DETECTING FRAUDULENT BANKING TRANSACTIONS USING MACHINE LEARNING

Dr. G.N.V. Vibhav Reddy[1],C Ramesh[2],G Srinivas[3],Sai Vijay Vardhan[4],G.Satwik Reddy[5],Mohd Shafiuddin Siddiqi[6]
Assosciate Professor, Department of Computer Science and Engineering[1]
Student, Department of Computer Science and Engineering[2,3,4,5,6]
Sree Dattha Institute of Engineering and Science, Sheriguda, Telangana. [1,2,3,4,5,6]

ABSTRACT

Vulnerabilities in the banking system have made us open to fraudulent actions, which not only hurt customers but also the bank's reputation and bottom line. Financial fraud in banks is responsible for the loss of an estimated significant quantity of money annually. The fraud may be reduced with the help of early identification, which enables the creation of a countermeasure and the restoration of such losses. To help with fraud detection, this study suggests a machine learning-based approach. The technology, which is based on artificial intelligence (AI), will speed up the check verification process to fight counterfeits and reduce damage. In this post, we looked at a bunch of smart algorithms that were trained on a public dataset to find out which factors are associated with fraud. The dataset utilised for this investigation is resampled to lower the high class of imbalance. For more precise results, the data is analysed using the proposed method.

Index Terms: Financial Fraud, Banking System Vulnerabilities, Machine Learning, Artificial Intelligence, Fraud Detection, Class Imbalance, Dataset Resampling.

DOI Number: 10.48047/nq.2024.22.4.nq24014 **NeuroQuantology 2024; 22(4):121-127**

1. INTRODUCTION

When compared to modern banks, those in the future will have much more advanced capabilities. Infrastructural, service, human, and skill set changes are to blame for these shifts. The introduction of new financial technology in the banking industry is solely responsible for this shift. The majority of financial institutions can easily incorporate new technology into their service delivery models, which allows us to shape the banking industry to our liking. Digital payment processing, blockchain [1], AI, big data, crowdsourcing, P2P lending, and robot advisers are all crucial to the delivery of financial services. When it comes to banking, why are all these technical advances necessary? When new technologies emerge, the banking industry is quick to incorporate them into their operations to better serve their customers. However, new initiatives in the sector have often been hampered by financial crises, making innovation a low priority. Meanwhile, a plethora of innovative technologies have emerged as potential game-changers in the quest to revamp the traditional banking system into one that prioritises the needs of its customers. There was still a disconnect between the customer's experience and the bank's offerings from their convenience and experience standpoint. In order to enhance the client experience via the use of AI technology, FinTech businesses assist various

financial operations, as shown in Figure (1) [2]. Lots of researchers looked into this void. As a result of clients' faith and confidence in new technologies, the conventional banking sector is also experiencing changes due to technological progress.

There are now hundreds of new financial technology businesses providing goods and services to banks, which helps to supplement this and improve technical assistance; There are a number of consumer-friendly options accessible to them via robo-advising platforms and peer-to-peer lending, both of which provide alternatives to loans offered by traditional banks. In addition to being affordable, these services are also extremely noticeable. They provide a graphical user interface (GUI) that is easy for customers to use, while traditional banks handle back-end procedures including consolidation, frequent reporting, post-dated settlement, and more. Because of this shift, the conventional banking function will remain a commodity utility supplier in the future, altering the future banking model. The front end and technical front end manage the consumer experience. There are a number of additional encouraging trends in the relevant industry sector that are associated with this banking technology breakthrough.

2.LITERATURE SURVEY

Adhakrishna Rambola, Prateek Varshney, and Prashant Vishwakarma proposed that the banking sector holds significant value in our daily lives. Every individual interacts with the banking sector in two primary ways: physically and online. Physical fraud can occur through activities such as stealing credit cards or sharing bank account details with corrupt bank employees. Online fraud involves sharing card details over the internet or phone with unauthorized persons, including spamming and phishing attempts. Such fraudulent activities pose risks to both customers and banks during transactions and adherence to bank policies.

- N. Malini and M. Pushpa highlighted that credit cards are a popular payment mode accepted both offline and online, facilitating cashless transactions. However, with technological advancements, credit card fraud has also increased significantly. Economic fraud is a global concern, resulting in substantial financial losses annually. These fraudulent activities mimic genuine transactions, necessitating efficient fraud detection methods to minimize chaos and ensure order in banking operations. Techniques such as machine learning, genetic programming, fuzzy logic, sequence alignment, K-nearest neighbor (KNN) algorithms, and outlier detection are employed to optimize fraud detection systems, reducing false alarm rates while increasing detection accuracy.
- Wencheng Cai, Shang Pan, and Lingyu Yan proposed a credit card fraud detection technology based on the Whale Optimization Algorithm (WOA) optimized Backpropagation (BP) neural network. This approach aims to address issues such as slow convergence, local optima, network deficiencies, and poor system stability inherent in BP neural networks. By utilizing WOA to optimize BP network weights, they enhance the accuracy of fraud detection.
- Ibtissam Benchaji, Samira Douzi, and Bouabid Ouahidi emphasized the escalating financial fraud associated with increased credit card usage, necessitating robust fraud detection methods to mitigate losses. They highlighted the challenge of imbalanced credit card fraud datasets, where fraudulent transactions are significantly fewer than legitimate ones. To tackle this, they proposed a sampling method combining K-means clustering and genetic algorithms to improve classification performance on minority class instances.
- Seeja K.R and Masoumeh Zareapoor introduced an intelligent credit card fraud detection model designed for highly

imbalanced and anonymous transaction datasets. Their model addresses the class imbalance problem using frequent itemset mining to identify patterns in legal and fraudulent transactions for each customer. A matching algorithm then determines which pattern (legal or fraudulent) best fits incoming transactions, achieving high fraud detection rates and balanced classification performance.

3.PROBLEM STATEMENT

The existing system for "Fraud Detection in Banking Transactions Using Machine Learning" incorporates a comprehensive approach to mitigating financial fraud within the banking sector. Initially, historical transaction data is collected, encompassing a diverse range of transactions, and undergoes rigorous preprocessing to handle missing data, address imbalances, and normalize features. The exploratory data analysis phase provides critical insights into patterns and correlations. Following this, relevant features are carefully selected to contribute to the fraud detection process. The model development phase employs machine learning algorithms, with a focus on continuous optimization through hyper parameter tuning. The AI-based model is implemented within the banking system for real-time or batch processing of transactions. Evaluation metrics, including accuracy, precision, recall, and AUC-ROC, are employed to assess the model's performance. Continuous monitoring mechanisms and feedback loops are established for adaptive improvements, ensuring the model remains effective against evolving fraudulent activities. The entire process is thoroughly documented, providing insights into data sources, preprocessing steps, model development, and evaluation metrics. Furthermore, security measures are integrated to safeguard both the model and the sensitive financial data it processes, encompassing encryption, access controls, and other relevant security best practices.

DRAWBACKS:

- Thesystem requires significant computational resources and expertise for continuous optimization and maintenance.
- Handling highly imbalanced data and ensuring effective real-time processing pose considerable technical challenges.

4.PROPOSED MODEL

The proposed system for "Fraud Detection in Banking Transactions Using Machine Learning" aims to overcome the limitations of the existing system by introducing innovative strategies and technologies. To address imbalanced data issues, the proposed system employs advanced resampling techniques to mitigate biases and enhance the model's ability to detect instances of fraud across various classes. A key focus lies in the continuous evolution of the fraud detection model to adapt to emerging patterns through regular updates facilitated by a dynamic learning mechanism. To mitigate overfitting, the proposed system integrates sophisticated regularization techniques and explores ensemble methods to improve the model's generalization to unseen data. Interpretability is enhanced through the incorporation of explainable AI techniques, ensuring that stakeholders can comprehend and trust the decision-making process of the model. Additionally, the proposed system places a strong emphasis on data quality and variability, implementing robust data validation and cleansing protocols. To address computational resource constraints, optimization strategies are explored to enhance the efficiency of processing, ensuring timely and cost-effective fraud detection. The system also incorporates mechanisms to fortify resilience against adversarial attacks, leveraging advanced security measures to protect against manipulative inputs. Regulatory compliance is integrated into the core of the proposed system, ensuring adherence to legal and ethical standards. User acceptance is fostered through comprehensive training and communication strategies to instill confidence in the reliability and effectiveness of the machine learning-

based fraud detection system. Through these advancements, the proposed system aims to not only enhance the accuracy and efficiency of fraud detection but also ensure adaptability, transparency, and compliance in the ever-evolving landscape of banking transactions.

ADVANTAGES:

- The proposed system enhances fraud detection accuracy and efficiency by employing advanced resampling techniques, dynamic learning mechanisms, and robust regularization methods.
- It ensures transparency and regulatory compliance through explainable AI techniques, comprehensive data validation, and stringent security measures, fostering stakeholder trust and confidence.

5.SYSTEM MODEL

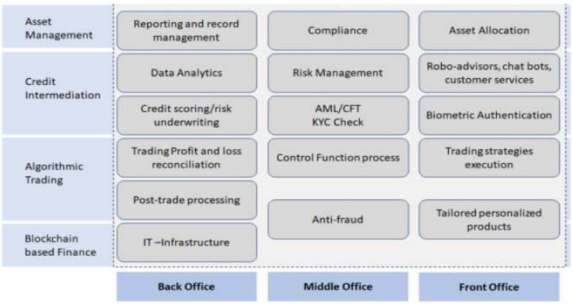


Figure.1.System Model

6.IMPLEMENTATION:

MODULES:

Data Collection and Preprocessing:

Gather relevant datasets containing banking transactions, ensuring a diverse representation of both genuine and fraudulent activities. Perform preprocessing tasks, including handling missing values, addressing outliers, and resampling to mitigate class imbalances.

Feature Engineering and Selection:

Identify and select features that are most relevant to fraud detection. This module involves analyzing the dataset to create new features or transform existing ones, enhancing the machine learning model's ability to discern patterns associated with fraudulent transactions.

Machine Learning Model Training:

Implement various machine learning algorithms, such as logistic regression, decision trees, random forest, support vector machines, or gradient boosting models. Train these models on the preprocessed dataset to learn and capture the patterns indicative of fraudulent activities.

Real-time Transaction Verification:

Develop a module for real-time transaction verification, leveraging the trained machine learning model. This module should facilitate the quick and efficient verification of transactions as they occur, ensuring timely detection and prevention of fraudulent activities.

Model Evaluation and Continuous Monitoring:

Assess the performance of the trained machine learning model using metrics like accuracy, precision, recall, and F1-score. Implement continuous monitoring mechanisms to track the model's effectiveness over time, enabling timely updates and adaptations to address emerging fraud patterns.

7.RESULTS

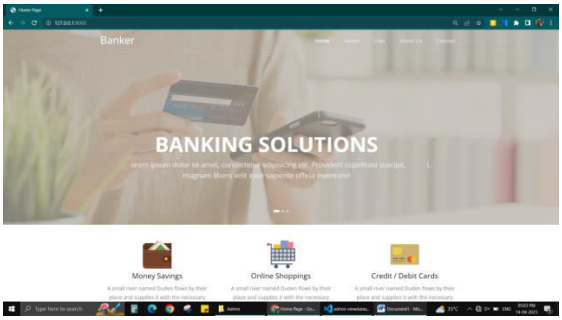


Figure.2. Home Page

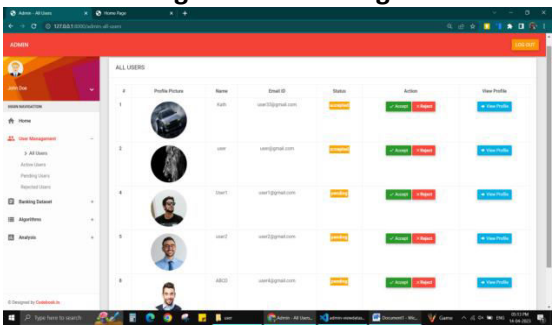


Figure.3. All users

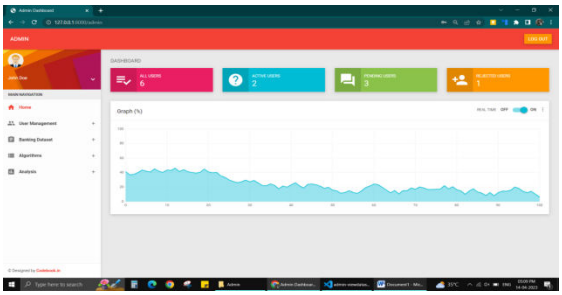


Figure.4. Admin Dash

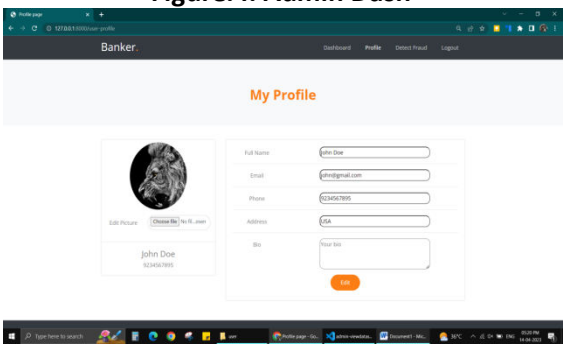


Figure.5.Profile Page

Figure.6.Active users

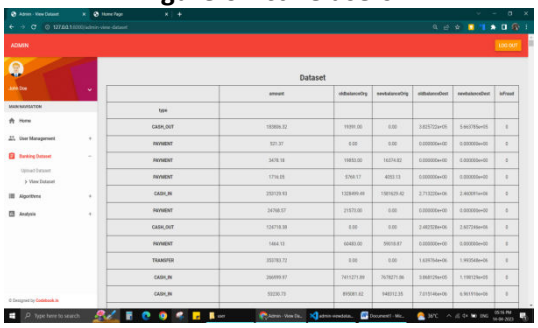


Figure7.View Data set

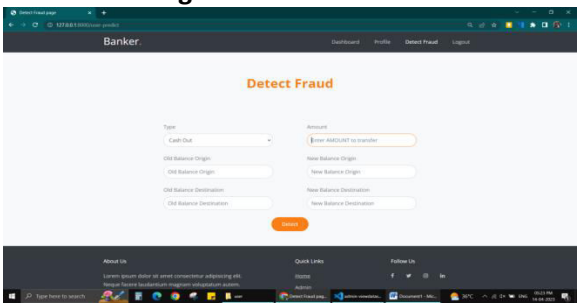


Figure.8.User Profile

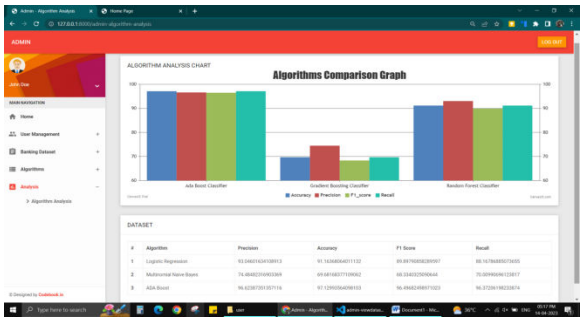
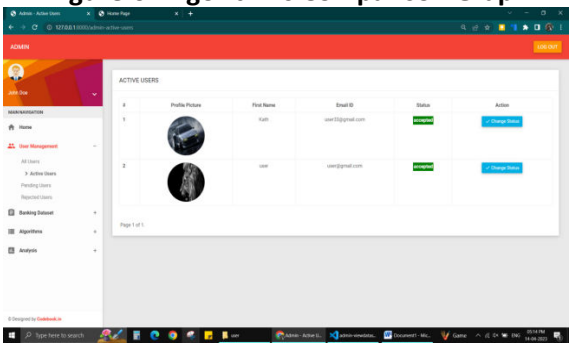


Figure.6. Algorithms Comparison Graph



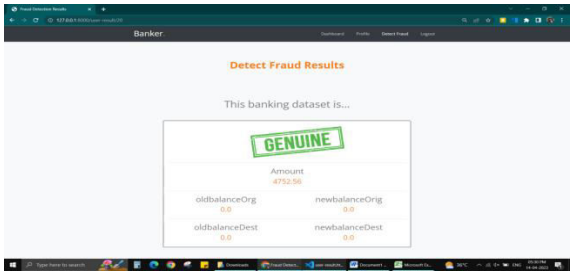


Figure.9 Genuine Result

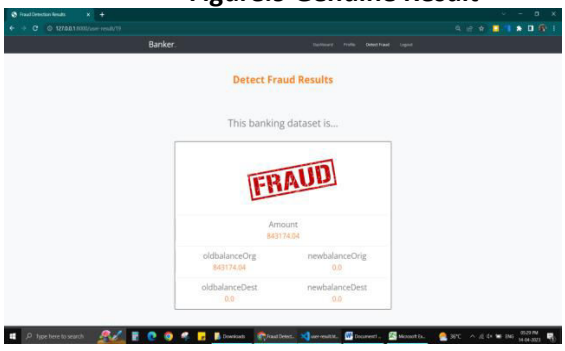


Figure.10 Fraud Result

8.CONCLUSION

The study utilized machine learning algorithms to detect fraud in banking applications using the UCI's publicly available dataset. Addressing the dataset's significant imbalance, SMOTE was employed to balance samples. Implementation challenges with KNN and Random Forest were mitigated by employing XGBoost for boosting.

The model achieved an impressive performance accuracy of 97.74%. Analysis revealed that individuals aged 19-25 years exhibit a higher likelihood of fraudulent behavior compared to other demographic group.

9.FUTURE ENHANCEMENT

Future research in banking fraud detection could focus on enhancing interpretability with Explainable AI, integrating real-time data processing for faster detection, exploring ensemble methods for improved accuracy, and analyzing demographic factors to better understand fraud patterns among different age groups.

10.REFERENCES

[1] R. Rambola, P. Varshney and P. Vishwakarma, “Data Mining Techniques for Fraud Detection in Banking Sector,” 2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2018, pp. 1-5, doi: 10.1109/CCAA.2018.8777535.

[2] N. Malini and M. Pushpa, “Analysis on credit card fraud identification techniques based on KNN and outlier detection,” 2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), Chennai, 2017, pp. 255-258, doi: 10.1109/AEEICB.2017.7972424.

[3] Ishan Sohony, Rameshwar Pratap, and Ullas Nambiar. 2018. Ensemble learning for credit card fraud detection. In Proceedings of the ACM India Joint International Conference on Data Science and Management of Data (CoDS-COMAD ’18). Association for Computing Machinery, New York, NY, USA, 289–294. DOI:https://doi.org/10.1145/3152494.3156815

[4] 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

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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

[10] 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

[11] 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

[12] 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

[13] 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

[14] 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_3Credit_Card_Fraud_Detection_Using_Deep_ Learning.pdf (accessed on 23 February 2021). DL

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

[16] 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.

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

[18] 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.

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

[20] 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.