2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

FRAUD DETECTION IN BANK TRANSACTIONS

M.Praveen[1], Kommu Yasasvini[2], Konda Nandini[3], Kukkudhu Ananya[4]

[1]Assistant Professor, School of CSE,Malla Reddy Engineering College For Women (UGC-Autonomous), Maisammaguda,Dhulapally,Secunderabad,Telangana-500100

[234]UG Student, School of CSE,Malla Reddy Engineering College for Women, (UGC-Autonomous), Maisammaguda,Dhulapally,Secunderabad,Telangana-500100

ABSTRACT

Fraud detection in banking systems is crucial to reduce financial loss and protect the reputation of the institution. Annual losses because of financial fraud indicate that there is a need for early and effective detection mechanisms. The paper discusses a machine learning approach to improve fraud detection to prevent financial fraud, especially speeding up the check verification process to detect forgery and thus avoid further damages. Multiple intelligent algorithms are trained and tested on a public dataset, resampled to reduce class imbalance, which is a very common problem in fraud detection scenarios. Balanced data ensures that the resampling will improve the accuracy of detecting fraudulent activities. Experimental results show that the proposed machine learning techniques significantly improve prediction accuracy for banks to proactively mitigate fraud risks. This paper demonstrates how artificial intelligence is used in improving detection capability, transaction security, and speeding up response to suspicious transactions within the banking industry.

Keywords-Fraud Detection; Banking System Security; Machine Learning; Check Verification; Forgery Detection; Financial Fraud.

I. INTRODUCTION

The banking industry is experiencing a profound transformation that is driven by rapid advancements in financial technology. Traditional banking models, once reliant on manual processes and legacy systems, are evolving into dynamic, technology-enabled ecosystems. This shift is necessitated by changing customer expectations, competitive pressures, and the growing complexity of financial services. Emerging technologies such as blockchain, artificial intelligence (AI), big data analytics, digital payment systems, peer-to-peer lending platforms, crowdfunding, and robo-advisors are transforming the way financial services are delivered and consumed.The rationale for this technological revolution is multifaceted. Customers want personalized, efficient, and secure banking experiences, while banks are trying to optimize their operations and stay competitive. FinTech solutions bridge the gap between traditional banking systems and customer expectations by offering innovative, user-friendly, and cost-effective alternatives. For example, robo-advisors provide personalized financial

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

guidance while peer-to-peer lending offers flexible loan options that are outside the conventional bank services. These innovations will not only provide high customer satisfaction but also facilitate a robust and scalable framework for the banks.Despite these, challenges exist. Financial crises have always inhibited innovation adoption and relegated technological integration to lower priorities. The gaps between traditional banking services and seamless and convenient services expected today remain unbridged. All these require a paradigm shift wherein banks begin to embrace FinTech as an ally rather than a competitor that is integrated into their present operation.

This project analyzes the potential of financial technologies in changing banking services. It aims to identify ways in which customer experience, operational efficiency, and trust in digital banking systems can be enhanced by analyzing the most critical technological advancements integrated into banking operations. Findings have pointed out how FinTech bridges the gap between traditional banking systems and the demand of modern customers, hence making the banking industry ready for a future full of innovation and growth.

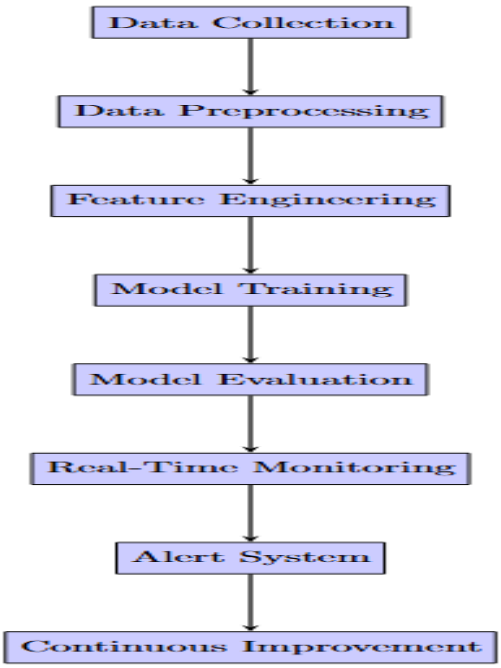

Fig 1: System Architecture

II. RELATED WORK

1. Data Mining Techniques towards Fraud Detection

Rambola et al. (2018) proposed research focused on data mining techniques for fraud detection in the banking industry. It discussed how classification models and anomalies detection models can be applied to uncover fraudulent patterns in transactional data. This approach emphasizes preprocessing techniques such as cleaning, normalization, and encoding categorical variables. Other common problems that have been associated with fraud detection, such as class imbalance, have been addressed by applying resampling techniques like SMOTE, which helps in giving more improved performance for the models developed. The authors conclude by noting the considerable reduction of false positives and false negatives as results of these techniques integration that

will make the paper an excellent source for improvement of the fraud detection system.

2. K-Nearest Neighbors and Outlier Detection

Malini and Pushpa (2017) discussed a fusion approach of K-Nearest Neighbors and outlier detection-based techniques to identify credit card fraud. The study showed that a KNN algorithm could classify the transactions under two major groups as fraudulent and nonfraudulent based on received transactional data, while the additional advantage provides for this algorithm with efficient handling of large datasets. The study also evaluates the use of outlier detection algorithms in identifying abnormal transaction patterns that were inconsistent with the expected behaviors. Their research emphasized that the basis of applying KNN with unsupervised anomaly detection techniques was useful in scenarios where labeled data was scarce. As such, it lends quite a value to the field of fraud detection.

3. Ensemble Learning for Fraud Detection

Sohony et al. (2018) proposed ensemble learning approaches to credit card fraud detection, using multiple classifiers combined to improve the overall accuracy of predictions. They explored algorithms like Random Forest and Gradient Boosting, which use multiple decision trees to make predictions. The results showed that ensemble learning methods are particularly effective for fraud detection, as they can reduce overfitting and improve generalization by learning from diverse patterns in the data. This method was proved to be more robust when working with highly imbalanced datasets; hence, the ensemble models are good for real-time fraud detection systems in the banking sector.

4. Optimization Algorithms in Fraud Detection

Wang et al. (2018) used optimization methods with neural networks for credit card fraud detection. In particular, they used WOA for BP-NN optimization. Their work showed that inclusion of the WOA in the BP-NN model resulted in a massive superiority in detection compared to traditional neural networks. The whale algorithm optimized the weights and biases of the BP-NN, which could converge better and even accelerate the speed of training. This work demonstrated that optimized deep learning algorithms could significantly enhance the performance of models in cases where fraud detection was to be conducted based on extremely complex and biased financial transaction datasets.

5. Genetic Algorithms for Imbalanced Data

Benchaji et al. (2018) used Genetic Algorithms to enhance the classification result of imbalanced datasets credit card fraud detection task. They have proposed an innovative approach where GA adjusts the decision thresholds of the classification models so that the detection models can capture the relatively tiny minority fraudulent class more sensitively. Over generations of evolutions, the model can better capture the small number of fraudulent transactions. The technique found proved successful in improving the recall of the fraud detection model so as to identify more of the fraudulent transactions without allowing false positives to increase significantly. This is very relevant in fraud

detection in banking, where the fraudulent class often forms a minority compared to the legitimate class.

III. IMPLEMENTATION

The fraud detection system for banking transactions utilizes machine learning algorithms to detect and prevent fraud. The system begins with a collection of transactional data, which include features such as amount, time, location, and user behavior. In preprocessing the data, one handles missing values, normalizes numerical features, and encodes categorical variables. Techniques like SMOTE (Synthetic Minority Oversampling) are used to balance the data so that the model gets trained on balanced data to address the class imbalance problem. New features like transaction velocity, location variance, and anomalies in transaction amounts help to identify patterns of fraudulent behavior through feature engineering. This uses various machine learning models like Logistic Regression, Random Forest, and Gradient Boosting trained on labeled datasets as well as anomaly detection methods such as Isolation Forest and Autoencoders in case of unlabeled data. Models are tested by metrics such as Precision, Recall, and ROC-AUC to ensure the ability of such models to correctly identify fraudulent transactions. To perform real-time monitoring, it integrates with the banking system using Flask APIs. Streaming data is also utilized in Apache Kafka, allowing alerts to be generated once suspicious transactions are detected. With such integration, flagged transactions get checked as soon as possible and also alerts both banks and customers involved in that transaction. In summary, the developed system uses machine learning techniques, real-time monitoring, and continuous updates to provide a robust solution to counter financial fraud in the banking industry.

IV. ALGORITHM

The proposed algorithm for fraud detection in banking transactions follows a structured approach using machine learning techniques to identify and prevent fraudulent activities.

Step 1: Data Collection and Preprocessing

Transactional data, including features such as amount, time, location, and user behavior, is collected from banking systems or public datasets. Data preprocessing includes handling missing values, removing outliers, normalizing numerical features (e.g., transaction amounts), and encoding categorical variables (e.g., location, transaction type). Class imbalance is addressed using techniques such as SMOTE (Synthetic Minority Oversampling) or undersampling to ensure a balanced dataset for training.

Step 2: Feature Engineering

New features are developed in fraud detection, including transaction velocity (transactions per unit of time), location variance (change in geographical patterns), and anomalies in the amount of transactions (amounts that vary from the average). Incorporation of historical fraud data to track recidivist fraudsters and observe patterns.

Step 3: Model Training

The dataset was split into training and testing sets. Supervised models like Logistic Regression, Random Forest, and Gradient Boosting were trained on labeled data, and unsupervised models such as Isolation Forest

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

and Autoencoders were used for anomaly detection to identify fraudulent transactions in labeled and unlabeled data.

Step 4: Model Evaluation

The models are evaluated with Precision, Recall, and ROC-AUC metrics. Precision is the fraction of correct fraud predictions, Recall evaluates the model's capability to recognize actual fraud cases, and ROC-AUC captures the overall performance of the model. The best-performing model is chosen for deployment.

Step 5: Real-Time Monitoring and Integration

The chosen model is deployed into the banking system with Flask for live predictions. Apache Kafka streams transaction data, and when suspicious transactions are detected, alerts are sent. The review team is notified, and customers and banks are informed of potential fraud.

Step 6: Continuous Improvement

The model is updated regularly with new data to keep abreast of emerging fraud patterns, ensuring ongoing improvement in fraud detection.

V. RESULT

Fig 1 : Login Service Provider

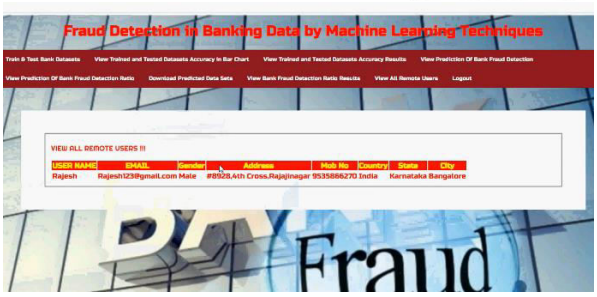

Fig 2 : View All Remote Users

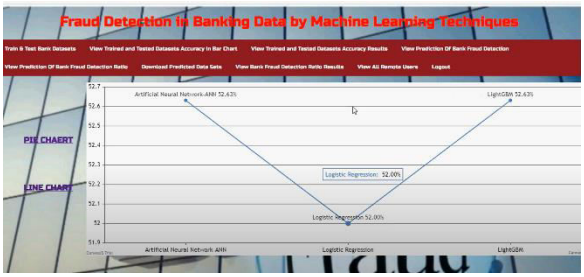

Fig 3 : Line Chart

Fig 4 : View Bank Fraud Prediction Type Details

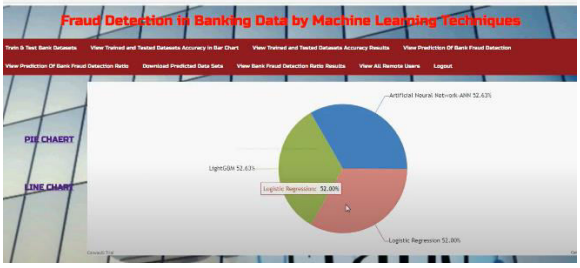

Fig 5 : Pie Chart

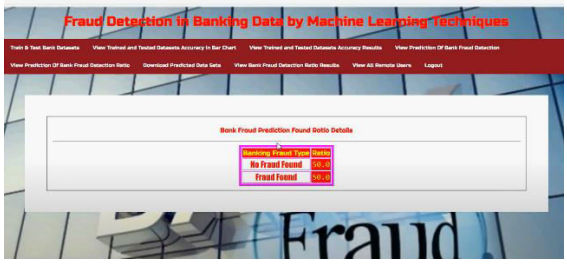

Fig 6 : Bank Fraud Ratio Details

VI. CONCLUSION

In conclusion, the proposed machine learning-based fraud detection system offers a holistic and effective solution for identifying and preventing fraudulent activities in banking transactions. By leveraging advanced techniques such as data preprocessing, feature engineering, and model training with both supervised and unsupervised algorithms, the system can accurately detect fraudulent patterns in real-time. It incorporates models such as Logistic Regression, Random Forest, Gradient Boosting, and anomaly detection techniques to ensure robust detection, even in imbalanced datasets. Long-term effectiveness is also provided through continuous monitoring and adaptation to new fraud patterns through regular updates. The real-time monitoring aspect of the system, using tools like Flask and Apache Kafka, is used to give immediate responses to suspicious activities and minimize the potential losses as well as enhance the transaction security. Furthermore, it uses historical fraud data, precision, recall, and ROC-AUC for model evaluation, hence making accuracy and reliability prime in fraud detection.

In short, this system is able to not only minimize financial risks to the banks but also increase the level of customer trust for the services in digital banking. This offers a scalable and efficient framework for the future of fraud prevention in the emerging financial technology landscape.

REFERENCES

[1]Fabrizio Carcillo, Andrea Dal Pozzolo, Yann-A¨el Le Borgne, Olivier Caelen, Yannis Mazzer, and Gianluca Bontempi. Scarff: a scalable framework for streaming credit card fraud detection with spark. Information Fusion, 41:182–194, 2018.

[2]Ma, T.; Qian, S.; Cao, J.; Xue, G.; Yu, J.; Zhu, Y.; Li, M. An Unsupervised Incremental Virtual Learning Method for Financial Fraud Detection. In Proceedings of the 2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA), Abu Dhabi, United Arab Emirates, 3–7 November 2019; pp. 1–6. Financial fraud detection.

[3]Galina Baader and Helmut Krcmar. Reducing false positives in fraud detection: Combining the red flag approach with process mining. International Journal of Accounting Information Systems, 2018.

[4]Ravisankar P, Ravi V, Raghava Rao G, and Bose, Detection of financial statement fraud and feature selection using data mining techniques, Elsevier, Decision Support Systems Volume 50, Issue 2, p491-500 (2011) SVM

[5] K. Seeja, and M. Zareapoor, “FraudMiner: A Novel Credit Card Fraud Detection Model Based on Frequent Itemset Mining,” The Scientific World Journal, 2014, pp. 1-10. KNN, SVM

[6] C. Tyagi, P. Parwekar, P. Singh, and K. Natla, “Analysis of Credit Card Fraud Detection Techniques,” Solid State Technology, vol. 63, no. 6, 2020, pp. 18057-18069. Credit card faud

[7] C. Chee, J. Jaafar, I. Aziz, M. Hassan, and W. Yeoh, "Algorithms for frequent itemset mining: a literature review," Artificial Intelligence Review, vol. 52, 2019, pp. 2603–2621. Litrature review AI

[8] S. Kiran, J. Guru, R. Kumar, N. Kumar, D. Katariya, and M. Sharma, "Credit card fraud detection using Naïve Bayes model based and KNN classifier," International Jounral of Advance Research, Ideas and Innovations in Technology, vol. 4, 2018, pp. 44-47. KNN Naïve Byers

[9] Pumsirirat, A.; Yan, L. Credit Card Fraud Detection Using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine. Available online: https://thesai.org/Downloads/Volume9No1/Paper_3-Credit_Card_Fraud_Detection_Using_Deep_Learning.pdf (accessed on 23 February 2021). DL

[10] PwC's Global Economic Crime and Fraud Survey 2020. Available online: https://www.pwc.com/fraudsurvey (accessed on 30 November 2020). Fraud surver.

[11] Pourhabibi, T.; Ongb, K.L.; Kama, B.H.; Boo, Y.L. Fraud detection: A systematic literature review of graph-based anomaly detection approaches. Decis. Support Syst. 2020, 133, 113303. Fraud detection.

[12] Lucas, Y.; Jurgovsky, J. Credit card fraud detection using machine learning: A survey. arXiv 2020, arXiv:2010.06479. Credit card fraud.

[13] Podgorelec, B.; Turkanovi´c, M.; Karakati˘c, S. A Machine LearningBased Method for Automated Blockchain Transaction Signing Including Personalized Anomaly Detection. Sensors 2020, 20, 147. Anomaly detection.

[14] Synthetic Financial Datasets for Fraud Detection. Available online: https://www.kaggle.com/ntnu-testimon/paysim1 (accessed on 30 November 2020). Fraud detection.

[15] C. Wang, Y. Wang, Z. Ye, L. Yan, W. Cai, and S. Pan, "Credit Card Fraud Detection Based on Whale Algorithm Optimized BP Neural Network," 2018 13th International Conference on Computer Science Education (ICCSE), Colombo, 2018, pp. 1-4, doi: 10.1109/ICCSE.2018.8468855

[16] I. Benchaji, S. Douzi and B. ElOuahidi, "Using Genetic Algorithm to Improve Classification of Imbalanced Datasets for Credit Card Fraud Detection," 2018 2nd Cyber Security in Networking Conference (CSNet), Paris, 2018, pp. 1-5, doi: 10.1109/CSNET.2018.8602972.

[17] John O. Awoyemi, Adebayo Olusola Adetunmbi, and Samuel Adebayo Oluwadare. Credit card fraud detection using machine learning techniques: A comparative analysis. 2017 International Conference on Computing Networking and Informatics (ICCNI), pages 1–9, 2017.

[18] Fabrizio Carcillo, Andrea Dal Pozzolo, Yann-A¨el Le Borgne, Olivier Caelen, Yannis Mazzer, and Gianluca Bontempi. Scarff: a scalable framework for streaming credit card fraud detection with spark. Information Fusion, 41:182–194, 2018.