e-ISSN:2582 - 7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 4, Issue 11, November 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 5.928

9710 583 466 9710 583 466 ijmrset@gmail.com www.ijmrset.com

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 5.928|

| Volume 4, Issue 11, November 2021 |

| DOI:10.15680/IJMRSET.2021.0411027 |

Credit Card Fraud Detection Using Machine Learning

Bhagya Rekha Kalukurthi[1]

R&D Engineer3, Broadcom Inc, India[1]

ABSTRACT: The traditional behaviour of a cardholder is used to teach an HMM at first. An incoming credit card transaction is considered fraudulent if the trained HMM does not accept it with a high enough probability. At the same time, we use an enhancement to ensure that valid transactions are not denied (Hybrid model). In subsequent sections,In further sections we compare different models and methods for fraud detection and prove that why HMM is more preferred method than other methods. Electronic commerce technology has advanced tremendously in recent years, and as a result, the use of credit cards has expanded considerably. As mastercard becomes the most widely used method of payment for both online and offline purchases, incidences of fraud involving it are on the rise. We offer the required theory for detecting fraud in mastercard transaction processing using a Hidden Markov Model in this work (HMM).

KEYWORDS: Machine Learning, Credit Card, Fraud Detection

I. INTRODUCTION

Credit cards are widely accepted around the world today, and organisations of all sizes are storing data in large quantities, in a wide variety of formats, at rapid speeds, and for a high value. This information is acquired from a variety of sources, including user purchase habits and social media followers, likes, and comments. All of this information was analysed and visualised to reveal the hidden data pattern. In the early stages of credit card analysis, general public databases, biometrics, and financial analyses were utilised. Credit cards are an easy and friendly target for fraudsters since a huge number of money may be obtained fast and without riskIn order to perpetrate credit card fraud, fraudsters seek to steal sensitive information such as credit card numbers, bank account numbers, and social security numbers. Because hackers try to make every fraudulent transaction appear legitimate, detecting fraud is challenging. Over 70% of consumers in the United States are vulnerable to fraudulent transactions, according to an increase in credit card transactions. Basically, Because there are usually more valid transactions than fraudulent transactions, the credit card dataset is considerably skewed. That is, it predictions with a high degree of accuracy while avoiding detecting a fraudulent transaction. A better way to address this type of problem is class distribution, which is a sample of minority classes. A class training example can be utilised to boost the algorithm's chances of producing a correct prediction while sampling the minority.

SYSTEM OVERVIEW AND DESIGN

By reducing redundant data and separating data into training and test sets, fraud and genuine transactions can be detected. Credit card fraud is detected in realtime using topological intelligence. Credit card fraud is detected using the Markov Model and Logistics Regression. We create a web application to read input and display output. The goal is to locate each and every fraudulent transaction. Because fraudulent credit card transactions account for a small percentage of total transactions, the system should be able to survive skewed distributions.

II. METHODOLOGY

A literature review was conducted in the first step to learn about credit card fraud detection algorithms and data classification methods, as well as their advantages and disadvantages. The proposed solution will be implemented in a PYTHON simulator with all relevant input and output settings. The performance of the implementation will be thoroughly examined and compared to existing models. Tobegin, we must collect data and examine the dataset's numerous properties. The next step is to undertake data preprocessing, which demonstrates how to handle unbalanced datasets. Finally, we must conduct data analysis to understand more about the dataset's many properties as well as the links between the dataset and other features. It describes the dataset view, which assists in the development of a better project model. Following that, the data should be divided into training and test sets, with the training set being supplied to the machine learning model. Following that, we should load our training data into the logistic regression and hidden

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 5.928|

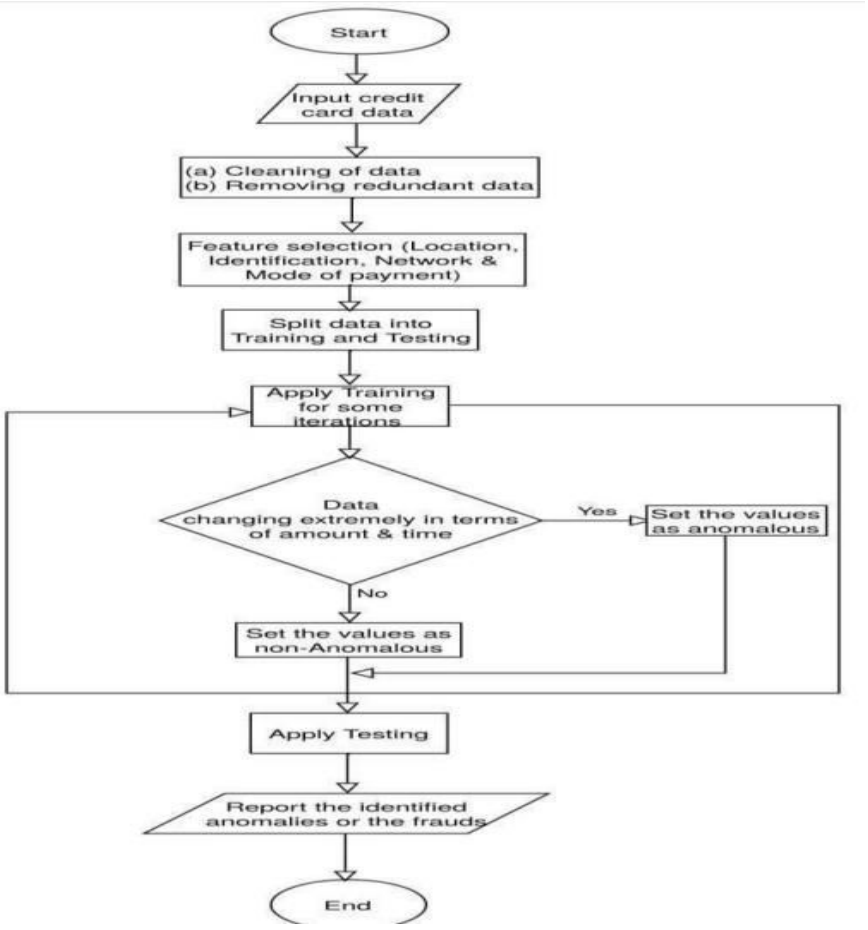
| Volume 4, Issue 11, November 2021 |

| DOI:10.15680/IJMRSET.2021.0411027 |

morkov models. This is the model that will be used in the project because the issue statement is a binary categorization. After training the model, it will be evaluated for performance by comparing it to the test model. This is the project's final phase after reviewing the training model's performance. We demonstrated fraud detection using webpages and graphs, and if the model detects fraud, it willsend an email and SMS to the consumer to confirm the transaction. If it detects a legitimate transaction, it will display a notification on the homepage stating that the transaction is secure and that you may proceed.

III. LITRATURE REVIEW

Unsupervised machine learning based scheme for fraud detection in credit card data: They are attempting to develop communication technologies and ecommerce in this study, which has resulted in credit cards becoming the most frequent method of payment for all types of purchases. As a result, it is critical to ensure that the system is secure and that no fraudulent transactions occur. It appears that fraud transactions in credit card transactions are increasing each year, necessitating study into strategies that might detect and prevent such frauds. This research suggests a strategy that uses a Neural Network (NN)-based unsupervised learning technique to detect fraud in credit card data. They attempted to offer a strategy that surpasses existing clustering methods such as Auto Encoder (AE), Local Outlier Factor (LOF), Isolation Forest (IF), and KMeans. They employed a NN-based fraud detection system that has a 99.87 percent accuracy, compared to 97 percent, 98 percent, and 99.75 percent accuracy for other existing methods. The fundamental idea behind this model is to use neural network methods to train a fraud detection system using a large number of well-known samples that can be used for testing. They introduced a neural network-based fraud detection strategy based on unsupervised machine learning approaches for detecting fraud in credit card data. They evaluated the performance of the proposed work to that of current schemes. They discovered that neural network- based techniques outperform existing systems in these tests. When compared to other existing methodologies, it was discovered that the neural network approach gives 99.98 percent accuracy. They were experimenting with unsupervised learning approaches for fraud detection, presenting their findings in the form of flow charts that depict the detection of fraud in credit card data and the performance of the fraud detection model



| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 5.928|

| Volume 4, Issue 11, November 2021 |

| DOI:10.15680/IJMRSET.2021.0411027 |

This survey aids in the development of a smart agriculture model utilizing robots and various sensors. While not all of the papers cited will be fully implemented, some of the characteristics may be improved. The major goal of this project is to create an integrated automated system that reduces manual field monitoring and provides an advanced approach to seed, plough, water, and cut crops with the least amount of manpower and effort, resulting in an efficient vehicle. To offer an idea for a comprehensive IoT system that spans the entire farming value chain and is based on self-contained IoT modules connected by a large cloud network.The method creates a network of commercial system users, agricultural technology providers, and IT professionals, with the goal of improving agricultural efficient.

- Credit card fraud detection using hidden Markov model: This study focuses on the rapid progress of electronic commerce technology in recent years, as well as the widespread use of credit cards. They claim that as DATASET SHIFT QUANTIFICATION FOR CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING Dept. of CSE, K.S. Institute of Technology Page 7 credit cards have become the most popular method of payment for all purchases, the number of cases of fraud associated with all transactions has increased dramatically. As a result, they're offering the essential theory for detecting fraud in credit card transactions using the Hidden Markov model (HMM). They claim that the credit card holder's normal behavior is used to train the Hidden Markov model. It demonstrates that if the HMM model does not accept the initial credit card transaction with a sufficient exception, the transaction is declared to be false. Using an upgrade to the Hidden Markov model, they were also attempting to verify that real transactions were not rejected. Finally, they analyzed the various strategies for fraud detection in relation to the Hidden Markov model to show that the Hidden Markov model is favored over other methods. This article explains how to create a highly efficient and accurate credit card fraud detection system, which is required by the millions of credit card transactions that occur every day. As a result, a significant quantity of research will be conducted, and a greater number of techniques will be offered to combat credit card fraud transactions.

Dataset shift quantification for credit card fraud detection: This paper discusses how to detect credit card fraud using techniques such as machine learning and data mining. However, it is understandable that fraudster behavior and techniques may evolve with time; this phenomenon of evolving fraudster strategies has been dubbed dataset shift or concept drift in the field of credit card fraud detection. As a result, they're attempting to propose a method for quantifying day-to-day dataset shifts in every credit card transaction made by a cardholder at any store. They tried classifying the days against one another and then evaluating the effectiveness of their classifications. They discovered that the more effective the categorization, the more distinct the behavior between two days is, and that if the classification is inefficient, the behavior is inefficient measure is found to be less it opposes the conditions. As a result, they attempted to obtain the distance matrix that characterizes the dataset shift. They discovered that the dataset shift patterns match calendar events during the time period using agglomerative clustering of the distance matrix. As a result, they added dataset shift knowledge to credit card fraud detection as a new feature, resulting in minor gains in credit card fraud detection. The proposed technique consisted of identifying each day's transaction and quantifying the covariate shift in a temporal dataset, as described above. The strategy, on the other hand, allows us to construct a distance matrix that characterizes the covariate shift between days. They finish this strategy by claiming that they are attempting to show the integrated knowledge of the sort of day previously detected using a Random Forest classifier, which boosts the accuracy recall AUC by minor percentage.

IV. CONCLUSION

Here we presented a neural network based fraud detection scheme for fraud detection in credit card data in which unsupervised machine learning is used. Performance comparison of proposed work with the existing schemes viz., Auto Encoder, Isolation Forest, Local Outlier Factor and K means clustering is done on a credit card dataset. It can be observed that neural network based approach performs better than the existing schemes.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 5.928|

| Volume 4, Issue 11, November 2021 |

| DOI:10.15680/IJMRSET.2021.0411027 |

REFERENCES

1. S.N. Kalid, K. H NG, G. K Tong, K. C Khore., "A Mult iple Classifiers System for Anomaly Detect ion in Credit Card Data With Unbalanced and Overlapped Classes", IEEE Access (2020), Vol. 8, pp.28210-28221
2. S. Makki, Z. A Assaghir,Y. Taher, R. Haque, M. S Hacid, H. Zeineddine, "AnExperimental Study With Imbalanced Classificat ion Approaches for Credit Card Fraud Det ect ion", Special Sect ion On Advanced Software And Data Engineering For Secure Societies, IEEE Access (2019), Vol 7, pp.93010-93022
3. https://www.avenga.com/magazine/anomaly-detection
4. Kola Vasista, "Foreign Capital Issuance and Participants in the Securities Market", International Journal of Research and Analytical Reviews, VOLUME 2, ISSUE 4, OCT. – DEC. 2015
5. Kola Vasista, "A Research Study On Major International Stock Market", International Journal of Research and Analytical Reviews, VOLUME 4, ISSUE 3, JULY – SEPT. 2017
6. Kola Vasista, "A Review On The Various Options Available For Investment", International Journal Of Creative Research Thoughts - IJCRT (IJCRT.ORG), Volume 7, Issue 2, April 2019, ISSN: 2320-2882
7. Satya Nagendra Prasad Poloju, "Data Mining As a Support For Business Intelligence Applications To Big Data", International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, Volume.7, Issue 2, pp.850-854, April 2019, Available at :http://www.ijcrt.org/papers/IJCRT1134576.pdf
8. Satya Nagendra Prasad Poloju, "Big Data Analytics: Data Pre-Processing, Transformation And Curation", International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, Volume.5, Issue 2, pp.835-839, May 2017, Available at :http://www.ijcrt.org/papers/IJCRT1134573.pdf
9. Satya Nagendra Prasad Poloju, "Applications Of Big Data Technology And Cloud Computing In Smart Campus", International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, Volume.1, Issue 2, pp.840-844, September 2013, Available at :http://www.ijcrt.org/papers/IJCRT1134574.pdf
10. Satya Nagendra Prasad Poloju, "Relevant Technologies of Cloud Computing System", International Journal of Engineering Research and Applications, ISSN: 2248-9622, Vol. 4, Issue 4, (Version-3) April 2014, pp. 74-78, Available at: https://www.ijera.com/pages/v4no4(v3).html
11. Adithya Vuppula, "Classification and Visualization of Data Mining Model", "International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering", Vol. 4, Issue 8, August 2015
12. Adithya Vuppula, "Data Mining: Convergence of Three Technologies", "International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering", Vol. 7, Issue 6, June 2018
13. Adithya Vuppula, "Initiatives of 5G Vision and 5G Standardization", International Journal of Innovative Research in Computer and Communication Engineering", Vol. 6, Issue 2, February 2018
14. Adithya Vuppula, "Security Mechanisms for IOT Services and Differences between IOT and Traditional Networks", "International Journal of Innovative Research in Science, Engineering and Technology", Vol. 6, Issue 2, February 2017
15. Adithya Vuppula, "Communication and Protocols towards IOT Based Security", "International Journal of Innovative Research in Science, Engineering and Technology", Vol. 3, Issue 10, October 2014

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 5.928

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

9710 583 466 | 9710 583 466 | ijmrset@gmail.com

www.ijmrset.com