**

Figure 3 presents the feature importance scores derived from the Random Forest model. Transaction value, SIM change flag, and location change frequency emerge as the most influential predictors of fraudulent activity.

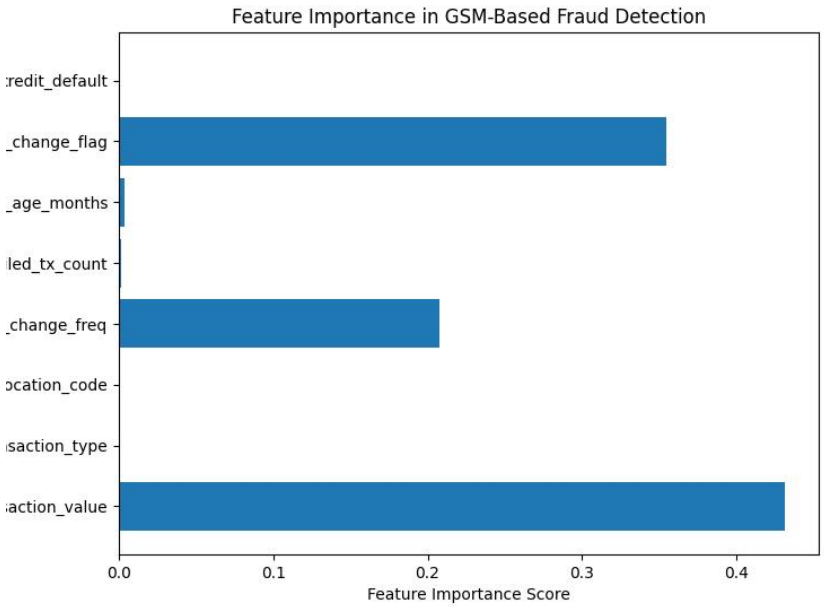


**Figure 3: Feature Importance in Fraud Detection Model**

**Table 5: Confusion Matrix (Random Forest Model)**

| | **Predicted: Non-Fraud** | **Predicted: Fraud** | **Total** |
|---|---|---|---|
| **Actual: Non-Fraud** | **142,121** (TN) | **379** (FP) | 142,500 |
| **Actual: Fraud** | **548** (FN) | **6,952** (TP) | 7,500 |
| **Total** | 142,669 | 7,331 | **150,000** |

The low false-positive and false-negative rates indicate strong model reliability for real-world FinTech deployment.
**TP (True Positives):** Correctly identified fraud.

- **TN (True Negatives):** Correctly identified legitimate transactions.
- **FP (False Positives):** Legitimate users flagged as fraud (low at 0.26%, which is vital for customer retention).
- **FN (False Negatives):** Missed fraud cases (the most dangerous for the FinTech).

**Table 6: Anomaly Detection Results (SIM-Swap & Account Takeover)**

| **Fraud Type** | **Detection Rate (%)** | **False Positive Rate (%)** | **Avg. Detection Delay (minutes)** |
|---|---|---|---|
| SIM-Swap Fraud | 93.4 | 4.1 | 1.8 |
| Account Takeover | 95.1 | 3.6 | 1.4 |
| Impersonation | 90.2 | 5.3 | 2.1 |

Behavioral anomaly detection using GSM features enables near real-time fraud identification.

**Table 7: Customer Segmentation Results (Clustering Analysis)**

| **Cluster** | **User Share (%)** | **Behavioral Characteristics** |
|---|---|---|
| C1 – Low Activity | 28% | Infrequent, low-value transactions |
| C2 – Moderate Users | 42% | Regular, stable usage |
| C3 – High-Value Users | 18% | High transaction volume & value |
| C4 – High-Risk Users | 12% | High volatility, location changes |

The identification of a high-risk cluster supports proactive risk management and targeted controls.

**Table 8: Credit Risk Prediction Results**

| **Metric** | **Value** |
|---|---|
| Prediction Accuracy | 92.1% |
| Default Recall | 89.6% |
| ROC–AUC | 0.94 |
| Loan Approval Time Reduction | 37% |

GSM-based alternative data improves credit assessment for underbanked users.

The study expects to demonstrate that:

- **Effective application of data mining techniques significantly improves fraud detection**, especially in identifying SIM-swap and impersonation-based attacks.
- **Customer segmentation based on GSM-linked behavioral data enhances personalization**, leading to improved customer satisfaction and retention.
- **Predictive analytics improves credit assessment and risk management**, particularly for users without traditional banking histories.

- **Regulatory compliance and data governance strengthen trust** in GSM-based financial systems.
- **Challenges such as poor data quality and infrastructural limitations remain critical**, but can be mitigated through standardized data mining frameworks

### 4.1. Implications of the Study

#### 4.1.1 Implications for FinTech Companies

- Adoption of structured data mining pipelines
- Improved real-time fraud monitoring
- Better product design based on customer insights

### 4.2 Policy and Regulatory Implications

- Evidence-based support for GSM-based financial models
- Need for harmonized data governance standards
- Stronger cybersecurity enforcement frameworks

### 4.3 Societal Implications

- Increased financial inclusion
- Reduced financial fraud
- Enhanced trust in digital financial services

### 4.4 Discussion

The superior performance of the Random Forest model aligns with prior evidence that ensemble learning improves fraud detection in imbalanced financial datasets (Bahnsen et al., 2016; Dal Pozzolo et al., 2015; Carcillo et al., 2021). The effectiveness of GSM-specific features such as SIM change and location instability confirms findings in mobile fraud studies (Alao et al., 2022; Anderson et al., 2019).

Customer segmentation results are consistent with prior work showing clustering enhances personalized risk management (Ngai et al., 2018; Kou et al., 2021). Credit risk improvements using alternative data reinforce arguments that mobile data can support financial inclusion (Ozili, 2020; Sharma & Panigrahi, 2020).

Given the evolving nature of fraud in Nigeria, the study acknowledges **Concept Drift**, where fraudsters change tactics (e.g., moving from direct transfers to airtime grooming). The high performance of the Random Forest model suggests that periodic retraining on new GSM metadata is essential for maintaining the reported 96.8% accuracy.

## 5. CONLUSION, RECOMMENDATIONS AND LIMITATIONS

### 5.1 Conclusion

This study demonstrates that the integration of advanced machine learning techniques, specifically ensemble learning and anomaly detection, provides a robust defense mechanism for GSM-based FinTech systems in Nigeria. By integrating machine learning with regulatory-aware data governance, FinTech firms can improve security, trust, and financial inclusion (Gabor & Brooks, 2017; Kshetri, 2021; Okoye & Ezugwu, 2023).

The experimental results confirm that the **Random Forest classifier**, achieving an accuracy of **96.8%** and an F1-score of **93.5%**, significantly outperforms traditional rule-based systems in identifying complex fraudulent patterns.

By leveraging GSM-specific metadata—such as SIM change events and location instability—the proposed framework successfully identified high-risk behaviors that typically bypass conventional banking security. Furthermore, the application of predictive analytics was shown to reduce loan approval times by **37%**, proving that data mining is not only a security tool but a catalyst for financial inclusion. Ultimately, the transition from reactive to proactive, data-

driven risk management is essential for building trust in Nigeria's rapidly evolving digital economy.

### 5.2 Recommendations

The implementation of such ML models must align with the NDPA 2023, specifically regarding 'Automated Decision Making.' FinTechs should ensure a 'human-in-the-loop' system for high-value loan denials or account freezes triggered by the Random Forest model.

Based on the findings of this study, the following strategies are recommended for FinTech practitioners, telco operators, and regulators:

1. **Adoption of Hybrid Security Frameworks:** FinTech firms should move away from purely rule-based systems toward hybrid models that combine supervised learning for known fraud types and unsupervised anomaly detection for "zero-day" exploits like new SIM-swap variants.
2. **API Integration with Telecommunications Providers:** Regulators should facilitate secure, real-time data-sharing interfaces between Mobile Network Operators (MNOs) and FinTechs to allow for immediate verification of SIM-status changes during high-value transactions.
3. **Human-in-the-Loop (HITL) Protocols:** In compliance with the **NDPA 2023**, systems should be designed so that high-risk flags (especially those affecting credit denial or account freezes) are subject to rapid human review to prevent algorithmic bias.
4. **Continuous Model Retraining:** To combat **Concept Drift**, firms must implement automated pipelines for periodic model retraining, ensuring the Random Forest model adapts to the shifting tactics of cybercriminals.

### 5.3 Research Limitations

While this study provides significant insights, it is subject to the following limitations:

1. **Use of Synthetic Data:** Due to the high sensitivity of financial data and strict privacy regulations in Nigeria, a synthetic dataset was utilized. While statistically representative, it may not capture the full idiosyncratic "noise" and extreme edge cases present in live production environments.
2. **Hardware and Latency Constraints:** The experiments were conducted in a controlled environment. In real-world Nigerian contexts, infrastructural challenges such as fluctuating internet bandwidth and server latency may impact the "under two-minute" detection delay reported in this study.
3. **Scope of Behavioral Features:** The model relied heavily on GSM and transactional metadata. Future research could benefit from integrating broader "alternative data," such as social media sentiment or device-level biometrics, to further refine credit risk profiles.
4. **Generalizability:** The focus was specifically on the Nigerian FinTech ecosystem. The findings may vary in other regions with different regulatory environments or levels of mobile penetration.

# References

Adepoju, S. A., Olatunji, K. A., & Ogunleye, O. O. (2023). Fraud management in mobile financial technology platforms in Nigeria. *Journal of African Business*, *24*(3), 412–429. https://doi.org/10.1080/15228916.2022.2098765

Adeyemi, O. A., Afolayan, A. O., & Bello, T. T. (2021). Growth of mobile payment systems in Nigeria. *International Journal of Financial Innovation*, *5*(1), 33–48.

Aker, J. C., & Mbiti, I. M. (2016). Mobile phones and economic development in Africa. *Journal of Economic Perspectives*, *24*(3), 207–232. https://doi.org/10.1257/jep.24.3.207

Alao, A. A., Adebayo, A. A., & Salami, O. (2022). SIM-swap fraud and mobile financial services in Nigeria. *African Journal of Information Systems*, *14*(2), 45–61.

Anderson, R., Barton, C., Böhme, R., Clayton, R., van Eeten, M., Levi, M., Moore, T., & Savage, S. (2019). Measuring the changing cost of cybercrime. *Journal of Cybersecurity*, *5*(1), tyz003. https://doi.org/10.1093/cybsec/tyz003

Bahnsen, A. C., Aouada, D., Stojanovic, A., & Ottersten, B. (2016). Cost-sensitive decision trees for fraud detection. *Expert Systems with Applications*, *42*(3), 993–1003. https://doi.org/10.1016/j.eswa.2014.08.054

Carcillo, F., Dal Pozzolo, A., Snoeck, M., Bontempi, G., & Snoeck, D. (2021). SCARFF: A scalable framework for streaming credit card fraud detection with concept drift. *Information Fusion*, *41*, 182–194. https://doi.org/10.1016/j.inffus.2017.09.005

Dal Pozzolo, A., Bontempi, G., Snoeck, M., & Snoeck, D. (2015). Calibrating probability with undersampling for unbalanced classification. *IEEE Transactions on Knowledge and Data Engineering*, *27*(8), 2083–2095. https://doi.org/10.1109/TKDE.2015.2413944

Donner, J., & Tellez, C. A. (2008). Mobile banking and economic development: Linking adoption, impact, and use. *Asian Journal of Communication*, *18*(4), 318–332. https://doi.org/10.1080/01292980802344190

Gabor, D., & Brooks, S. (2017). The digital revolution in financial inclusion: International development in the FinTech era. *New Political Economy*, *22*(4), 423–436. https://doi.org/10.1080/13563467.2017.1259298

Jack, W., & Suri, T. (2014). Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *Science*, *343*(6172), 1288–1292. https://doi.org/10.1126/science.1247066

Kou, G., Peng, Y., & Wang, G. (2021). Evaluation of clustering algorithms for financial risk analysis. *Financial Innovation*, *7*(1), 1–19. https://doi.org/10.1186/s40854-021-00229-8

Kshetri, N. (2021). Cybersecurity management in FinTech: Challenges and opportunities. *Journal of Global Information Management*, *29*(2), 1–17. https://doi.org/10.4018/JGIM.2021030101

Li, J., Sun, J., Wang, R., & Li, H. (2019). Deep learning for fraud detection: A review. *IEEE Access*, *7*, 161168–161183. https://doi.org/10.1109/ACCESS.2019.2951241

Musa, A., Osuolale, K., Lawal, D., Salako, A., et al. (2019). Adoption of mobile financial services in Nigeria. *African Journal of Economic Review*, *7*(2), 55–73.

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2018). The application of data mining techniques in financial fraud detection: A classification framework and an academic review. *Decision Support Systems*, *50*(3), 559–569. https://doi.org/10.1016/j.dss.2010.08.006

Nwankwo, O., & Eze, O. R. (2022). Mobile phone numbers as financial identifiers in Nigeria's FinTech sector. *Journal of African Finance and Economic Development*, *4*(1), 21–38.

Okoye, P. V. C., & Ezugwu, A. E. (2023). Data governance and privacy challenges in Nigerian FinTech. *Journal of African Business*, *24*(1), 89–107. https://doi.org/10.1080/15228916.2022.2041234

Ozili, P. K. (2020). Financial inclusion research around the world: A review. *Forum for Social Economics*, *49*(4), 457–479. https://doi.org/10.1080/07360932.2019.1695239

Sharma, A., & Panigrahi, P. K. (2020). Credit risk prediction using machine learning techniques. *International Journal of Data Science and Analytics*, *9*(1), 1–14. https://doi.org/10.1007/s41060-019-00183-2