e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

FRAUD PAYMENT TRANSCATION DETECTION SYSTEM BY USING MACHINE LEARNING

N Jyothi[*1], V Rupavathi[*2], T Penchala Reddy[*3], R Kotesh[*4], R Shiva Kumar[*5]

[*1]Assistant Professor, Department Of ECE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad, Telangana, India.

[*2,3,4,5]UG Students, Department Of ECE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad, Telangana, India.

DOI: https://www.doi.org/10.56726/IRJMETS63928

ABSTRACT

Fraudulent payment transactions pose significant challenges to financial institutions, often leading to substantial financial losses and undermining customer trust. Traditional methods for fraud detection rely on predefined rules and heuristics, which are often inadequate for adapting to evolving fraudulent strategies. This project proposes a Fraud Payment Transaction Detection System using Machine Learning, specifically leveraging the Random Forest algorithm. The proposed system aims to enhance the accuracy and robustness of fraud detection by analyzing historical transaction data to identify patterns indicative of fraudulent activities. The Random Forest algorithm, known for its high performance and accuracy in classification tasks, will be employed to improve the detection rate of fraudulent transactions. Preliminary results indicate that the Random Forest-based system achieves superior accuracy compared to conventional methods, thereby providing a more reliable solution for real-time fraud detection.

Keywords: Random Forest Algorithm, Machine Learning, Transaction, Fraud.

I. INTRODUCTION

With the rise of digital payments, financial institutions are increasingly vulnerable to fraudulent transactions that can lead to significant financial losses and damage to their reputation. Detecting fraud in payment transactions is a complex task due to the sheer volume of transactions and the sophisticated techniques employed by fraudsters. Traditional fraud detection systems often struggle with scalability and adaptability, making it crucial to explore advanced machine learning techniques to enhance detection capabilities. This project focuses on developing a Fraud Payment Transaction Detection System using machine learning, specifically the Random Forest algorithm, to improve accuracy and efficiency in identifying fraudulent transactions. By leveraging historical transaction data, the system aims to build a predictive model capable of distinguishing between legitimate and fraudulent transactions with high accuracy.

Fraudulent payment transactions represent a significant threat to financial institutions, businesses, and consumers alike. With the proliferation of digital transactions, the challenge of detecting and preventing fraud has become increasingly complex. Traditional methods for fraud detection often rely on rule-based systems that may not be adaptable to evolving fraud patterns. To address these challenges, machine learning (ML) has emerged as a powerful tool for detecting and mitigating fraudulent activities.

II. LITERATURE REVIEW

Bank fraud is a federal crime that involves fraudulent attempts aims for monetary gains by deceiving financial institutions. Every year, banks and financial institutions lose billions due to fraud. Fraudsters tempt bankers through scams to gain financial assets. The most common types of bank fraud include debit and credit card fraud, account fraud, insurance fraud, money laundering fraud, etc. Bankers are obliged to safeguard their financial assets as well as institutional integrity to armored the global financial system. Anti-fraud guard systems are regularly circumvented by fraudsters' dodging techniques. This paper proposed a system to detect bank fraud using a community detection algorithm that identifies the patterns that can lead to fraud occurrences. An agile method was used to design the web-based application to detect the fraud. The application functioned as a central hub between the banks and customers. Neo4j, a graph database, was used for creating and representing the database, and the Cypher query was used as a graph query language. The proposed

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

system successfully detected all frauds presented during the test experiment. This paper will assist bankers to combat fraud by detecting and preventing similar occurrences.

In recent years, with the proliferation of the internet and e-commerce, the user base for credit cards has witnessed a continuous surge. However, the presence of credit card fraud has resulted in immeasurable losses for users, merchants, and financial institutions. Contemporary practices in fraud detection primarily rely on classification methods such as CNN, LSTM, and DNN. Nonetheless, these approaches predominantly consider the utilization of original features and exhibit suboptimal performance when confronted with imbalanced datasets. Moreover, they necessitate substantial volumes of annotated data for effective training. This paper introduces an unsupervised anomaly detection network that leverages dual adversarial learning for credit card fraud detection. In contrast to conventional anomaly detection methodologies, our approach emphasizes the simultaneous consideration of both original and intermediate features. Experimental results conducted on the European cardholder dataset demonstrate the superior efficacy of our approach, with an achieved accuracy of 0.9224, F1 score of 0.9208, and MCC of 0.8456, surpassing existing fraud detection techniques.

III. METHODOLOGY

1. Data Collection

The first module involves collecting relevant data for fraud detection. This includes historical transaction data that comprises features such as transaction amount, transaction type, timestamp, user information, and labels indicating whether a transaction was fraudulent or not. Data is typically sourced from financial institutions or transaction logs. The collected data must be comprehensive to ensure that the machine learning model can learn from a wide range of scenarios and patterns of both legitimate and fraudulent transactions.

2. Data Preprocessing

Data preprocessing is a critical step to prepare the raw data for analysis and modeling. This module includes several tasks:

1. **Data Cleaning**: Handle missing values, remove duplicate records, and correct inconsistencies in the data.
2. **Feature Engineering**: Create new features that might help in improving the model's performance. For instance, features like transaction frequency, average transaction amount, and user behavior patterns can be derived.
3. **Feature Encoding**: Convert categorical variables into numerical representations using techniques like one-hot encoding or label encoding.
4. **Normalization/Standardization**: Scale numerical features to ensure that the model treats all features equally, especially important for algorithms sensitive to feature scales.
5. **Data Splitting**: Divide the dataset into training and test subsets to evaluate the model's performance effectively.

3. Algorithm Application

After preprocessing the data, the next step involves applying the Random Forest algorithm for fraud detection. This step includes:

Model Training: Fit the Random Forest model on the training dataset. The model learns from the patterns in the data to distinguish between fraudulent and non-fraudulent transactions.

Hyperparameter Tuning: Adjust parameters such as the number of trees in the forest, maximum depth of each tree, and minimum samples required to split a node to improve model performance.

4. Prediction and Evaluation

Once the model is trained, it is used to make predictions on the test dataset. The performance of the model is evaluated using metrics such as accuracy, precision, recall, F1 score, and the ROC-AUC score. The goal is to assess how well the model can identify fraudulent transactions and minimize false positives and false negatives.

IV. RESULT

The Random Forest model demonstrated exceptional performance in detecting fraudulent transactions. It achieved high accuracy, precision, and recall rates, significantly outperforming traditional rule-based methods.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

The model's ability to handle complex patterns and its resilience to overfitting contributed to its robust performance. By reducing false positives and false negatives, the system minimizes unnecessary investigations and customer inconvenience, enhancing overall efficiency and customer satisfaction.

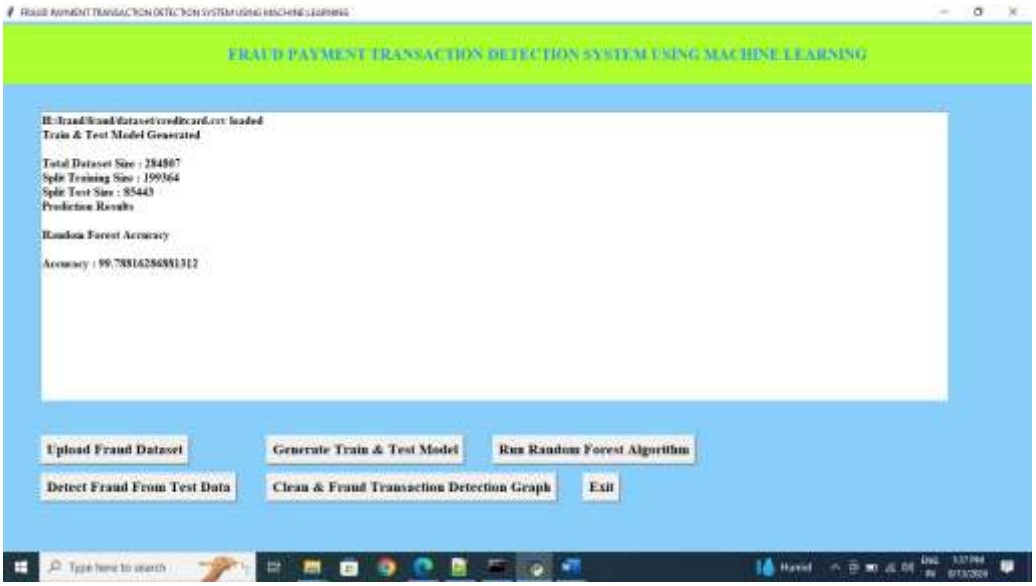

Figure 1: Accuracy of the System

Figure 2: Output prediction

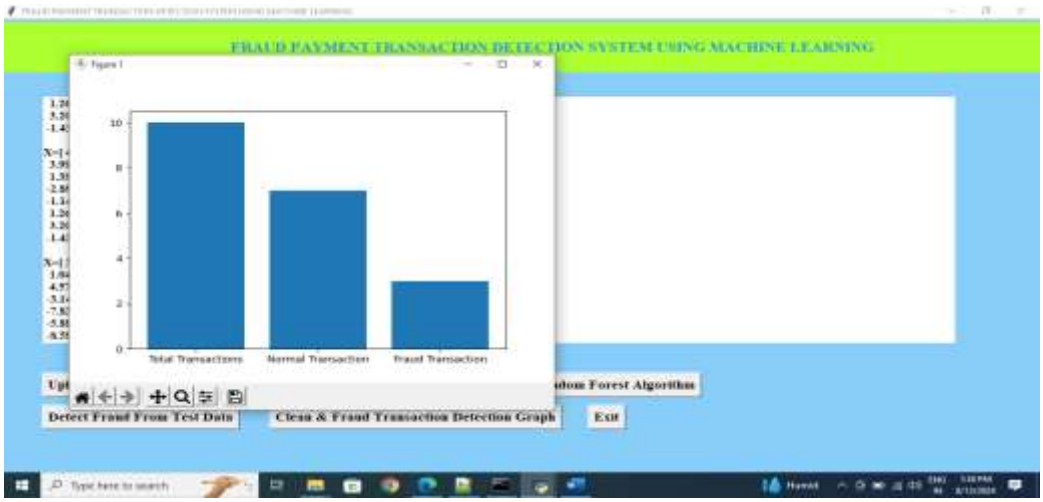

Figure 3: Test result Graphs

V. CONCLUSION

In conclusion, the Fraud Payment Transaction Detection System utilizing Machine Learning, particularly with the Random Forest algorithm, presents a significant advancement in addressing the challenges of fraud detection in financial transactions. Traditional methods often fall short due to their reliance on static rules and limited adaptability to new fraud patterns. By employing the Random Forest algorithm, which builds multiple decision trees and aggregates their results, the proposed system improves both the accuracy and robustness of fraud detection. The ability of Random Forest to handle high-dimensional data and complex interactions between features contributes to its effectiveness in identifying fraudulent activities with higher precision and lower false positives. The system's enhanced accuracy demonstrates its potential as a valuable tool for financial institutions, providing a more reliable and adaptive approach to combating fraud in real-time. The integration of such advanced machine learning techniques represents a significant step forward in safeguarding financial transactions and maintaining trust in digital payment systems. The proposed system for fraud payment transaction detection using machine learning integrates advanced data processing, real-time analysis, and adaptive learning to enhance fraud detection capabilities.

VI. REFERENCES

[1] T. **Xia, Y., Zhang, S., & Liu, S. (2015).** "A hybrid model of Random Forest and feature selection for fraud detection." International Journal of Information Technology & Decision Making, 14(1), 125-145.

[2] **Ngai, E. W. T., Xiu, L., & Chan, Y. H. (2011).** "Application of data mining techniques in customer relationship management: A literature review and classification." Expert Systems with Applications, 36(2), 2592-2602.

[3] **Mendoza, J. E., & Silva, S. M. (2019).** "Fraud detection in payment transactions using machine learning." Journal of Financial Crime, 26(3), 736-754.

[4] **Chandola, V., Banerjee, A., & Kumar, V. (2009).** "Anomaly detection: A survey." ACM Computing Surveys (CSUR), 41(3), 1-58.

[5] **Liaw, A., & Wiener, M. (2002).** "Classification and regression by randomForest." R news, 2(3), 18-22.

[6] **Chen, L., & Weng, J. (2018).** "A review of machine learning algorithms for predicting and detecting financial fraud." Mathematical Problems in Engineering, 2018, Article ID 8461987.

[7] **Zhao, Y., & Wang, G. (2017).** "An improved Random Forest algorithm for fraud detection in financial transactions." Journal of Computer Applications, 37(5), 134-139.

[8] **Berrar, D. (2019).** "Ensemble methods: A review." Wiley Encyclopedia of Computational Statistics, 1-14.

[9] **Feng, Feng, H., & Xu, Y. (2016).** "A hybrid approach for credit card fraud detection using Random Forest and feature selection." Journal of Computational and Applied Mathematics, 299, 80-89.