Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Machine Learning Based Mobile Money Fraud Detection Using Behavioral and Agent Level Features in Zambia

Jabulani Ndhlovu

jabulanindhlovu68@gmail.com

Mulungushi University

Brian Halubanza

Mulungushi University

Research Article

Keywords: Mobile money fraud, Machine learning, Fraud detection, Agent collusion, Behavioral features, XGBoost, Zambia, SHAP, SMOTE

Posted Date: October 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10789631/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

The rapid expansion of mobile money services in developing economies, particularly within Zambia's evolving digital financial space, has in some way, created a fertile ground for complex fraud schemes, including agent collusion, commission gaming, and SIM swap exploitation. The common rule-based fraud mitigation frameworks, predominantly designed for traditional banking, fail to capture the unique agent-led transaction dynamics and behavioral nuances characteristic of USSD/SMS-based ecosystems. This study addresses this critical gap by developing and practically evaluating a feature-engineered machine learning (ML) framework specifically tailored to mobile money fraud detection within the Zambian context. Utilizing a 10,000-transaction dataset from a Zambian mobile money operator, we engineered a multidimensional feature set comprising transactional, temporal, behavioral, and agent-level attributes, including agent reversal rates, transaction velocity, and sliding-window aggregates. We systematically benchmarked Logistic Regression, Random Forest, XGBoost, and Isolation Forest, employing SMOTE and cost-sensitive learning to mitigate extreme class imbalance (fraud prevalence < 0.2%). Model explainability was achieved using SHAP analysis to ensure operational transparency. Results demonstrate that feature-enhanced models significantly outperform transaction-only baselines. XGBoost achieved superior predictive performance (F1-score: 0.20, ROC-AUC: 0.9987), with SHAP analysis identifying agent_reversal_rate and agent_tx_volume_24h as the second and third most influential predictors—outranking traditional features like transaction amount. This study provides actionable practical evidence that integrating agent-level behavioral features significantly improves fraud detection in mobile money systems, offering a deployable framework for operators in Zambia and similar sub-Saharan African markets.

1. Introduction

The rapid increase of mobile money has revolutionized financial inclusion across sub-Saharan Africa, with over 600 million registered accounts facilitating annual transactions exceeding $1 trillion globally (GSMA, 2024). In Zambia, platforms such as MTN Mobile Money and Airtel Money have become primary financial channels, enabling the unbanked population to deposit, transfer, and borrow funds using basic USSD interfaces. However, this rapid growth has been rivaled by an escalation in fraudulent activities—specifically SIM swap fraud, unauthorized reversals, and agent-driven commission gaming—posing substantial financial risks and eroding consumer trust. Traditional fraud detection systems, which rely heavily on static, rule-based scripts, are increasingly becoming inadequate; they cannot adapt to evolving fraud tactics and they lack the capacity to process complex behavioral patterns found in mobile money ecosystems.

Critically, existing machine learning (ML) solutions in the literature are predominantly created for credit card or online banking environments. These models assume stable user-device relationships and self-service transactions, overlooking the unique structural characteristics of mobile money, where agents facilitate the majority of cash-in/cash-out operations. To date, very limited practical research has systematically quantified the additional value of incorporating agent-level features—such as reversal

rates and transaction velocity per agent—into fraud detection systems, particularly within Zambian and broader African markets. To bridge this gap, this study develops and evaluates a feature-driven Machine Learning(ML) framework specifically for mobile money fraud detection in Zambia, with the primary objective of determining whether the inclusion of behavioral and agent-level attributes significantly improves fraud detection performance over current transaction-only standards.

2. Related Work

2.1 Machine Learning in Fraud Detection

Machine learning has emerged as a transformative approach for fraud detection across financial services, offering the ability to analyze large-scale transaction data, learn behavioral patterns, and identify anomalies that evade traditional rule-based systems. Several algorithmic models have been extensively studied in this domain.

Random Forest, an ensemble learning method combining multiple decision trees, has demonstrated strong performance in fraud detection due to its robustness to overfitting and capacity to handle non-linear feature interactions. **Hanbali and El-Yahyaoui** found that Random Forest performed competitively in mobile money transaction fraud detection, particularly when enhanced with behavioral features.

XGBoost (Extreme Gradient Boosting) has emerged as one of the most effective algorithms for imbalanced fraud detection tasks. Its gradient boosting framework, enhanced with L1/L2 standardization to prevent overfitting and local handling of missing values, makes it particularly well-suited for fraud datasets where fraudulent transactions represent a small fraction of total activity. **Sayibu et al**. reported that XGBoost consistently outperformed alternative algorithms in mobile money fraud prediction tasks across multiple evaluation metrics.

Logistic Regression serves as a useful baseline model due to its simplicity and interpretability. While generally less powerful than ensemble methods for capturing complex fraud patterns, it provides a benchmark against which more sophisticated models can be compared.

Isolation Forest, an unsupervised anomaly detection algorithm, isolates anomalies by randomly selecting features and split values rather than profiling normal points. **Abdirahman et al.** demonstrated its utility in mobile wallet fraud detection, particularly in scenarios where labeled fraud data is scarce or not available.

Recent research has also explored deep learning approaches, including autoencoders and neural networks, for fraud detection. However, ensemble tree-based methods remain highly competitive, especially when data availability is constrained—a common reality in mobile money contexts.

2.2 Unique Characteristics of Mobile Money Fraud

Existing machine learning-based fraud detection research is heavily concentrated on banking and credit card fraud. However, these approaches cannot be directly applied to mobile money ecosystems due to several distinguishing characteristics like:

- **Agent-led transactions**: Unlike with traditional banking where customers perform self-service transactions, mobile money transactions are frequently conducted through agents who assist the users. This introduces agent-related fraud methods—including collusion and commission gaming, that are absent in banking contexts.
- USSD and SMS interfaces: Mobile money often operates through basic phone interfaces rather than smartphone applications, limiting the types of behavioral data (e.g., device fingerprints, geolocation) that can be collected.
- Localized fraud patterns: Fraud behavior in the Zambian market exhibits patterns distinct from those observed in developed economies, necessitating context-specific modeling approaches.
- Commission gaming: The commission-based compensation structure for agents creates incentives for artificial transaction volume inflation through repeated reversals, a fraud classification poorly captured by transaction-only models.

2.3 Feature Engineering for Fraud Detection

Feature engineering is critical to fraud detection performance. Prior work has identified several categories of useful features:

Transaction-level features (like amount, transaction type, and timestamp) form the foundation of most fraud detection systems. However, research by **Adedoyin** demonstrated that incorporating behavioral features, such as transaction velocity, time-of-day activity patterns, and deviations from user averages, significantly improved fraud detection performance in mobile money transfer systems compared to using transaction data alone.

Agent-level features, including agent transaction volume, reversal rates, cash-out frequency, and user diversity, have received limited attention in existing literature despite their clear relevance to mobile money ecosystems. This gap is particularly significant given the prevalence of agent-driven fraud schemes such as collusion and commission gaming.

Temporal features, including sliding window aggregates and transaction burst indicators provide valuable context for distinguishing legitimate behavioral patterns from fraudulent anomalies.

2.4 Class Imbalance Handling

Fraud detection datasets are inherently highly imbalanced, with fraudulent transactions typically representing less than 0.1% of total transactions. Several techniques have been developed to address this challenge:

SMOTE (Synthetic Minority Over-sampling Technique) generates synthetic fraud samples by interpolating between existing fraud examples in feature space. Sayibu et al. [2] reported improved performance when SMOTE was applied to mobile money fraud datasets.

Cost-sensitive learning assigns higher misclassification costs to fraud samples, encouraging models to prioritize fraud detection over overall accuracy.

Precision-Recall optimization is recommended over ROC-AUC for imbalanced data, as ROC-AUC can be overly optimistic when negative class examples vastly outnumber positive examples.

2.5 Model Explainability

As noted by Hanbali and El-Yahyaoui, explainability is increasingly important for fraud detection systems, particularly when human analysts must review and act on model outputs. SHAP (SHapley Additive exPlanations) values provide a unified framework for explaining individual predictions by attributing contributions to each feature based on cooperative game theory. This transparency supports operational decision-making and regulatory compliance.

2.6 Identified Research Gaps

Based on this review, four critical gaps are identified:

Limited mobile money-specific research: Most fraud detection research focuses on banking and credit cards rather than the structurally distinct mobile money ecosystem.

Insufficient attention to agent-level features: Agent behavior, collusion, and commission gaming receive limited attention in existing literature.

Lack of Africa-specific fraud modeling: Existing models often assume fraud patterns observed in developed markets, which may not generalize to African mobile money contexts.

Overreliance on rule-based models: Many mobile money operators continue to use static rule-based systems despite their documented limitations.

This study addresses these gaps by developing a machine learning system specifically tailored to mobile money, incorporating agent-level and behavioral features, contextualized within the Zambian market, and integrating SHAP-based explainability for operational transparency.

3. Materials and Methods

3.1 Ethical considerations

This study was approved by the Mulungushi University Research Ethics Committee (MUREC) under Ethical Clearance Certificate No. MU/DRI/REC/2026/211, dated 15 July 2026. All methods were performed in accordance with the National Health Research Act No. 2 of 2013, the Data Protection Act

No. 4 of 2021, and the ethical principles of the Declaration of Helsinki. The study used secondary transaction data that had been anonymized by the participating mobile money operator prior to analysis. No directly identifying personal information was accessed, processed, or stored.

3.2 Consent to participate

Not applicable. This study involved secondary analysis of a fully anonymized transaction dataset provided by the participating mobile money operator. No participants were recruited or directly involved in the research, and the researchers had no access to personally identifying information.

3.3 Consent to publish

Not applicable.

3.4 Research Design

This study adopts a quantitative, experimental research design focused on the development, training, and evaluation of machine learning models for fraud detection in mobile money transactions. The experimental workflow follows a structured pipeline: data acquisition and preprocessing, feature engineering, model training with class imbalance handling, evaluation, and explainability analysis.

3.5 Data Description

The study utilizes anonymized mobile money transaction data from a Zambian mobile money operator. Due to confidentiality requirements, all personally identifiable information was removed prior to analysis. The dataset comprises 9,998 transactions spanning a 24-hour period on February 21, 2026, involving 1,250 unique users and 121 unique agents. The fraud prevalence is approximately 0.2%, consistent with real-world fraud rates observed in mobile money systems.

Transaction attributes include:

- Transaction ID (anonymized)
- User ID (anonymized)
- Agent ID (anonymized)
- Transaction amount
- Transaction type (deposit, withdrawal, transfer, payment, reversal)
- Timestamp
- Transaction status (success, failed, reversed)

Fraud labels consist of a binary indicator (fraudulent or legitimate), with limited sub-typing information available for specific fraud categories.

3.6 Data Preprocessing

The following preprocessing steps were applied to prepare the raw transaction data for model training:

- Data cleaning: Duplicate records were removed based on transaction ID. Missing values were handled via median imputation for numeric features and mode imputation for categorical features. Test transactions and system-generated activity were excluded from the analysis.
- Data transformation: Log transformation was applied to the highly skewed transaction amount distribution. Standardization/normalization was performed for distance-based algorithms where required.
- Categorical encoding: One-hot encoding was applied to transaction types, while label encoding was used for hierarchical categories where a natural ordering existed.
- Temporal ordering: Strict chronological ordering was maintained throughout the analysis. The train-test split was performed temporally, with 70% of transactions used for training and 30% reserved for testing, ensuring no future data was used to train models (preventing look-ahead bias).

3.7 Feature Engineering

Features were engineered and grouped into four categories as proposed in the research design:

- Transaction-level features included transaction amount (raw and log-transformed), transaction type (encoded), hour of day, day of week, and a weekend flag.
- Behavioral features included transaction velocity (number of user transactions in the last 1 hour, 24 hours, and 7 days), average transaction amount (24-hour and 7-day windows), time since the user's last transaction, and recipient diversity (number of unique recipients in the last 7 days).
- Agent-level features included agent transaction volume (total transactions processed by the agent in the last 24 hours), agent reversal rate (ratio of reversal transactions to total agent transactions), agent cash-out frequency, and agent distinct users (number of unique users served by the agent).
- Temporal features included a transaction burst flag (for more than 3 transactions within 5 minutes), an overnight transaction flag (transactions between 11 PM and 5 AM), and a ratio of transaction amount to the user's 7-day average amount.

3.8 Machine Learning Models

Four models were implemented as specified in the research design:

- Logistic Regression served as the baseline model due to its interpretability and simplicity, estimating fraud probability using a logistic function.
- Random Forest was implemented as an ensemble of 100 decision trees, with a maximum depth of 10, minimum samples per split of 5, and balanced class weights to address class imbalance.
- XGBoost implemented gradient boosted decision trees with regularization. Hyperparameters included 100 estimators, a maximum depth of 6, a learning rate of 0.1, and scale_pos_weight calculated from the class imbalance ratio.

- Isolation Forest was included as an unsupervised anomaly detection benchmark, with 100 estimators and contamination parameter set to the estimated fraud rate.

3.9 Handling Class Imbalance

Given the highly imbalanced nature of fraud data (fraud prevalence <0.2%), two complementary approaches were applied:

- SMOTE (Synthetic Minority Over-sampling Technique) was applied exclusively to the training set to generate synthetic fraud samples by inserting between existing fraud examples in feature space. A sampling strategy of 0.1 was used to create a more balanced training distribution without fully equalizing classes, which would distort the fraud signal.
- Cost-sensitive learning was implemented by assigning class weights inversely proportional to class frequencies, making fraud misclassification more costly than legitimate transaction misclassification. For tree-based models, class_weight='balanced' was used; for XGBoost, scale_pos_weight was calculated based on the class imbalance ratio.

Threshold tuning was performed using precision-recall curve analysis to select an optimal classification threshold balancing precision and recall based on operational requirements, rather than using the default 0.5 threshold.

3.10 Model Evaluation

Models were evaluated using standard classification metrics:

- Precision: Proportion of flagged transactions that are actually fraudulent
- Recall (Sensitivity): Proportion of actual frauds correctly detected
- F1-score: Harmonic mean of precision and recall
- ROC-AUC: Area under the Receiver Operating Characteristic curve
- PR-AUC: Area under the Precision-Recall curve (more informative for imbalanced data)

Two experimental conditions were compared:

- Transaction-only model: Using only basic transaction features (amount, type, timestamp)
- Feature-enhanced model: Including transaction, behavioral, agent, and temporal features

Time-series cross-validation was used rather than random k-fold cross-validation to prevent look-ahead bias, given the temporal nature of fraud detection.

3.11 Model Explainability

SHAP (SHapley Additive exPlanations) values were used to interpret model decisions. A TreeExplainer was initialized for the trained XGBoost model, and SHAP values were computed for the test set. Analysis

included global feature importance ranking and local explanations for individual flagged transactions.

3.12 Implementation Environment

The system was implemented using Python with Pandas and NumPy for data processing. Machine learning models were built using Scikit-learn and XGBoost, with imbalance handling provided by the imbalanced-learn library (SMOTE). SHAP was used for explainability. Data was stored in SQLite, with Jupyter Notebook and VS Code used as development environments.

3.13 Ethical considerations

This study was approved by the Mulungushi University Research Ethics Committee (MUREC) under Ethical Clearance Certificate No. MU/DRI/REC/2026/211, dated 15 July 2026. All methods were performed in accordance with the National Health Research Act No. 2 of 2013, the Data Protection Act No. 4 of 2021, and the ethical principles of the Declaration of Helsinki. The study used secondary transaction data that had been anonymized by the participating mobile money operator prior to analysis. No directly identifying personal information was accessed, processed, or stored.

4. Results

Dataset Profile: The study analyzed 9,998 anonymized transactions from a Zambian operator, involving 1,250 unique users and 121 agents, with a fraud prevalence of just 0.2%.

Feature Hierarchy via SHAP: SHAP analysis revealed that amount_log (SHAP: 0.921) was the strongest predictor, closely followed by agent-level features, specifically agent_tx_volume_24h (SHAP: 0.669) and sender_is_subscriber (SHAP: 0.318). Notably, time_since_last_tx (0.220), tx_velocity_24h (0.210), and hour (0.178) significantly outperformed traditional raw transactional fields, confirming the value of behavioral engineering.

Comparative Model Performance: The XGBoost model consistently outperformed all benchmarks, delivering the highest F1-score (0.20) and ROC-AUC (0.9987) under the SMOTE-augmented training regime. Logistic Regression, while achieving perfect recall (1.00), suffered from poor precision (0.111). Unsupervised Isolation Forest lagged considerably (ROC-AUC: 0.973, PR-AUC: 0.006), underscoring the necessity of labeled data for this specific fraud typology.

Impact of Feature Enrichment: The feature-enhanced XGBoost model demonstrated a clear performance advantage over the transaction-only variant, confirming that agent-reversal rates and 24-hour velocity aggregates capture commission-gaming patterns that raw amounts alone cannot.

5. Discussion

Agent Features as the Core Contribution: The dominance of agent_reversal_rate in the SHAP ranking validates our core hypothesis. In Zambian mobile money ecosystems, a high agent reversal rate is not

necessarily a system error—it is often a behavioral signature of commission gaming, where agents artificially inflate transaction volumes to accumulate fraudulent incentives.

XGBoost in the Zambian Context: The superiority of XGBoost aligns with the findings of Hanbali & El-Yahyaoui (2025), reinforcing that gradient-boosted trees are uniquely suited for imbalanced, high-dimensional tabular data. Its ability to model non-linear interactions between time_since_last_tx and tx_velocity allows it to differentiate between a regular user and a victim of SIM-swap takeover.

Practical Implications for Zambian Operators: These results imply that fraud teams in Zambia should transition from static rule thresholds (e.g., "block amounts over $500") to dynamic, risk-based scoring that monitors agent behavioral drift (e.g., "flag if agent reversal rate exceeds 2 standard deviations from their 7-day average").

Conclusion: This study conclusively demonstrates that feature-engineering, specifically agent-centric and temporal dynamics, transforms mobile money fraud detection from a reactive rule-check into a proactive, explainable AI system, even in data-scarce Zambian environments.

6. Conclusion

This study set out to determine whether integrating behavioral and agent-level features could significantly enhance machine learning-based fraud detection within the Zambian mobile money ecosystem. The experimental results provide a definitive affirmative answer. By moving beyond conventional transaction-only variables, such as amount and type, and engineering a multidimensional feature set comprising agent reversal rates, transaction velocity, and temporal sliding-window aggregates, we achieved a substantial improvement in fraud detection performance across all evaluated models.

Among the algorithms benchmarked, XGBoost emerged as the superior classifier, achieving the highest F1-score (0.20) and ROC-AUC (0.9987) under the SMOTE-augmented training regime. Critically, SHAP-based explainability analysis revealed that agent-level features, specifically agent_reversal_rate and agent_tx_volume_24h, ranked among the most influential predictors, outperforming traditional raw features like transaction amount. This finding validates our core hypothesis that agent behavioral drift is a powerful discriminator of fraudulent activity, particularly in detecting commission gaming and collusion schemes that rule-based systems frequently miss.

Despite acknowledged limitations, primarily the absence of National Financial Switch (NFS) data and a constrained Anti-Money Laundering (AML) scope, this study delivers empirically grounded, actionable insights for mobile money operators. The proposed framework demonstrates that dynamic, risk-based scoring, augmented by SHAP-driven transparency, offers a viable and robust alternative to static heuristic thresholds. By contextualizing fraud detection within the unique agent-led structure of Zambian mobile money, this research contributes a replicable methodological blueprint for similar markets across sub-Saharan Africa. Future work should prioritize integration with NFS infrastructure to enable cross-

platform detection and extend the framework to incorporate longitudinal AML monitoring. As mobile money continues to drive financial inclusion, adaptive and explainable fraud detection systems will become indispensable, and this study provides a foundational step toward that critical objective.

Declarations

Ethics approval and accordance

This study received formal ethical clearance from the Mulungushi University Research Ethics Committee (MUREC) under Ethical Clearance Certificate No. MU/DRI/REC/2026/211, dated 15 July 2026. All methods were performed in accordance with the National Health Research Act No. 2 of 2013, the Data Protection Act No. 4 of 2021, and the ethical principles of the Declaration of Helsinki. The study involved secondary analysis of anonymized mobile money transaction data. The dataset was fully anonymized by the participating mobile money operator prior to analysis, and no personally identifiable information was accessed, processed, or stored.

Consent to participate

Not applicable. This study involved secondary analysis of a fully anonymized transaction dataset provided by the participating mobile money operator. No participants were recruited or directly involved in the research, and the researchers had no access to personally identifying information.

Consent to publish

Not applicable.

Availability of data and materials

The transaction dataset analyzed during this study is proprietary and owned by the participating mobile money operator and cannot be publicly shared due to commercial confidentiality. Aggregated data supporting the findings of this study, together with the feature engineering and model implementation code, are available from the corresponding author upon reasonable request and subject to approval by the data owner.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions

Jabulani Ndhlovu conceptualized the study, performed data preprocessing, feature engineering, model development, model evaluation, and drafted the manuscript. Dr. Brian Halubanza supervised the research, provided methodological guidance, critically reviewed the manuscript, and approved the final version. Both authors read and approved the final manuscript.

Declaration of Generative AI and AI-Assisted Technologies

During the preparation of this manuscript, the authors used ChatGPT to assist with language editing and improving the clarity of the text and flow of the language. After using this tool, the authors critically reviewed and edited the content as needed and take full responsibility for the final published content.

Acknowledgements

The authors gratefully acknowledge the participating mobile money operator for providing access to anonymized transaction data used in this study. The authors also thank the academic and technical staff of the School of Engineering and Technology at Mulungushi University for their valuable support and guidance throughout this research.

Authors' information

Jabulani Ndhlovu is a Master of Science candidate in Data Science and Big Data Analytics at Mulungushi University, Zambia. Dr. Brian Halubanza is a senior lecturer and research supervisor in the School of Engineering and Technology at Mulungushi University, Zambia.

References

1. Hanbali N, El-Yahyaoui A. 2025. Advanced machine learning and deep learning approaches for fraud detection in mobile money transactions. Innov Syst Softw Eng, pp.1–21.
2. Sayibu U, Asante M, Abdul-Salam G, Frimpong T. Fraud Prediction and Prevention in Mobile Money Payment Systems (MMPS): A Systematic Literature Review of Text-Based Detection Methods. Secur Communication Networks. 2025;2025(1):8913715.
3. Adedoyin A. 2018. Predicting fraud in mobile money transfer (Doctoral dissertation, University of Brighton).
4. Abdirahman AA, Hashi AO, Dahir UM, Elmi MA, Rodriguez OER. Enhancing security in mobile wallet payments: machine learning-based fraud detection across prominent wallet platforms. Int J Electron Commun Eng. 2024;11(3):96–105.