IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

Credit Card Fraud Detection Using Machine Learning

Lohith E[1] and Mrs. Priyanka Mohan[2]
Student MCA, IVth Semester[1]
Assistant Professor, Department of MCA[2]
Dayananda Sagar Academy of Technology and Management, Udaypura, Bangalore, Karnataka, India
lohith.4518@gmail.com and priyanka-mca@dsatm.edu.in

Abstract: *The financial industry is very concerned about credit card theft because it can result in large losses. The Random Forest technique, which is renowned for its accuracy and capacity to handle unbalanced datasets, is used in this project to create a strong fraud detection system. Based on several characteristics, the model that was trained on a labeled dataset of credit cardtransactions can differentiate between authentic and fraudulent activity. Using Flask, HTML, CSS, and JavaScript for a responsive front end and Python and JavaScript for a reliable back end, the system is integrated into a full-stack web application. For safe data storage, MySQL is used. The web tool uses sophisticated data visualization to provide insights while enabling real-time transaction monitoring, fraud warnings, and historical data analysis*

Keywords: Credit Card Fraud, Random Forest, Data Visualization, Fraud Detection, Real- Time Monitoring, and Web Development

I. INTRODUCTION

Credit card fraud is a serious problem that can result in large losses in terms of money as well as erode consumer confidence. Recognizing fraudulent activity has become increasingly important as the number of online transactions rises. Real-time fraud detection systems seek to spot questionable transactions to stop illegal access and minimize losses.

Recognized for its accuracy and resilience, the Random Forest algorithm works wonders for identifying fraudulent activity. To identify fraudulent transactions among several valid ones, it builds multiple decision trees and outputs their mode for classification, managing large datasets and imbalanced data.

D. Tanouz et al,. [1] suggest the use of credit cards has increased, which has increased the number of fraud cases and made the creation of precise fraud detection algorithms necessary. To address imbalanced datasets and enhance detection accuracy, precision, recall, F1 score, and ROC-AUC score, this study suggests machine learning classification algorithms such as logistic regression, random forest, and Naive Bayes.

Maniraj S P et al,.[2] suggest to keep customers safe, credit card companies need to identify fraudulent transactions. Using methods like the local outlier factor and the isolation forest, this project uses machine learning to model historical transactions to detect fraud. The goal is to achieve 100% detection while decreasing false positives.

Our system is a full-stack web application that uses Random Forest. The front end offers real- time monitoring and alarms through the use of Flask, HTML, CSS, and JavaScript. The back end, which was created using JavaScript and Python, handles data processing and securely connects to a MySQL database.

II. LITERATURE SURVEY

Parvati R et al,. [3] Introduced significant financial losses have resulted from an increase in fraud due to the growing use of credit cards. This research uses deep learning and machine learning techniques to identify these types of frauds. Through the use of techniques including Random Forest, SVM, XGBoost, CNN architectures, and logistic regression, the project aims to reduce false alarms, increase accuracy, and successfully deploy credit card fraud detection in the real world.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

Ruttala Sailusha et al,. [4] describe the prevalence of unlawful use of credit card information as increased due to the rise in online purchases and credit card fraud. The goal of this research is to use machine learning algorithms—Random Forest and Ada boost in particular—to identify fraudulent activity. The best approach is found by comparing the algorithms' performances in terms of accuracy, precision, recall, and F1 score.

The rise in e-commerce has raised the danger of credit card fraud. This paper is by Vaishnavi Nath Dornadula et al,. [5] designed which grouped cardholders, utilizing a sliding window approach to enhance classifier performance, and examining customer behaviour patterns, this paper seeks to create a novel fraud detection method for streaming transaction data. It also addresses concept drift.

Credit card fraud has surged due to the growth of e-commerce, underscoring the necessity for efficient detection. This paper is by Yanxia Sun & Zenghui Wang et, al. [6] suggests that in contrast to current techniques, this research suggests a machine learning-based fraud detection engine that selects features using the genetic algorithm, greatly increasing detection accuracy.

III. METHODOLOGY

Existing Method

Various machine learning techniques are used in existing approaches for credit card fraud detection to improve accuracy and performance. In addition to classifiers like decision trees, random forests, logistic regression, artificial neural networks, and naive Bayes, techniques like genetic algorithms for feature selection are employed. To further enhance detection accuracy and mitigate idea drift, techniques such as the sliding window strategy for streaming data and the comparative analysis of Random Forest and Ada boost are utilized.

Proposed Method

In our project, we created a machine-learning application to control fraudulent credit card transactions. The suggested approach to credit card fraud detection combines real-time data analysis, optimal feature selection, and sophisticated machine learning algorithms. The goals encompass applying classifiers such as Random Forest, Decision Tree, ANN, and Naive Bayes, using genetic algorithms for feature selection, and creating a sliding window approach for streaming data analysis.

Based on the value of their transactions, cardholders are grouped, and algorithms are compared. Concept drift is addressed using a feedback mechanism that has been verified on an actual dataset of European cardholders.

We gathered information from many credit card transactions, including the following: expiration date, card type, gender, credit card balance, credit card limit, credit card over the limit, excess over the limit, and class (target variable). To detect credit card fraud, we use machine learning techniques like K-means clustering and random forest. The powerful supervised learning algorithm Random Forest constructs many decision trees during training, which it then mixes to improve prediction accuracy.

We start by pre-processing the transaction data, handling missing values, encoding category variables, and normalizing numerical features. By choosing pertinent qualities for training, the Random Forest algorithm discovers patterns and anomalies from historical data to predict fraudulent transactions. Many metrics are used to evaluate the efficacy of the model, such as accuracy, precision, recall, and F1-score.

Random Forest: Regression and classification are two applications for the ensemble learning system Random Forest. To find trends and anomalies in credit card fraud detection, transaction data is analyzed. It improves prediction accuracy, lowers overfitting, and prioritizes feature importance by constructing numerous decision trees and merging their outputs, offering a reliable and scalable method of identifying fraudulent transactions.

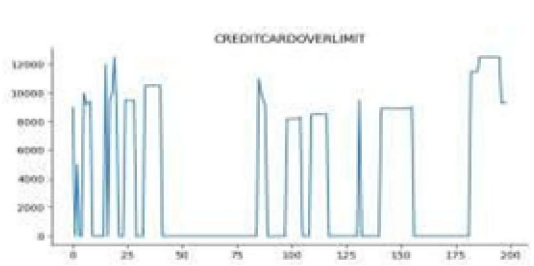
K-Means Clustering: K-Means clustering is an unsupervised learning algorithm that groups data into clusters based on feature similarity. In credit card fraud detection, K-Means helps by identifying patterns and deviations in transaction data. Transactions that significantly deviate from cluster centroids or belong to anomalous clusters can be flagged as potential fraud, aiding in early detection.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

IV. RESULTS AND DISCUSSIONS

Random Forest outperforms K-means clustering in credit card fraud detection because of its increased accuracy and resilience in classification tasks. K-means clusters similar data points together; however, it has trouble explicitly classifying fraudulent data. On the other hand, Random Forest is particularly good at managing unbalanced datasets and offering accurate fraud detection through the creation of several decision trees for trustworthy classification.

Model	Algorithm	Accuracy	Description
Model 1	K-means Clustering	96%	Utilized for initial training, clustering data points to identify patterns in Transaction predictions.
Model 2	Random Forest	98%	Used for final training, leveraging multiple decision trees to improve prediction accuracy.



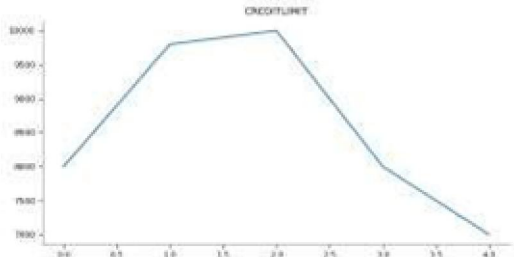


Figure 1: Demonstration of Credit Card Overlimit Figure 2: Demonstration of Credit Limit

Figure 1: The graph illustrates the credit card limit for various transactions. The model will likely flag instances of significant credit limit exceedance as potential fraud. Large spikes in the graph indicate moments when the credit limit has been crossed, triggering the fraud detection model to identify these transactions as suspicious and potentially fraudulent.

Figure 2: The credit limit across different transactions is displayed on the graph. Transactions that stay under the credit limit—as seen by the graph's values—will probably be flagged by the model as not fraudulent. Because of the consistent pattern and lack of notable spikes, the fraud detection algorithm concludes that credit consumption is within allowable bounds and labels these transactions as genuine rather than suspicious.

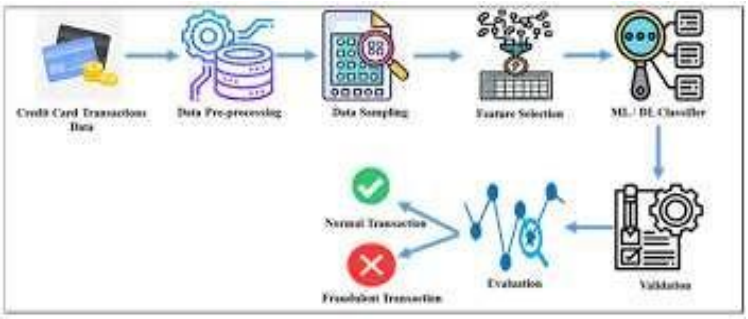


Figure 3: Demonstration of model training

Figure 3: The steps involved in training a model to detect credit card fraud are shown in this diagram. The first step is gathering the raw transaction data, which is then cleaned and normalized by data pre-processing. Class imbalance is addressed via data sampling, and the most important variables are the focus of feature selection. After that, this refined data is used to train classifiers for machine learning (ML) or deep learning (DL). Metrics like accuracy and recall are used to assess the model's performance, and its suitability for generalizing to new data is confirmed. Ultimately, the verified model enhances the accuracy of fraud detection by categorizing transactions as either legitimate or fraudulent.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

V. CONCLUSION

The Algorithm that Random Forest models operate in detecting credit card fraud, attaining an accuracy of 98% compared to another algorithm. Our goal is to improve transaction security and lessen financial losses from fraud by utilizing cutting-edge machine-learning techniques. Our goal is to enhance fraud detection skills by utilizing predictive analytics and real-time transaction monitoring to optimize detection models. To resist increasing fraud strategies and revolutionize security measures in financial transactions, the future scope will involve incorporating more advanced deep learning architectures and investigating adaptive algorithms.

REFERENCES

[1] D Tanouz et al. "Credit Card Fraud Detection Using Machine Learning", May 2021, http://dx.doi.org/10.1109/ICICCS51141.2021.9432308

[2]. Maniraj S P et al,. "Credit Card Fraud Detection Using machine learning And Data science" September 2019, http://dx.doi.org/10.17577/IJERTV8IS090031

[3] Parvathi R et al,. "Credit Card Fraud Detection using Machine Learning", May 2023 https://www.questjournals.org/jses/papers/Vol9-issue-5/09055560

[4] Ruttala Sailusha et al,. "Credit Card Fraud Detection Using Machine Learning", May 2020 http://dx.doi.org/10.1109/ICICCS48265.2020.9121114

[5] Vaishnavi Nath Dornadula et al,. "Credit Card Detection Fraud Detection Using Machine Learning", January 2019, http://dx.doi.org/10.1016/j.procs.2020.01.057

[6] "Credit Card Fraud Detection Using GA Algorithm and Other Features", February 2019 https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00573-8