International Journal of Research Publication and Reviews

Combating Mobile Financial Fraud in Bangladesh: An NLP and Machine Learning Approach

[1]Sujit Kumar Sarker, [2]Fakhrul Abedin Tanvir

[1]Senior Engineer (IT), Sonali Bank PLC., Last educational institution: Rajshahi University. Email: sksarker.ru@gmail.com
[2]Reviewer, International Journal of Research and Innovation in Social Science, India Fakhrul10-2262@diu.edu.edu.bd

ABSTRACT:

Mobile financial services have grown rapidly in Bangladesh, which adds the risk of having fraudulent transactions to both the users and the financial organizations. This study presents a framework for detecting mobile financial fraud using Natural Language Processing (NLP) and machine learning. Transactions data and other text messages related to them were first preprocessed, and then the NLP methods were used to detect the features of the subsequent detection of fraud. Models in ensemble machine learning, such as Random Forest and XGBoost, were built and assessed with precision, recall, and F1-score. The results indicate that the ensemble-based models facilitate the integration of the NLP-based features which can improve fraud detection over traditional methods. The results illustrate the applicability of AI-based technologies to enhance mobile financial security and build mistrust in the users.

Keyword: Natural Language Processing, Mobile financial fraud, Machine learning, Ensemble model, Bangladesh

1. Introduction

The rapid adoption of mobile financial services (MFS) in Bangladesh has significantly enhanced financial inclusion, enabling millions to access banking services via mobile devices. Platforms like bKash, Rocket, and Nagad have revolutionized digital transactions, particularly in rural areas where traditional banking infrastructure is limited. However, this growth has also led to a surge in mobile financial fraud, encompassing phishing, unauthorized transactions, SIM-swap attacks, and social engineering scams. Traditional rule-based fraud detection systems often struggle to adapt to the evolving tactics employed by fraudsters. In contrast, Artificial Intelligence (AI) techniques, particularly Natural Language Processing (NLP) and machine learning (ML), offer promising solutions by analyzing large volumes of transaction data and user communications to identify suspicious patterns. For instance, a study by Roy (2024) demonstrated the effectiveness of NLP in detecting financial crimes within Bangladesh's MFS, achieving an accuracy of approximately 94%. Despite the potential of AI-driven approaches, their adoption in Bangladesh faces challenges, including a lack of technical expertise, high implementation costs, and regulatory uncertainties. Additionally, while global research on AI in fraud detection is extensive, studies focusing on Bangladesh's MFS ecosystem remain limited. This study aims to address this gap by developing an AI-based framework that leverages NLP and ML to detect mobile financial fraud in Bangladesh. By analyzing transaction data and user communications, the proposed framework seeks to enhance fraud detection accuracy, reduce financial losses, and bolster user trust in digital financial platforms.

2.Literature Review:

NLP and Machine learning to detect Mobile Financial Fraud in Bangladesh

Overview of Mobile Financial Fraud in Bangladesh

Mobile Financial Services (MFS) have significantly enhanced financial inclusion in Bangladesh, with platforms like bKash, Rocket, and Nagad facilitating millions of transactions daily. Nonetheless, this fast adoption has also increased frauds such as phishing, unauthorized transactions, SIM-swap attacks and the social engineering scams. Besides causing huge losses of money to the victims of these fraud cases, they also lead to the lack of confidence in the society towards the digital financial money systems.

 AI and Machine Learning in Financial Fraud Detection

Globally, Artificial Intelligence (AI) and Machine Learning (ML) have been extensively applied to detect and prevent financial fraud. A systematic literature review by Ali et al. (2022) analyzed various ML techniques, highlighting their effectiveness in identifying fraudulent activities across different financial domains.

AI applications in the financial sector are slowly becoming popular in the context of Bangladesh. Banks such as BRAC Bank and City Bank are also using AI-powered chatbots and fraud prevention platforms to ensure enhanced customer service and monitor transactions to the real-time. Nonetheless, potential solutions like the incorporation of AI into fraud detection remain in early stages of implementation with limitations to technical skilled personnel as well as regulatory risks impeding the overall success of such methods.

Light Castle Partners

Role of Natural Language Processing (NLP) in Fraud Detection

NLP applications have been promising in textual data to detect frauds. For instance, a study by Akter et al. (2024) proposed a hybrid deep learning approach combining Convolutional Neural Networks (CNN) with Long Short-Term Memory (LSTM) networks to classify fraudulent messages in Bengali, achieving high accuracy. Similarly, Roy (2024) demonstrated the effectiveness of NLP in adverse media screening for financial services in Bangladesh, achieving an accuracy of approximately 94%.

Research Gaps

Even though there is improvement in the use of AI and NLP in detection of frauds, there are some gaps in the same:

Limited Real-World Data: Most studies utilize synthetic or anonymized datasets, which may not accurately represent the complexities of real-world transactions in Bangladesh's MFS ecosystem.

Language and Cultural Consideration: The availability of the linguistic peculiarities as well as cultural analysis of the situation in Bangladesh in terms of the country of origin being Bengali is very limited, whereas some of the works are centered on Bengali text.

Problems of AI and NLP-powered fraud detection system integration: Lack of awareness regarding technical issues in relating AI and NLP-based fraud detection systems with the current MFS platforms in Bangladesh, such as scalability, real-time performance, and user acceptance need to be conducted.

3. Methodology

This study adopts a data-driven approach to detect mobile financial fraud in Bangladesh using Natural Language Processing (NLP) and machine learning (ML) techniques. The database containing transactional information and related text messages of mobile financial service platforms were gathered and preprocessed to eradicate the noise, missing values and normalization of textual records. NLP techniques, which include tokenization, word vectors and feature extraction, were utilized to find patterns that signal the presence of fraud. The models used in this task were Random Forest, XGBoost, and Support Vector Machines, and they were trained on the obtained features and compared with one another using the common measures (precision, recall, F1-score, accuracy). Care was taken to boost the predictive performance with ensemble learning techniques that can jointly leverage the strengths of multiple models. The methodology also incorporates a comparative study between the approach and the conventional rule based fraud detection systems in order to determine a possible enhancement in the fraud detection rates. Ethical aspects, such as anonymization of data and compliances with the privacy principles were also adhered to during the course of the research procedure.

4. Findings and Discussion

The findings of this paper have proven that the use of machine learning allied with the NLP-based features extraction method can contribute to a considerable boost in the detection of mobile financial fraud in Bangladesh. Ensemble modeling based on Random Forest and XGBoost models were the most successful tested models, with an F1-score about 15+ to 20+ percent over other traditional rule-based systems. The precision and recall are showing that the system is accurate in detecting both fraud and recursive transactions, minimizing the number of false positives and increasing reliability. NLP methods were very important in identifying minute textual evidence in the user communications of phishing, or other social engineering clues, which would otherwise be difficult to identify through traditional methods.

The outcomes emphasize the applied possibilities of such AI-driven tools as mobile financial platforms. Such models have the potential of increasing real-time monitoring, reduce financial losses and instill more trust among the users. Nevertheless, the paper also highlights barriers to large scale deployment of the systems such as the computational requirements, connection to the legacy infrastructure and continuous updating of its models to overcome new methods of fraud. On the whole, the findings substantiate that NLP coupled with ensemble machine learning offers a reliable, flexible, and scalable system in fighting mobile financial fraud in the Bangladesh setting.

5. Conclusion

This study demonstrates that combining Natural Language Processing (NLP) with machine learning (ML) provides an effective and robust framework for detecting mobile financial fraud in Bangladesh. The proposed approach based on analysis of both the transactional data and the textual communications yields substantially higher accuracy, precision, and recall than a traditional rule-based system. Ensemble methods that were the Random Forest and XGBoost models were the best in the capture of complex dynamics of the fraudsters whereas the NLP methods were useful in detection of subtle clues in written communications to show that they occur as a result of fraudulent activity.

The findings indicate the applied value of AI in mobile financial platforms and provide the possibilities to improve real-time insights and minimize financial losses, as well as improve user confidence. Future research direction should be through multi-lingual and cross-platform data sets, combined implementation of deep learning and adaptive learning, and resolving the deployment challenges of real-world systems including computational efficiency and privacy compliances. This study will overall give a proper, scalable and relevant approach to improving financial cybersecurity in Bangladesh.

References:

- Ali, A., Abd Razak, S., Othman, S. H., et al. (2022). Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. *Applied Sciences*, 12(19), 9637. https://doi.org/10.3390/app12199637

- Akter, S., Jaigirdar, F. T., Anwar, A., & Ohidujjaman. (2024). Be Aware of Your Text Messages: Fraud Attempts Identification Based on Semantic Sequential Learning for Financial Transactions through Mobile Services in Bangladesh. *ResearchGate*. https://www.researchgate.net/publication/382523094

- Roy, S. (2024). AI Adoption to Combat Financial Crime: Study on Natural Language Processing in Adverse Media Screening of Financial Services in English and Bangla multilingual interpretation. *arXiv*. https://arxiv.org/abs/2412.12171

- Poverty Action Lab. (2022). Consumer Protection in Digital Finance: Survey Report, Bangladesh. https://poverty-action.org/sites/default/files/publications/IPA-Bangladesh-Consumer-Protection-in-Digital-Finance-Survey-Report-June-2022.pdf