IJIRCCE

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

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Special Issue 2, March 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 |

| International Conference on Recent Innovations in Engineering and Technology (ICRIET'23)|

| Sharadchandra Pawar College of Engineering, Pune, India |

|| Volume 11, Special Issue 2, March 2023 ||

Credit Card Fraud Detection using AI/ML

Saba Inamdar, Dr.Monika D.Rokade

PG Student, Dept. of Computer, SPPU, Pune, India

Assistant Professor, Dept. of Computer, SPPU, Pune, India

ABSTRACT: In order to prevent customers from being charged for products they did not buy, credit card firms must be able to recognise fraudulent credit card transactions. Data Science may be used to solve these issues, and coupled with machine learning, it is of utmost relevance. With the use of credit card fraud detection, this research aims to demonstrate the modelling of a data set using machine learning. Modelling previous credit card transactions using information from those that turned out to be fraudulent is part of the Credit Card Fraud Detection Problem. Afterwards, a new transaction is analysed using this model to determine whether it is fraudulent or not. The goal here is to minimise inaccurate fraud categories while detecting 100% of the fraudulent transactions.

KEYWORDS— Credit card fraud, applications of machine learning, data science, isolation forest algorithm, local outlier factor, automated fraud detection.

1.INTRODUCTION

Financial institutions lose billions of dollars annually as a result of credit card fraud, which is a serious issue. The complexity and ongoing evolution of fraud make traditional rule-based approaches for detecting it increasingly ineffective. As a result, using machine learning algorithms to spot fraudulent transactions in real time has grown in popularity.

Algorithms that use machine learning can learn from data and spot trends that point to fraudulent activity. These algorithms have the capacity to examine massive amounts of data, spot anomalies, and generate precise forecasts. They are appropriate for dynamic and changing contexts since they can also adapt and learn from fresh data.

In this research, we provide a machine learning-based approach for detecting credit card fraud. The suggested method classifies transactions as either fraudulent or non-fraudulent using a dataset of credit card transactions and a variety of machine learning algorithms. The experimental findings demonstrate that the suggested method detects fraudulent transactions with high accuracy.

The usefulness of several machine learning techniques for detecting credit card fraud, such as logistic regression, decision trees, random forests, support vector machines (SVM), neural networks, and deep learning, is examined in this paper. The paper also investigates the significance of feature selection, ensemble approaches, and data pre-treatment in enhancing the efficiency and accuracy of the algorithms.

The suggested method can help financial institutions and customers by providing a dependable and affordable way to identify credit card fraud in real-time. The technology can assist lower fraud-related losses, boost client happiness, and enhance the overall safety of the financial system.

Overall, by addressing the difficulties and potential future paths for study in this area, this paper offers insights into the use of machine learning algorithms for credit card fraud detection.

II.LITERATURE REVIEW

Detecting credit card fraud is a serious problem for the financial sector since fraudulent transactions cost billions of dollars each year. The complexity and ongoing evolution of fraud make traditional rule-based approaches for detecting it increasingly ineffective. As a result, using machine learning algorithms to spot fraudulent transactions in real time has grown in popularity.As a result, using machine learning algorithms to spot fraudulent transactions in real time has grown in popularity.

The efficiency of machine learning algorithms in detecting credit card fraud has been the subject of numerous studies. For instance, Li et al. (2017) assessed the performance of different machine learning algorithms on a dataset for credit card fraud, including decision trees, random forests, support vector machines (SVM), neural networks, and deep learning.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

| International Conference on Recent Innovations in Engineering and Technology (ICRIET'23)|

| Sharadchandra Pawar College of Engineering, Pune, India |

|| Volume 11, Special Issue 2, March 2023 ||

A credit card fraud detection system using an ensemble of classifiers, including k-nearest neighbours, decision trees, and SVM, was proposed by Phua et al. (2010) in another study. The study demonstrated that the ensemble method outperformed individual classifiers in terms of accuracy.

Machine learning-based credit card fraud detection also heavily depends on feature selection and data pretreatment. As an illustration, Oommen and Cherian (2018) suggested a feature selection strategy based on correlation analysis and reciprocal information gain. According to the study, employing only a few of the features led to greater accuracy than using all of them.

It has also been demonstrated that ensemble methods, such as bagging and boosting, enhance the effectiveness of machine learning algorithms in detecting credit card fraud. For instance, to increase the accuracy of spotting fraudulent transactions, Kim and Kim (2017) suggested a hybrid model that integrated SVM and gradient boosting.
In order to detect credit card fraud, deep learning methods like convolutional neural networks and recurrent neural networks are increasingly in demand. For instance, a CNN-based model that demonstrated excellent accuracy in spotting fraudulent transactions was proposed by Nandakumar and Soman (2017).
The body of literature demonstrates that machine learning algorithms can accurately and quickly identify credit card fraud in real time, with ensemble approaches and deep learning algorithms having the highest levels of accuracy. The performance of machine learning algorithms can be enhanced by data pre-processing and feature selection. Future studies may examine how to enhance the precision and interpretability of credit card fraud detection systems by utilising more sophisticated deep learning algorithms, like RNN and attention-based models.

III. METHODOLOGY

The method this study suggests makes use of the most recent machine learning techniques to identify unusual behaviours, also known as outliers.
First of all, we got our dataset from Kaggle, a website for data analysis that offers datasets. There are 31 columns total in this dataset, 28 of which are labelled v1–v28 to safeguard sensitive information. Time, Amount, and Class are represented by the other columns. Time displays the elapsed time between the first transaction and the next. The amount of money exchanged is called the amount. Class 0 denotes a legitimate transaction, while Class 1 denotes a fraudulent one.

To visually understand the dataset and look for inconsistencies in it, we plot various graphs.
After studying the graphs closely we observed the spending pattern/transaction pattern in the dataset. We used this transaction pattern to identify the frauds among the transactions. The observation of data was Done on the following points: -

1) Number of fraudulent transaction compared to legitimate ones
2) The average time of the day when the transaction is done.
3) The amount that was transacted per transaction.
4) Location of the Transaction.

After collecting all the data we trained the machine according to points specified above. The machine got used to this spending analysis pattern of the user.
Now here is the part where it detects the transaction

1) Consider a user who usually does a transaction between 50$ and 80$ and suddenly one day he does a transaction of about 1000$. Don't you think there is 0.01% of this transaction to be fraud.
2) Consider there is customer who resides in India and does his every transaction from India and suddenly one day does a transaction from China again, there is a few chance of fraud here.
3) Also the average time of the transaction of a customer is 8am to 8pm but suddenly one day he does the transaction on 3am.Don't you think there is very less probability of fraud transaction here.

Now, if a circumstance like this occurs, the computer will verify the customer's authenticity before proceeding with the transaction. The machine will put the transaction on hold until it receives the customer's confirmation before allowing the transaction to be completed. The system will notify the relevant authorities if the user disputes the transaction.
In order to confirm the accuracy and precision of several algorithms, we are comparing their outputs.

IV. IMPLEMENTATION IN REAL WORLD APPLICATION

Implementing a credit card fraud detection system using machine learning in a real-world application involves several challenges, such as dealing with imbalanced datasets, ensuring scalability and real-time performance, and preventing false positives that could lead to legitimate transactions being declined.

Here are some tips for implementing a credit card fraud detection system using machine learning in a real-world application:

1. Define your problem and requirements: Start by defining your problem and requirements. What types of fraud do you want to detect? What performance metrics are important to you? What are your latency requirements?
2. Collect and preprocess data: Collect a dataset of credit card transactions that includes a mix of legitimate and fraudulent transactions. Preprocess the data to clean it, remove missing values, and convert categorical variables into numerical ones.
3. Feature engineering: Create features that can help in detecting fraud. This can involve creating new variables based on existing ones, such as calculating the ratio of the transaction amount to the average transaction amount for that user.
4. Model selection and training: Evaluate different machine learning models and select the one that best suits your requirements. Train the model using your preprocessed dataset.
5. Model evaluation and fine-tuning: Evaluate the performance of your model using various metrics such as accuracy, precision, recall, and F1 score. Fine-tune the model by adjusting the hyperparameters to improve its performance.
6. Deployment and integration: Integrate the model into a real-time system that can analyze credit card transactions in real-time and flag potentially fraudulent ones. Ensure that the system is scalable and can handle large volumes of data.
7. Monitor and update: Monitor the system regularly to ensure that it is detecting fraud accurately and efficiently. Update the model and system as necessary to keep up with new fraud trends and maintain high accuracy.

Overall, implementing a credit card fraud detection system using machine learning in a real-world application requires careful planning, data collection and preprocessing, feature engineering, model selection and training, evaluation and fine-tuning, deployment and integration, and ongoing monitoring and updates.

V.CONCLUSION

Unquestionably, using a credit card fraudulently is a criminal act of dishonesty. The most popular fraud schemes, as well as how to spot them, are listed in this article, which also reviews recent research in the area. Together with the algorithm, pseudocode, description of how it is implemented, and results of experimentation, this work has also provided a detailed explanation of how machine learning can be used to improve fraud detection.

The algorithm does achieve over 99.6% accuracy, however when only a tenth of the data set is considered, its precision is still only 28%. The precision increases to 33% when the system is fed the whole dataset, though. This high degree of accuracy is predicted given the stark disparity between the amount of transactions that are valid and those that are genuine.

REFERENCES

1. Iwasokun GB, Omomule TG, Akinyede RO. Encryption and tokenization-based system for credit card information security. Int J Cyber Sec Digital Forensics. 2018;7(3):283–93.
2. Burkov A. The hundred-page machine learning book. 2019;1:3–5.
3. Maniraj SP, Saini A, Ahmed S, Sarkar D. Credit card fraud detection using machine learning and data science. Int J Eng Res 2019; 8(09).
4. Dornadula VN, Geetha S. Credit card fraud detection using machine learning algorithms. Proc Comput Sci. 2019;165:631–41.
5. Thennakoon, Anuruddha, et al. Real-time credit card fraud detection using machine learning. In: 2019 9th international conference on cloud computing, data science & engineering (Confluence). IEEE; 2019.
6. Robles-Velasco A, Cortés P, Muñuzuri J, Onieva L. Prediction of pipe failures in water supply networks using logistic regression and support vector classification. ReliabEngSystSaf. 2020;196:106754.

7. Liang J, Qin Z, Xiao S, Ou L, Lin X. Efficient and secure decision tree classification for cloud-assisted online diagnosis services. IEEE Trans Dependable Secure Comput. 2019;18(4):1632–44.

8. Ghiasi MM, Zendehboudi S. Application of decision tree-based ensemble learning in the classification of breast cancer. Comput in Biology and Medicine. 2021;128:104089.

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