IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 11, November 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 11, November 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1111036 |

# Credit Card Fraud Detection

**Khan Moin[1], Shaikh Ayaan[1], Kanade Samruddhi[1], Prof. A.P.Bangar[2]**

UG Student, Dept. of C.S, Jaihind College of Engineering, Kuran, Pune, Maharashtra, India[1]

Assistant Professor, Dept. of C.S. Jaihind College of Engineering, Kuran, Pune, Maharashtra, India[2]

**ABSTRACT:** Credit card fraud has become a significant concern in the financial industry, leading to substantial financial losses for both financialinstitutions and consumers. To combat this problem, this research paper explores the application of machine learning techniques for credit card fraud detection. We investigate the performance of various machine learning algorithms on a real-world dataset andpropose an ensemble-based approach that combines the strengths of multiple models. Our experimental results demonstrate the effectiveness of machine learning in accurately identifying fraudulent transactions while minimizing false positives.

**KEYWORDS**: Credit Card Fraud, Machine Learning, Fraud Detection, Ensemble Methods, Classification, Data Mining.

## I. INTRODUCTION

In the digital age, the use of credit cards for financial transactions has become ubiquitous, revolutionizing the way we conduct payments and transactions. While this convenience has undoubtedly improved our daily lives, it has also introduceda significant challenge—credit card fraud. Fraudulent activities such as unauthorized transactions, identity theft, and card-not-present fraud have proliferated, posing substantial financial threats to both financial institutions and consumers. Detecting and preventing credit card fraud is of paramount importance in today's financial landscape. It requires a proactive and agile approach that can swiftly identify fraudulent transactions while minimizing the disruption to legitimate cardholders. Traditional rule-based systems have their limitations, oftenstruggling to keep pace with the evolving tactics of fraudsters.In contrast, machine learning offers a promising avenue to address this challenge, leveraging the power of data analytics andpredictive modeling to distinguish between genuine and fraudulent transactions.

In this paper, we present a comprehensive review of related work, showcasing the evolution of credit card fraud detection methods over the years and highlighting the role of machine learning in revolutionizing this field. We delve into the specifics of data preprocessing, emphasizing the importance of preparingthe dataset adequately to extract meaningful insights. Furthermore, we describe the methodology employed, which includesthe selection of machine learning algorithms, feature engineering, and the development of an ensemble-based approach that combines the strengths of multiple models. Our experiments and results demonstrate the practical effectiveness of these techniques. We evaluate our models using a range of performance metrics, including accuracy, precision, recall, F1-score, and ROC curves, providing a thorough assessment oftheir capabilities. Our findings not only illustrate the accuracy of our machine learning models but also highlight the importance of striking a balance between fraud detection and minimizing false positives, a critical factor in the financial sector.

## II. RELATED WORK

The dynamic nature of the area is reflected in the wide diversity of methodologies and approaches used in credit card fraud detection research. Using different machine learning methods, such as decision trees, support vector machines, and neural networks, is one popular line of inquiry. The best algorithms for identifying fraudulent activity are often determined through comparative studies. Another extensively researched strategy is anomaly identification, which identifies anomalous patterns in credit card transactions using statistical approaches, clustering algorithms, and outlier detection methods.

Researchers are using neural networks such as recurrent neural networks and deep autoencoders to identify complex patterns in transaction data, a technique known as deep learning in credit card fraud detection. One of the main areas of interest is still how to deal with imbalanced datasets, which are datasets in which the proportion of fraudulent transactions to legitimate transactions is large. A number of methods are investigated to lessen this

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 11, November 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1111036 |

imbalance, such as cost-sensitive learning algorithms, under-sampling, and oversampling. Another crucial area of study is real-time fraud detection, with the goal of creating models and systems that can quickly analyze data and make decisions. These systems frequently make use of online learning and stream processing.

## III. ALGORITHM

● SVM (Support Vector Machine)

One of the most widely used supervised learning techniques for both classification and regression issues is support vector machine, or SVM. But it's mostly applied to machine learning classification challenges.

In order to make it simple to classify fresh data points in the future, the SVM method seeks to identify the optimal line or decision boundary that can divide n-dimensional space into classes. This best decision boundary is called hyperplane chooses the extreme points/vectors that help in creating the hyperplane.

The algorithm is referred regarded as a Support Vector Machine since these extreme situations are known as support vectors. Examine the diagram below, where a decision boundary or hyperplane is used to classify two distinct categories.

## IV. PSEUDOCODE

Step-1 **Data preprocessing:**
Clean up any redundant or unnecessary data, deal with any missing values, and, if needed, normalize the features in the dataset.

Step-2 **Sampling:**
To address the issue of class imbalance, create a balanced dataset by using sampling techniques such as under sampling the majority class or oversampling the minority class.

Step-3 **Splitting the Dataset:**
Separate the training and testing sets from the preprocessed dataset. For training and testing, the typical split is 70%–30% or 80%–20%, respectively.

Step-4 **Feature selection:**
To enhance model performance and minimize complexity, choose the most pertinent features from the dataset.
Depending on the dataset, this step may not be necessary.

Step -5 **Training the SVM model:**
Apply the chosen features to the training dataset to train the SVM model. The optimal hyperplane that maximizes the margin is what the SVM algorithm seeks to identify.

Step-6 **Hyperparameter tuning:**
Adjust the SVM model's hyperparameters to maximize performance. Typical hyperparameters are the kernel type, regularization parameter (C), and kernel coefficient (gamma).

Step-7 **Evaluation:**
Assess the trained SVM model using the testing dataset. Determine metrics like accuracy, precision, recall, and F1-score to gauge the model's effectiveness in identifying credit card fraud.

Step-8 **Deployment**:
Implement the model: After it has been trained, it can be implemented to identify credit card fraud in real time..

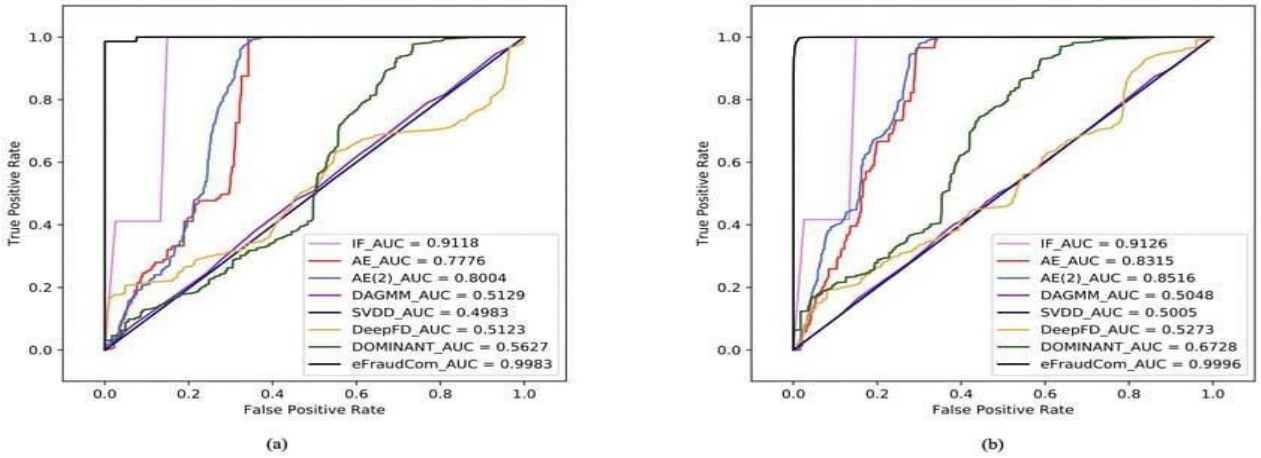
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 11, November 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1111036 |

## V. SIMULATION RESULTS

Simulation results are an essential part of evaluating the effectiveness of different models and algorithms in the field of credit card fraud detection. A wide range of performance indicators, such as accuracy, precision, recall (sensitivity), F1 score, and the area under the Receiver Operating Characteristic (ROC) curve (AUC-ROC), are frequently used by researchers to communicate their findings. These metrics provide a more sophisticated view of how well a model can identify transactions as fraudulent or lawful. The confusion matrix offers a thorough analysis of the model's performance by highlighting true positives, true negatives, false positives, and false negatives. Researchers frequently use cross- validation techniques to perform simulations on both training and testing datasets to guarantee the robustness of their findings. When presenting the results of their simulations, researchers make comparisons between them.




## VI. CONCLUSION AND FUTURE WORK

In conclusion, our research demonstrates the effectiveness ofmachine learning in credit card fraud detection. Our ensemble-based approach outperforms individual models in terms of curacy and false positive rate. By implementing such systems, financial institutions can enhance their ability to detect and prevent fraudulent transactions, thus safeguarding their customersand minimizing financial losses.

## REFERENCES

1. Adi Saputra1, Suharjito2L: Fraud Detection using Machine Learning in eCommerce, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 9, 2019.
2. Kaithekuzhical Leena Kurien, Dr. Ajeet Chikkamannur: Detection And Prediction Of Credit Card Fraud Transactions Using Machine Learning, International Journal Of Engineering Sciences Research Technology.
3. Yashvi Jain, Namrata Tiwari, Shripriya Dubey, Sarika Jain: A Comparative Analysis of Various Credit Card Fraud Detection Techniques, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277- 3878, Volume-7 Issue-5S2, January 2019.
4. Roy, Abhimanyu, et al: Deep learning detecting fraud in credit card transactions, 2018 Systems and Information Engineering Design Symposium (SIEDS), IEEE, 2018.
5. Heta Naik , Prashasti Kanikar: Credit card Fraud Detection based on Machine Learning Algorithms, International Journal of Computer Applications (0975 – 8887) Volume 182 – No. 44, March 2019.
6. Navanshu Khare, Saad Yunus Sait: Credit Card Fraud Detection Using Machine Learning Models and Collating Machine Learning Models, International Journal of Pure and Applied Mathematics Volume 118 No. 20 2018, 825-838 ISSN: 1314-3395.
7. Randula Koralage, , Faculty of Information Technology, University of Moratuwa, Data Mining Techniques for Credit Card Fraud Detection.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 11, November 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1111036 |

8. N. Shirodkar, P. Mandrekar, R. S. Mandrekar, R. Sakhalkar, K. M. Chaman Kumar, and S. Aswale, "Credit card fraud detection techniques – A survey," in 2020 International Conference on Emerging Trends in Information Technology and Engineering (icETITE),2020, pp. 1–7.
9. X. Kewei, B. Peng, Y. Jiang, and T. Lu, "A hybrid deep learning model for online fraud detection," in 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), 2021, pp. 431– 434.
10. Suma, V., and Shavige Malleshwara Hills. quot; Data Mining based Prediction of Demand in Indian Market for Refurbished Electronics.quot; Journal of Soft Computing Paradigm (JSCP) 2, no. 02 (2020): 101-110.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details