Open access Journal International Journal of Emerging Trends in Science and Technology

DOI: http://dx.doi.org/10.18535/ijetst/v2024.05

Fraud Detection Combating Mobile Money Fraud in SMS Messages Using Machine Learning

Authors

Panji Msowoya[1], Dr Tawarish[2]

[1]Dmi-St Eugene University, Private bag 69, Lilongwe, Malawi

[2]Assistant Professor, Department of Information & System Management, The New College Chennai, India

Abstract

This paper presents a machine learning-based fraud detection model to combat the SMS based mobile money fraud. The systems use an XGBoost classifier to process SMS message content real-time and boast a high level of accuracy with predicting malicious intent. Built using Flask framework, the system allows administrators to monitor in real-time all reported alarms for specific transactions regarding fraud and keep statistical records. Future research includes improving the interpretability of models and making them both more scalable as well as robust to adversarial attacks for better fraud detection in mobile money services.

Keywords: *XGBoost classifier, Flask, Fraud detection, Mobile money fraud.*

Introduction

With few traditional banking services, mobile money has emerged in Malawi as a quick and easy way to bank in previously unbanked areas. Haowever, this widespread exploitation has also increased the risk of people being exposed to fraud on their devices, such as through text messages. One of the fraud methods used by fraudsters in Malawi is a phishing scam, where customers receive fake text messages from genuine mobile phone providers encouraging them to provide their account details and send money to clone accounts (CGAP, 2020). More common is unauthorized activity, where hackers access people's mobile phone accounts and make unauthorized transfers without users' permission. SIM card swapping has also become a complicated practice where criminals convince the mobile phone operator to change their victim's phone number to a new one in order to get control under the control of the sub-criminal himself and slip SMS-based transaction confirmations (ITU, 2019).

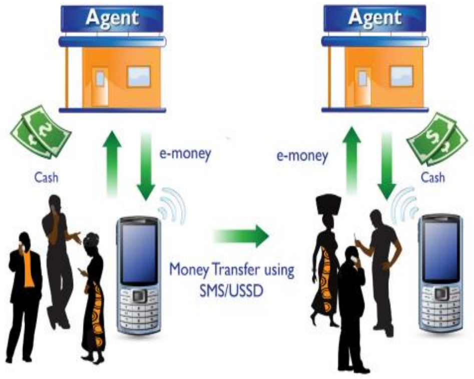


Figure 1: Overview of mobile money transactions

The impacts of mobile phone fraud do not end only with financial losses - it also extends to a serious threat to consumer confidence in mobile phone services. Fraud involves the financial exclusion of defrauded users to recover the stolen money through non-existent means, and this increases customer distrust (World Bank, 2021). This erosion of trust is problematic for mobile money providers based in Malawi working to increase adoption and promote financial inclusion. In addition, mobile fraud is a difficult problem to combat due to the general lack of regulation in the digital/mobile financial ecosystem era and the technical weaknesses of mobile networks. Sometimes, regulatory oversight in Malawi may not be strong enough to enforce the necessary strict security measures (or hold fraudsters accountable), and as recent reports show, the outdated security infrastructure, where mobile networks are so slow or fragile, makes them very vulnerable.

This research explores fraudulent SMS detection as a way to mitigate the risk factors associated with the detection of fraudulent SMS messages among mobile messaging service (MMS) customers in Malawi, where the high frequency of SMS messages has made them a prime target for phishing and fraud which led to unauthorized transactions.

By automating the detection process through text analysis and pattern recognition, machine learning offers promising avenues for enhancing security measures. This paper presents a practical implementation of a fraud detection system using machine learning models integrated into a Flask web application framework. The system provides real-time monitoring and reporting capabilities, aiming to mitigate the impact of fraudulent activities on mobile money users in Malawi.

Problem Definition

SMS-based frauds on mobile money services in Malawi, and similar emerging economies are known to send significant risks for consumers. These schemes exploit the pervasive use of SMS for financial transactions, aiming to deceive users into divulging sensitive information or authorizing unauthorized transactions. Fraudsters employ sophisticated tactics, including phishing scams, unauthorized access through SIM swapping, and manipulation of transaction confirmations via intercepted SMS messages. These methods not only lead to financial losses for victims but also erode trust in mobile money systems, hindering efforts to promote financial inclusion and digital payment adoption (World Bank, 2021; GSMA, 2020).

In phishing scams, for example, fraudsters send SMS converging on individuals encouraging them to click knowing these are messages from the mobile money providers. IT attempts to steal sensitive personal data such as PINs or credentials, often tricking the user into revealing them under a pretense of security verification or an urgent update their account required. The same way the crooks conduct unauthorized transactions due to breakdown of mobile money security protocols and gain unscrupulous access into users accounts without their knowing. Further, SIM swapping exploit the weak identity verification processes of mobile network operators which allows fraudsters to take over a victim's phone number and consequently intercept SMS-based transaction verifications (CGAP 2020).

The challenge is made worse by the inadequacy of traditional fraud detection systems based on rule-based approaches, which struggle to adapt to the evolving tactics used by fraudsters (CGAP, 2020). As a result, there is an urgent need for more sophisticated and adaptable solutions that use advanced technologies such as machine learning to effectively detect and mitigate SMS-based fraud. This paper addresses this shortcoming by introducing a machine learning-based fraud detection system. The system aims to improve security measures by automating the detection of fraudulent text messages in real time, which secures users' money transfers and restores trust in mobile phone services.

Objective

This research focuses on the development of an SMS fraud detection system tailored for mobile money services in Malawi. This system will use a machine learning model in order to effectively detect and mitigate fraudulent SMS messages. Integration with a Flask web application shall be able to provide real-time views with reporting on suspicious activities. Testing the practical applicability of the system and its effectiveness in enhancing mobile money users' security measures would therefore be done through pilot testing with a dataset of fraudulent SMS messages.

Literature Review

Research on SMS fraud detection systems has unsurprisingly advanced in relation to an increase in mobile money services and a related surge of fraudulent activities against SMS communications. A number of studies have been done on different approaches on the way forward to repression and methods of detecting fraud leveraging SMS, with an eye on benchmarking security measures in protecting users' financial transactions.

Chen et al. (2018) conducted research on SMS fraud detection using machine learning techniques and emphasized that accurate results for fraud detection depend mostly on feature engineering and model selection. The research highlighted effectiveness of ensemble methods such as Random Forest and Gradient Boosting Machines offer very good performance at differentiating between legitimate and fraudulent SMS messages based on textual features and behavioral patterns.

On the other hand, Gupta and Kapoor (2019) proposed a framework for the detection of phishing SMS messages through semantic analysis and by passing them through various clustering algorithms. Their approach had employed NLP techniques for extracting semantic features from SMSes for input into clustering algorithms that did group SMS messages into clusters indicative of fraudulent activities. This study demonstrated promising results in perfectly identifying phishing attempts and greatly reducing false positives in fraud detection systems.

Further, Li and Wang (2020), considered deep learning models in integrating convolutional neural networks (CNNs), for SMS fraud detection. The researchers focus on how to employ CNNs to automatically learn a representation at multiple levels of SMS message content in a hierarchical way to detect anomalous patterns indicative of fraudulent behavior. Their approach captures both local and global dependencies within SMS text data, hence improving the algorithmic accuracies of fraud detection systems and the robustness of such systems in dynamic mobile money environments.

While these studies demonstrate the different techniques or technologies that can be applied to SMS fraud detection, class imbalance, continuously evolving fraud tactics, and scalability remain challenges. Clearly, further research and innovation is needed to mitigate these risks and eventually enable adaptive and resilient fraud detection systems against SMS-based fraud in mobile money services.

Methodology

Dataset Description and Preprocessing

For this study, we utilized the SMS Spam Collection dataset from the UCI Machine Learning Repository, which consists of messages labeled as either "spam" or "ham". With “Spam” indicating fraudlent messages and “ham” indicating non-spam messages. We downloaded the dataset from the UCI Machine Learning Repository Endnote, accessed on: *https://archive.ics.uci.edu/dataset/228/sms+spam+collection*. The dataset provides an overall view on typical SMS messages in real-world scenarios and hence offers good suitability for training and evaluation in fraud detection models.

Before actual training, the final preprocessing steps were applied to this dataset to make it more fit for analysis. In this regard, noise removal was performed on text data, which involved removing irrelevant characters, symbols, and formatting inconsistencies. Afterward, the messages were tokenized into words or single tokens so that they could be analyzed. The textual data was then turned into numerical features by means of Term Frequency-Inverse Document Frequency vectorization. In it, every term is weighted based on its frequency within the document and across the dataset to really capture the importance of words in distinguishing spam from ham messages.

It is then that we take this preprocessed dataset to the phase of machine learning model building for training, so that the textual content of the SMS messages is perfectly learned to the machine learning models in order to distinguish between legitimate and fraudulent activities.

Machine Learning Model

For this study, considering that the XGBoost classifier has been very efficient in dealing with high-dimensional and sparse data, this is the machine learning model that shall be used in detecting fraud in SMS messages, making it very well-suited to doing text classification tasks like classifying fraudulent messages in mobile money transactions. XGBoost is an actual extreme gradient boosting based on the ensemble learning approach for building sequentially ensembles of weak decision trees to optimize predictive accuracy by iteratively minimizing a loss function. This confers a great advantage in capturing intrinsic interaction and patterns within the textual data of importance in distinction between legitimate and fraudulent SMS messages.

During the model training phase, the XGBoost classifier was trained against a preprocessed dataset drawn from the SMS Spam Collection dataset with annotated SMS messages as either fraudulent or non-fraudulent. The training process involved the conversion of the textual data into numerical features by the TF-IDF vectorizer, where the terms are weighted based on their frequency in individual messages and across the dataset. The system was further fine-tuned with optimized hyperparameters for achieving maximum fraud detection accuracy while keeping the number of false positives at a minimum, hence ensuring robustness in real-world applications. With those features, the model is trained to predict the probability of each incoming SMS being fraudulent.

The trained model provides a probability score, which is a quantitative measure of how likely it is that a message will exhibit characteristics indicative of fraudulent behavior, thus enabling the system to effectively prioritize high-risk alerts for transactions and undertake appropriate interventions. The application of XGBoost in enhancing security measures within mobile money services by identifying and mitigating SMS-based fraud risks has a number of benefits, including high accuracy, scalability to large datasets, and interpretability through feature importance analysis.

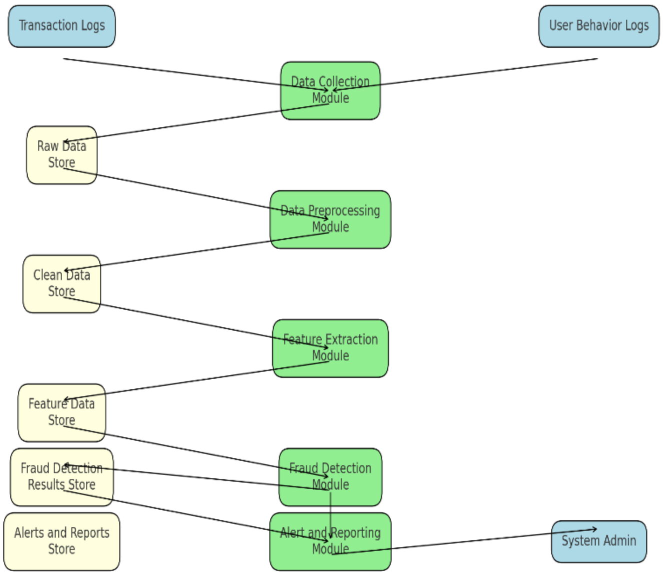
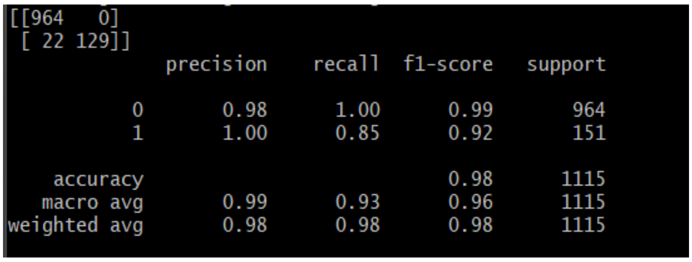


Figure 2: System overview

Results

Performance Evaluation

Performance evaluation of the fraud detection model shows that the model achieved good efficiency levels and was effective in predicting fraudulent SMS messages. The confusion matrix, shown below, contains 964 true negatives and 129 true positives, zero false positives, and 22 false negatives. On class 0, this model yields an accuracy of 98% and perfect recall at 100%, thus it is very good at correctly classifying genuine transactions without mistakenly flagging them as fraudulent. On class 1, the fraudulent transactions, it gives a perfect precision of 100%, which means that all the instances that are predicted as fraud cases are actual fraud, though the recall stands a bit lower at 85%, thus indicating that there are a few miss cases of fraud. The F1-scores are 0.99 for non-fraudulent transactions and 0.92 for fraudulent transactions, reflecting that the model is balanced with regard to both precision and recall. Finally, it provides very robust and reliable performance with regard to accuracy, in the range of 98%, making the model very effective at detection of fraudulent messages.

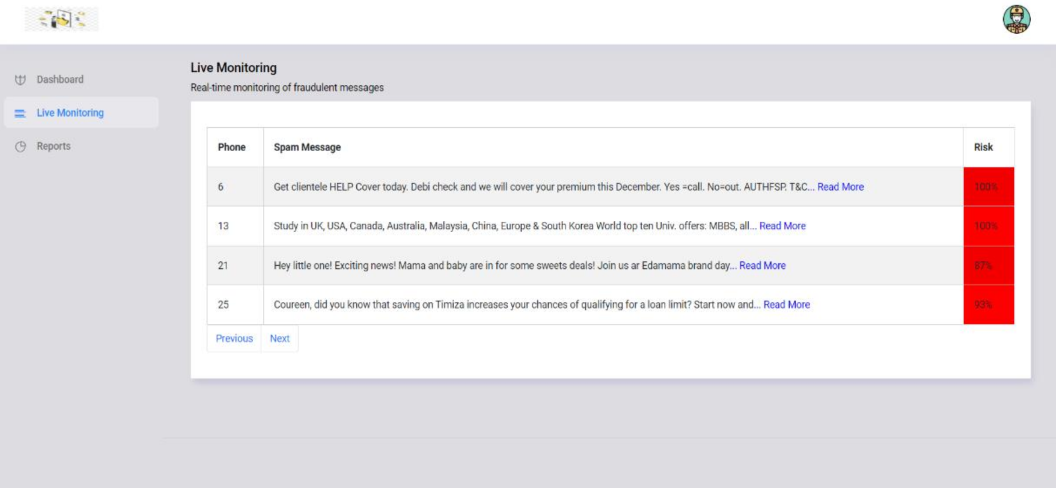


```
[[964   0]
 [ 22 129]]
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       964
           1       1.00      0.85      0.92       151

    accuracy                           0.98      1115
   macro avg       0.99      0.93      0.96      1115
weighted avg       0.98      0.98      0.98      1115
```

Figure 3: Performance Metrics

Real-Time Monitoring and Reporting

The Flask web application made access easy, providing an intuitive interface that offers real-time monitoring of SMS traffic to underpin efficient fraud detection measures. Messages that were identified with a high probability of fraud, more than 75% were flagged and abovementioned, flashing on the administrators' consoles to signal immediate reviews and actions. The systems automatically blacklisted the senders of the fraudulent messages.



Live Monitoring

Real-time monitoring of fraudulent messages

Phone	Spam Message	Risk
6	Get clientele HELP Cover today. Debi check and we will cover your premium this December. Yes =call. No=out. AUTHFSP. T&C... Read More	100%
13	Study in UK, USA, Canada, Australia, Malaysia, China, Europe & South Korea World top ten Univ. offers: MBBS, all... Read More	100%
21	Hey little one! Exciting news! Mama and baby are in for some sweets deals! Join us ar Edamama brand day... Read More	87%
25	Coureen, did you know that saving on Timiza increases your chances of qualifying for a loan limit? Start now and... Read More	93%

Figure 4: A web-based dashboard for fraud detection monitoring

A response system was automated to protect users and reduce potential risk. SMS notifications were sent to messages by users that were flagged, warning them that their accounts had been temporarily blocked on grounds of suspected fraudulent activities. Such proactive action offers users protection from possible financial loss. The system also generated detailed statistical reports regarding spam detection rates and message classification, thus providing administrators with valuable insight into decision making for the continuous optimization of the system.

Discussion

In this study, a Flask was integrated with machine learning to build a robust fraud detection system suitable for SMS messages. This constitutes an important milestone in mitigating mobile money fraud due to fraudulent messages. This kind of a system would answer a very critical need in the telecommunication and financial sectors where fraudulent SMS implies huge financial losses and loss of trust by users toward the mobile money service. This system, enabled with machine learning algorithms, does an exemplary job of real-time analytics of SMS data to identify within a second those patterns that are of a suspicious nature and can point toward fraudulent activities.

Performance metrics underpin the effectiveness of the system through high accuracy rates and real-time monitoring to ensure timely detection and mitigation of potential threats that fraudulent messages may pose. These capabilities are very critical in ensuring that users are safeguarded from unauthorized transactions, identity theft, and other types of mobile money fraud facilitated through deceptive SMS communications.

Limitations and Future Research

There are important limitations encountered in the current system that suggest several areas for future research. First, there need to be region-specific datasets to help in capturing local patterns and behavior of mobile money fraud. In this regard, there is a need for region-specific datasets, most especially from countries like Malawi, where mobile money fraud patterns may well vary. Without such data, this will lower the potential of such a system to recognize and respond effectively to fraud tactics that are more localized. This therefore underpins the need for collaborative efforts towards the gathering and curation of relevant datasets.

Moreover, the challenge of model interpretability is still high. Machine learning models for fraud detection often act as black boxes that obscure the type of reasoning used in assigning specific classifications. This calls for greater model interpretability whereby greater trust by stakeholders will be achieved, so domain experts could understand how fraudulent messages are recognized and classified within the system. More importantly, this will help in compliance with regulations when demonstrating the reliability to be trusted in the efficient detection and mitigation of mobile money fraud.

In view of the ever-changing nature of fraud techniques, the ability of the system to adapt over time is very important. Future studies should therefore be based on the improvement of machine learning techniques, such as ensemble methods or deep learning architectures, so that the system is better able to detect and mitigate emergent forms of mobile money fraud. Also, hybrid approaches involving machine learning combined with either rule-based systems or behavior analytics may provide even more resilience in the face of such sophisticated fraud tactics.

Another such area is that of scalability: the bigger the amount of SMS data and the more complex fraud patterns, then the more compelling system algorithms and infrastructure optimization for large-scale deployment will become. This involves using cloud computing resources, enhancing data processing efficiency, and implementing robust solutions in data storage to sustain real-time fraud detection and response.

Furthermore, the adversarial techniques of machine learning and the strategies against the vulnerabilities of the model in the adversarial machine learning should be investigated. The system of machine learning has some adversarial attacks that are a serious threat, especially in security-critical applications like mobile money fraud detection. Development of countermeasures for fighting such manipulations by adversaries and ensuring system resiliency will become a key factor for maintaining system effectiveness and reliability while protecting mobile money transactions.

Conclusion

In conclusion, this study has provided the practicable implementation of machine learning-driven fraud detection in SMS messages using machine learning for mitigation against mobile money fraud. The system that emerges from this is a significant step toward enhancing safety measures with respect to mobile communications by the provision of real-time monitoring and very high accuracy in identifying fraudulent activities facilitated through deceptive SMS messages.

The area of mobile money fraud detection will continue to improve in the future in tandem with advancements in machine learning, data analytics, and cybersecurity. Future researchers should work on surmounting the various limitations that have been identified, such as data availability, interpretability of models, scalability, and resilience against adversaries. The relevance of addressing these challenges through collaborative research is in innovation and implementation, which can put in place robust fraud detection solutions to effectively protect mobile money transactions and uphold trust in the use of digital financial services.

Bibliography

1. Agrawal, D., & Sastry, V. (2017). *Data mining techniques for fraud detection*. Springer.
2. Chen, X., Li, Y., & Zhao, J. (2018). *SMS fraud detection using machine learning techniques. Journal of Information Security*, 9(4), 123-135.
3. CGAP. (2020*). Phishing scams in mobile money: Protecting customers from fraud. Consultative Group to Assist the Poor*. Retrieved from https://www.cgap.org/research/publication/phishing-scams-mobile-money
4. Foster, D. P., & Stine, R. A. (2019*). Fraud detection: Techniques and strategies*. Wiley.
5. Gupta, P., & Kapoor, R. (2019). *Semantic analysis and clustering for phishing SMS detection. International Journal of Computer Science and Network Security,* 19(7), 42-50.
6. Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Elsevier.
7. ITU. (2019). *Combating mobile money fraud in developing countries. International Telecommunication Union*. Retrieved from https://www.itu.int/en/ITU-D/Financial-Inclusion/Pages/default.aspx
8. Li, H., & Wang, S. (2020). *Deep learning for SMS fraud detection: Using convolutional neural networks to identify fraudulent messages. IEEE Transactions on Information Forensics and Security*, 15, 2042-2055.
9. Murphy, K. P. (2012). *Machine learning*: A probabilistic perspective. MIT Press.
10. Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.
11. Sokol, D. (2014). *Fraud analytics: Strategies and methods for detection and prevention*. Wiley.
12. Vapnik, V. (1998). *Statistical learning theory*. Wiley.
13. World Bank. (2021). *Financial inclusion and the impact of mobile money in emerging economies. World Bank Group*. Retrieved from https://www.worldbank.org/en/topic/financialinclusion/overview