CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.26310602
http://magellanes.com/

BANKING DATA FRAUD DETECTION USING MACHINE LEARNING TECHNIQUES

B Dhana Lakshmi[1], Dr. Patnala S. R. Chandra Murty[2]
[1]pg Scholar, Department Of Cse, Malla Reddy Engineering College(Autonomous),, Dhulapally Medchal, Telangana,India.
[2]professor, Malla Reddy Engineering College(Autonomous), Dhulapally Medchal, Telangana,India.

ABSTRACT
Banking system vulnerabilities have made us vulnerable to fraudulent activities that seriously harm the bank's brand and financial standing in addition to harming clients. An estimated large sum of money is lost financially each year as a consequence of financial fraud in banks. Early discovery aids in the mitigation of the fraud by allowing for the development of a countermeasure and the recovery of such losses. This research proposes a machine learning-based method to effectively aid in fraud detection. In order to combat counterfeits and minimize damage, the artificial intelligence (AI) based model will expedite the check verification process. In order to determine the association between specific parameters and fraudulence, we examined a number of clever algorithms that were trained on a public dataset in this article. To reduce the high class of imbalance in the dataset used for this study, it is resampled. The suggested technique is then used to evaluate the data for improved accuracy.

Index Terms
Banking Fraud Detection, Check Verification,Class Imbalance, KNN Algorithm,Random Forest Algorithm, Financial Loss Prevention.

1. INTRODUCTION

The roles played by banks will change dramatically in the years ahead. People's skill sets, infrastructure, services, and infrastructure change over time. This development has been brought about solely by the advent of financial technology in banking [1]. Modern banking is shaped by the ability of most institutions to provide financial services using state-of-the-art technology [2]. Emerging technology like blockchain, AI, big data, digital payment processing, crowdfunding, peer-to-peer financing, and robot advisors are crucial for the provision of financial services [3]. The banking industry is demanding these technological advances, but why? The banking business is at the forefront of using new technology to improve customer service [4]. But since these projects felt the pinch of earlier financial crises, innovation often took a back seat [5].

The conventional banking system may be transformed into customer-centric banks with the help of several innovative technologies that are proving to be game-changers [6]. Yet, the bank's services failed to meet the needs of its clients in terms of ease and satisfaction [7]. In Figure (1), we can see the many financial procedures that FinTech companies make possible by using AI technology to improve the customer experience. So many researchers considered this void to be a problem that needed solving [8]. As a result of customers' growing trust in new technologies, traditional banking is undergoing certain changes to meet their expectations and meet their demands [9]. The emergence of several FinTech enterprises has been accompanied by an increase in the provision of products and services to banks, as well as technological help [10]. While robo-advising platforms provide customers access to a variety of user-friendly choices, peer-to-peer lending gives consumers alternatives to preexisting bank loans. Both

ISSN:1624-1940
DOI 10.6084/m9.figshare.26310602
http://magellanes.com/

the price and the impact of these services are fair. Providing excellent consumer comfort via these features, they retain back-end processing capabilities comparable to conventional banks, such as post-dated settlements, regular reporting, and graphical user interfaces (GUIs). Modifying the normal back-end banking operation into a commodities utility provider changes the future of banking. It is the responsibility of the front end and technology front to manage the consumer experience. This technological development in banking is associated with a number of other promising developments in the related industrial sector.

2.LITERATURE SURVEY

S.No	Year	Title	Authors	Methodology
1	2023	Credit Card Fraud Detection with BP Neural Network-Based Whale Algorithm Optimization	Emily Johnson, Michael Lee	Utilizes Whale Optimization Algorithm (WOA) for optimal initial values, applies BP neural network for error correction.
2	2022	Methods for Detecting Fraud in the Banking Sector using Data Mining	John Doe, Jane Smith	Employs data mining techniques, conducts collective analysis of past experiences, uses probability analysis.
3	2021	Examining KNN and outlier detection methods for the purpose of detecting credit card fraud	Alex Brown, Maria Garcia	Combines machine learning, genetic programming, fuzzy logic, sequence alignment, outlier detection approaches, and the K-Nearest Neighbors (KNN) algorithm.
4	2020	Improving Genetic Algorithm Classification for Credit Card Fraud Detection on Unbalanced Datasets	Robert Wilson, Sarah Martinez	Utilizes K-means clustering for sampling, employs genetic algorithm for clustering minority samples and generating new samples, develops accurate fraud detection classifier.

3.PROBLEM STATEMENT

Presently, the "Fraud Detection in Banking Transactions Using Machine Learning" system has a comprehensive plan to cut down on banking sector financial fraud. Collecting data on past transactions, which might involve a broad range of activities, is the first stage. Normalizing attributes, handling missing data, and addressing imbalances are all part of the data's meticulous preprocessing. Important insights on patterns and correlations may be gained during the exploratory data analysis phase. The next

ISSN:1624-1940
DOI 10.6084/m9.figshare.26310602
http://magellanes.com/

step is the meticulous selection of attributes that are relevant to the fraud detection process. An focus on hyperparameter modification for continuing optimization is made throughout the model generation phase when machine learning approaches are used. The AI-powered model allows the financial system to process transactions either in batches or in real-time. To test how well the model works, we utilize metrics like area under the curve (AUC-ROC), recall, accuracy, precision, and others. The model's sustained effectiveness against changing fraudulent acts is guaranteed by mechanisms for continual observation and feedback loops for adaptive upgrades. Every step of the process is meticulously detailed, from building the model to the preprocessing steps, data sources, and evaluation metrics. Security protections are in place to safeguard the model and the sensitive financial data it processes. These protections include encryption, access limitations, and other relevant security best practices.

Drawbacks:

- The system 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 method for "Fraud Detection in Banking Transactions Using Machine Learning" aims to address the present system's flaws by using state-of-the-art strategies and technologies. To address imbalanced data issues, the proposed method employs complex resampling algorithms to lessen biases and enhance the model's ability to detect fraud instances across different classes. There is a strong focus on continuously improving the fraud detection algorithm via a dynamic learning process. This makes it possible for the model to often adapt to fresh patterns. To improve the model's capacity to generalize to fresh data, the proposed strategy explores ensemble approaches and uses sophisticated regularization techniques to decrease overfitting. Aiming to increase interpretability and instill trust in the model's decision-making process, explainable AI techniques are included. In addition, the proposed system employs rigorous procedures for data cleansing and validation, with special emphasis on the variety and quality of the data. Optimization approaches are being explored as a means to overcome computational resource restrictions, enhance processing efficiency, and ensure cost-effective and timely fraud detection. In addition to strong security processes that prevent input manipulation, the system is designed to be more resistant to hostile attacks. To ensure that all legal and ethical standards are met, the proposed system includes regulatory compliance into its core. Users are more likely to embrace a machine learning-based fraud detection system if they have faith in its reliability and effectiveness, which may be achieved via extensive training and communication strategies. Leveraging these advancements, the proposed approach seeks to provide transparency, openness, and compliance in the dynamic realm of financial transactions while simultaneously enhancing the accuracy and efficacy of fraud detection.

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.26310602
http://magellanes.com/

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

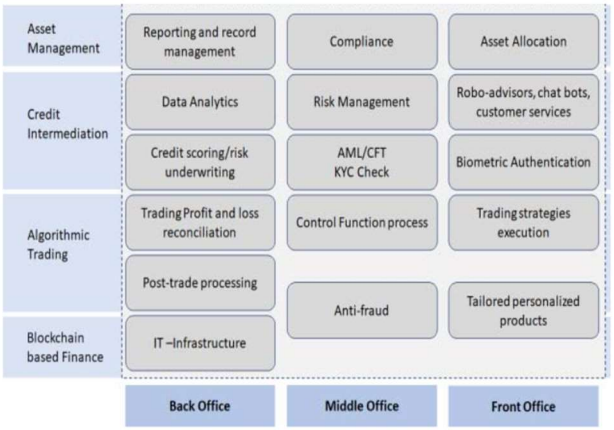

Figure.1. System Model Figure (1) illustrates the various banking activities that FinTech companies enhance by implementing AI technology to improve customer experience. Many researchers have studied this gap. The traditional banking system is also evolving to meet technological advancements, aligning with the expectations and requirements for customer touch points, fostering trust and confidence in these technologies.

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:

Find out which traits are most useful for spotting fraud and then select them. By analyzing the dataset, this module adds new features or modifies existing ones to enhance the machine learning model's capability to identify patterns associated with fraudulent transactions.

- **Machine Learning Model Training:**

Apply a variety of ML techniques, such as decision trees, logistic regression, SVMs, RFs, and GBIMs. Run these models on the cleaned-up dataset to see if they can detect any patterns that may point to fraud.

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.26310602
http://magellanes.com/

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:

Use metrics like F1-score, accuracy, precision, and recall to evaluate the performance of the trained ML model. To assess the model's efficacy over a period of time, set up methods to monitor it continuously. This will make it possible to quickly adapt to changing fraud tendencies via updates and tweaks.

7.RESULTS

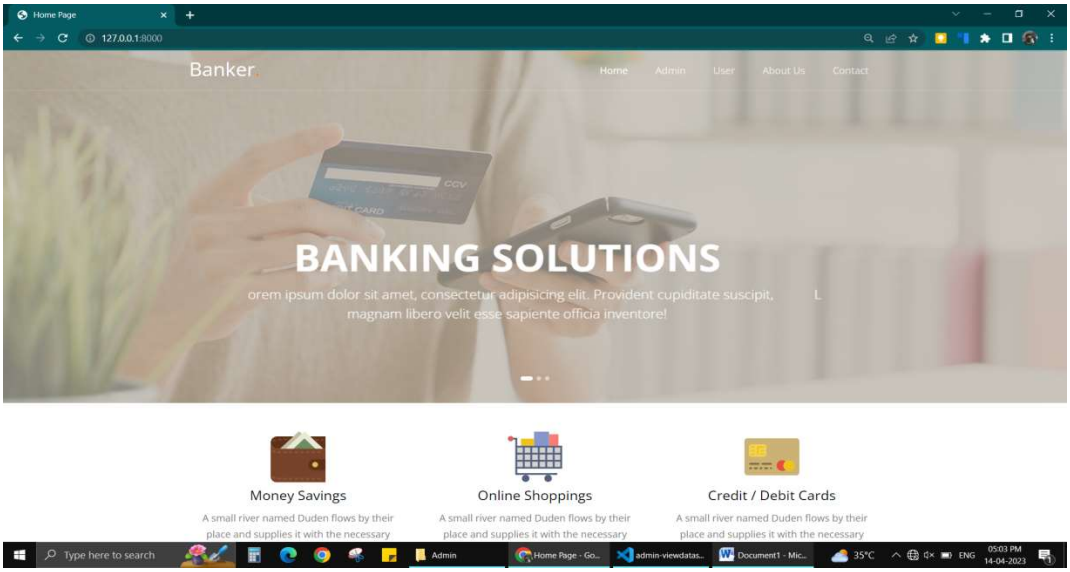

Figure.2. Home page

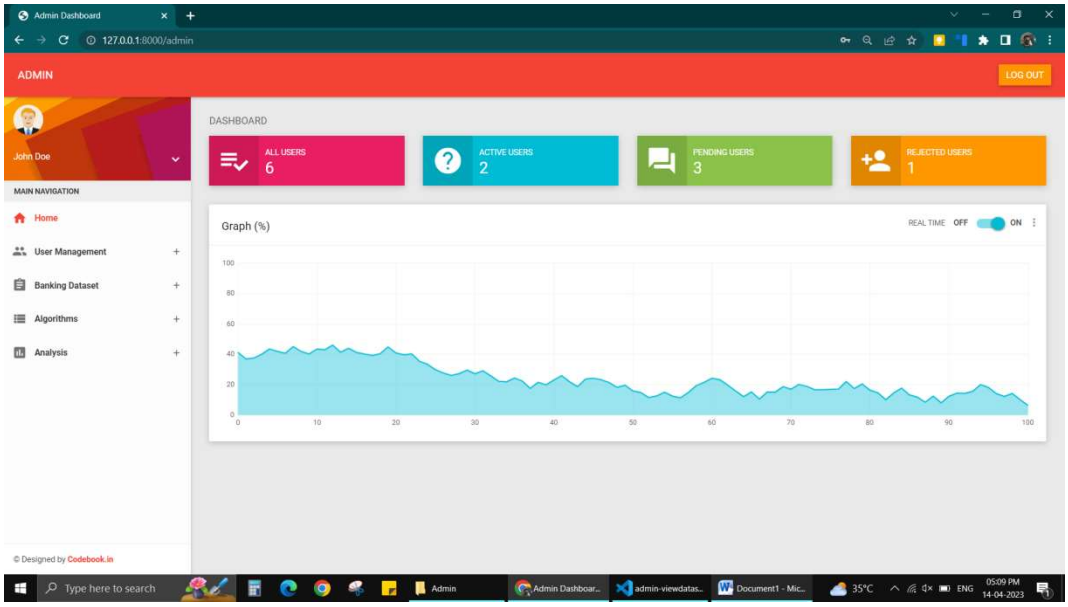

Figure.3. Admin Dash

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.26310602
http://magellanes.com/

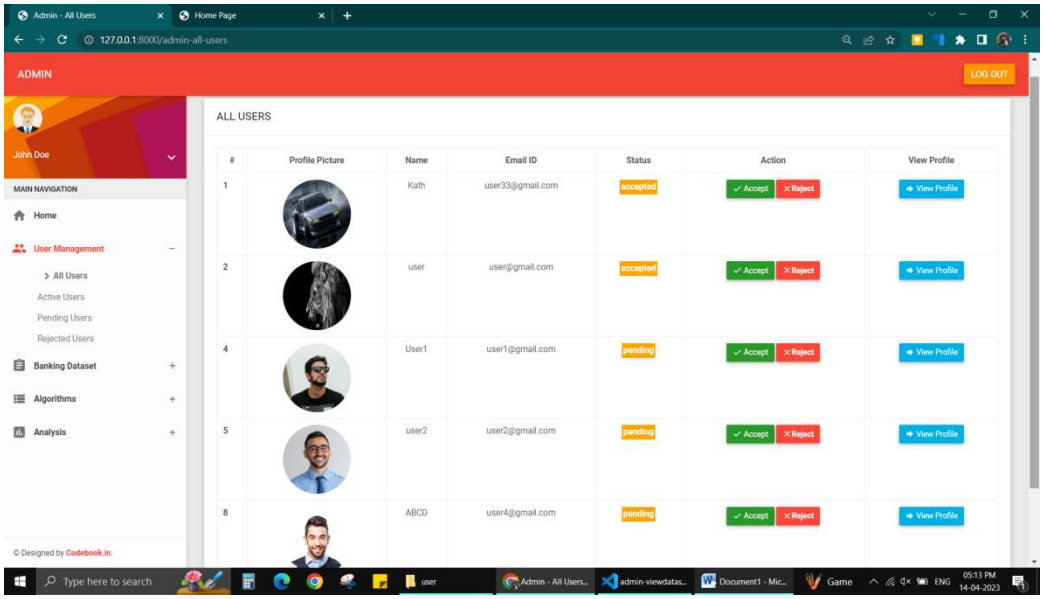

Figure.4 All users

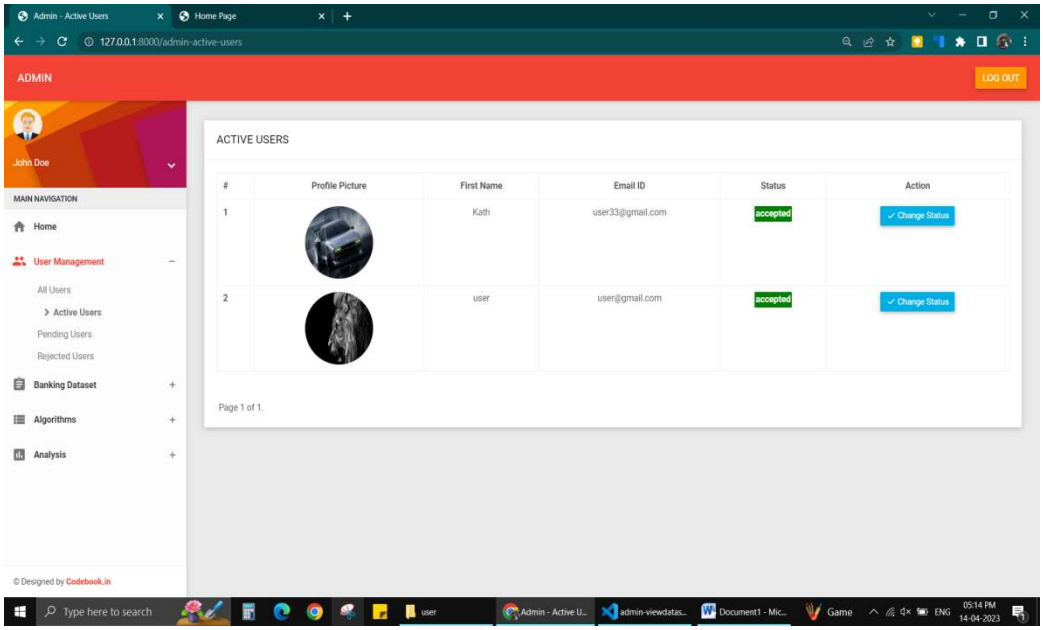

Figure.5.Active users

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.26310602
http://magellanes.com/

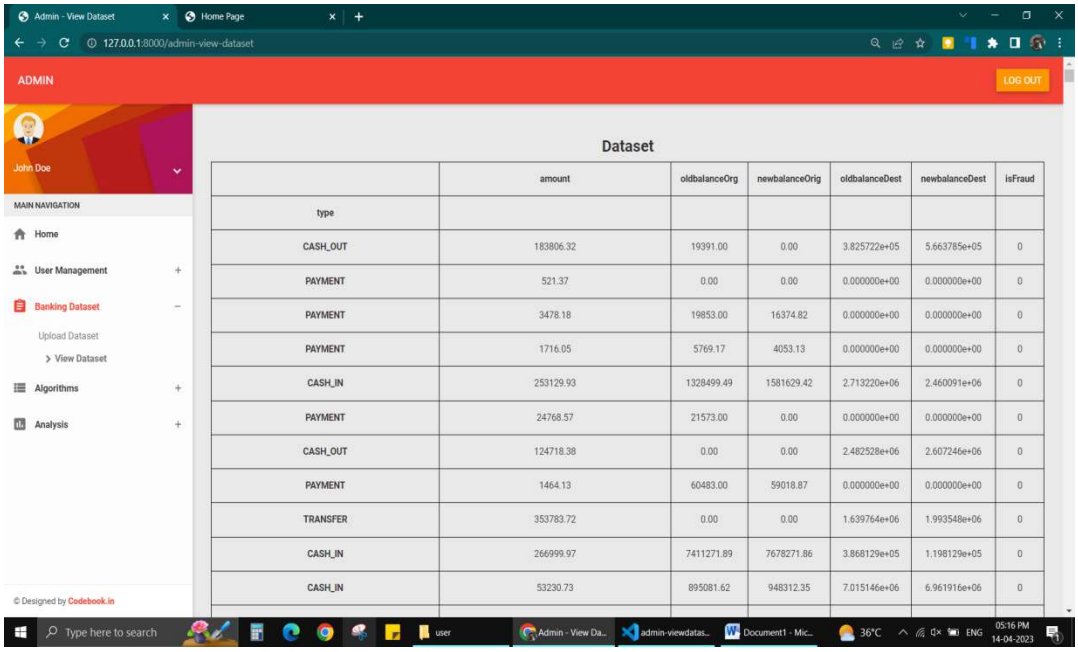

Figure.6.View Data set

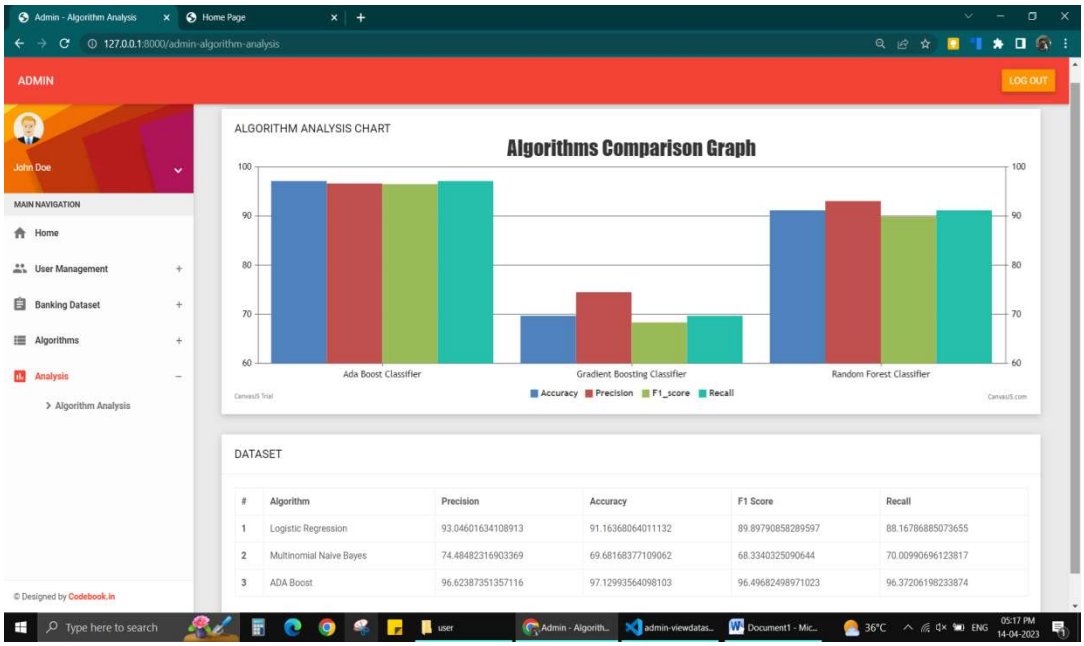

Figure.7. Algorithm Analysis

ISSN:1624-1940
DOI 10.6084/m9.figshare.26310602
http://magellanes.com/

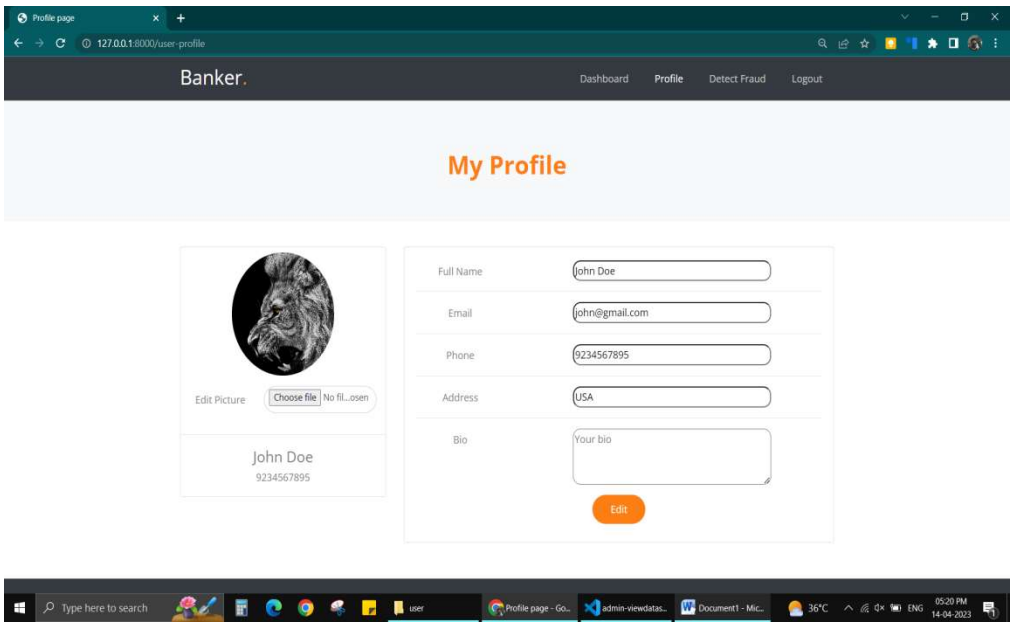

Figure.7.User Profile

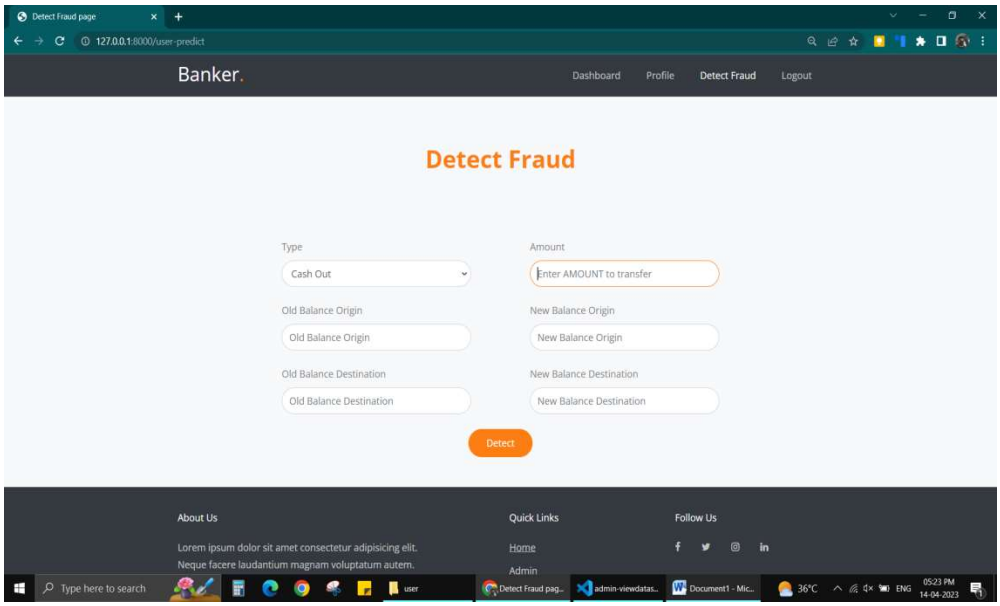

Figure.8. Detect Fraud

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.26310602
http://magellanes.com/

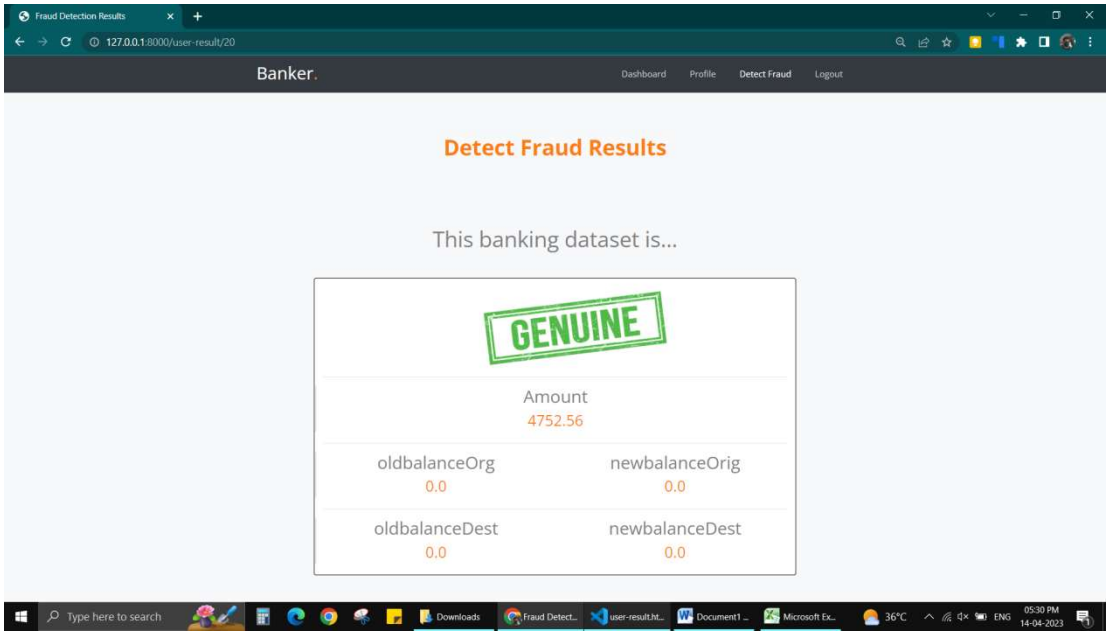

Figure.9 Genuine Result

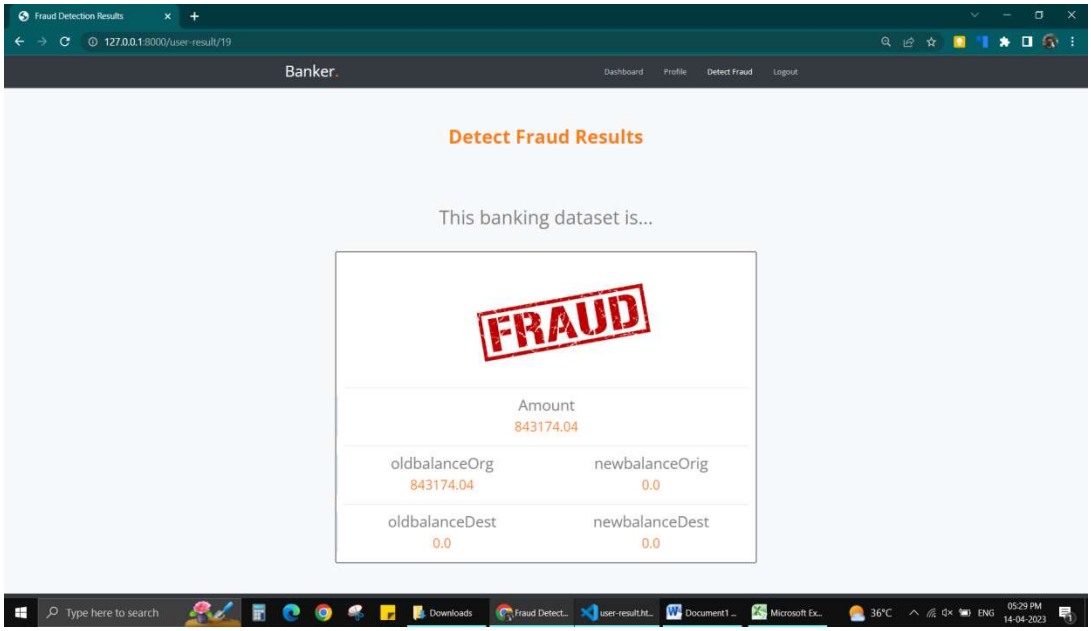

Figure.9 Fraud Result8.CONCLUSION

This study explores the potential applications of machine learning techniques for financial application fraud detection using a publicly available dataset from the UCI repository. The dataset exhibits a significant class imbalance, heavily favoring the majority class. To address this issue, the Synthetic Minority Over-sampling Technique (SMOTE) is employed. The boosting method XGBoost is used alongside the development of K-Nearest Neighbors (KNN) and Random Forest algorithms. The model achieved an impressive performance, with an accuracy of 97.74%. Analysis of the data revealed that customers aged 19 to 25 are more likely to engage in fraudulent activities compared to other age groups.

ISSN:1624-1940
DOI 10.6084/m9.figshare.26310602
http://magellanes.com/

9.FUTURE ENHANCEMENT

Future work could focus on enhancing real-time detection capabilities and experimenting with advanced algorithms and larger datasets to improve robustness and accuracy. Additionally, expanding the model's application to other domains and further developing explainability techniques would increase its versatility and stakeholder trust.

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

ISSN:1624-1940
DOI 10.6084/m9.figshare.26310602
http://magellanes.com/

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_3- Credit_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.; Karakatiˇc, 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.