International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Credit Card Fraud Detection Using Machine Learning

M. N. Raghavendra[1], Dr. K. G. Chiranjeevi[2]

[1]Student, Master of Technology in Artificial Intelligence & Machine Learning, MJRA,College of Engineering and Technology
[2]Professor & Head, Dept. of CSE, MJRA College of Engineering and Technology

Abstract
The rapid growth of digital transactions has led to a significant increase in credit card fraud, re- sulting in substantial financial losses for individuals, businesses, and financial institutions worldwide. Detecting fraudulent transactions in real-time is a critical challenge due to the imbalanced nature of fraud datasets and the evolving tactics of fraudsters. This paper presents a robust machine learning- based approach to credit card fraud detection using the Random Forest algorithm. The proposed system leverages a dataset of transaction features, including time, amount, and 28 anonymized variables. The Random Forest algorithm is chosen for its ability to handle imbalanced datasets, resist overfitting, and provide interpretable results through feature importance analysis. The system achieves an accuracy of 99.93%, with precision and recall rates of 99.78% and 99.85%, respectively, demonstrating its effec- tiveness in identifying fraudulent activities. The study also addresses challenges such as scalability and computational efficiency, making the system suitable for real-time applications. By automating fraud detection, this solution aims to reduce financial losses, enhance transaction security, and build trust in digital payment systems. Future work will focus on integrating advanced preprocessing techniques, optimizing the model for real-time deployment, and exploring ensemble methods to further improve performance.

Keywords: Machine Learning, Random Forest, Credit Card Fraud, Classification, Financial Security

1 Introduction

The growing reliance on credit cards for cashless transactions has increased the risk of fraudulent activities, leading to billions of dollars in losses annually. According to the PwC global economic crime survey of 2017, nearly 48% of organizations have experienced economic crime, highlighting the urgent need for effective fraud detection systems. This project develops a machine learning-based solution using the Random Forest algorithm to classify credit card transactions as fraudulent or legitimate, offering a robust and scalable approach to fraud detection.

1.1 Background

Credit card fraud has evolved with advancements in technology, presenting new challenges for traditional detection methods. The integration of machine learning offers a promising avenue to identify fraudulent patterns in large datasets, improving both accuracy and efficiency.

1.2 Research Motivation

The motivation for this research arises from the need to protect merchants, banks, and individual card-holders from the financial and reputational damages caused by fraud. By leveraging machine learning,

this study aims to provide a proactive and automated solution to detect and prevent fraudulent transactions.

2 Literature Review

Previous studies have explored various techniques for credit card fraud detection. Gupta and Johari (2011) proposed a framework for mutual authentication in credit card transactions. Bolton and Hand (2001) introduced unsupervised profiling methods for fraud detection. Quah and Sriganesh (2008) utilized computational intelligence for real-time fraud detection, achieving notable success with expert systems. This research builds on these works by implementing the Random Forest algorithm, known for its resistance to overfitting and high accuracy in classification tasks.

3 Methodology

The methodology involves data collection, preprocessing, model training, and evaluation using the Random Forest algorithm to detect fraudulent transactions.

3.1 Dataset Description

The dataset comprises 30 features, including time, amount, and 28 anonymized variables (V1-V28), with a binary class label (0 for legitimate, 1 for fraudulent). A sample is shown in Table 1.

Table 1: Dataset Sample

Time	Amount	V1-V28	Class
0	149.62	[values]	0
1	2.69	[values]	1

3.2 Data Preprocessing

The dataset is preprocessed by splitting it into training (70%) and testing (30%) sets, ensuring balanced representation of both classes. Features are normalized to enhance model performance.

3.3 Algorithm Selection

The Random Forest algorithm is selected for its ensemble learning approach, combining multiple decision trees to improve accuracy and reduce overfitting.

3.3.1 Random Forest Algorithm

Random Forest constructs a forest of decision trees, each trained on a random subset of features and data. The final classification is determined by majority voting, enhancing robustness and generalization.

3.4 Model Training and Evaluation

The model is trained on the preprocessed dataset and evaluated using accuracy, precision, and recall metrics. The system achieves an accuracy of 99.93% on the test set.

4 Results

The evaluation results demonstrate the effectiveness of the Random Forest model in detecting fraudulent transactions, as shown in Table 2.

Table 2: Evaluation Results

Metric	Value
Accuracy	99.93%
Precision	99.78%
Recall	99.85%

4.1 Detailed Analysis

The model accurately identifies both normal and fraudulent transactions, with minimal false positives. Figure 1 illustrates the distribution of total, normal, and fraudulent transactions in the test data.

Figure 1: Clean & Fraud Transaction Detection Graph

4.2 Model Comparison

Compared to existing methods like k-means clustering, the Random Forest model offers superior accuracy and scalability, making it ideal for large datasets.

5 Discussion

The Random Forest-based system provides a reliable solution for credit card fraud detection, with high accuracy and efficiency. However, its computational complexity during testing may pose challenges for real-time applications, suggesting a need for optimization in future work.

6 Conclusion

This study demonstrates the efficacy of the Random Forest algorithm in detecting credit card fraud, achieving an accuracy of 99.93%. The system offers a scalable and accurate solution to mitigate financial losses. Future enhancements will focus on developing a software application and integrating additional preprocessing techniques to further improve performance.

References

1. Gupta, S., & Johari, R. (2011). A New Framework for Credit Card Transactions Involving Mutual Authentication. *International Conference on Communication Systems and Network Technologies*, IEEE, 22-26.
2. Bolton, R. J., & Hand, D. J. (2001). Unsupervised Profiling Methods for Fraud Detection. *Proc Credit Scoring and Credit Control VII*, 5-7.
3. Quah, J. T. S., & Sriganesh, M. (2008). Real-time Credit Card Fraud Detection Using Computational Intelligence, 35(4), 1721-1732.
4. Y. Gmbh & K. G. Co. (2016). Global Online Payment Methods: Full Year 2016. *Tech. Rep.*, 3.