International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Credit Card Fraud Detection and Prevention using Face Authentication

Dhanmiga S[1], Ragapriya M[2], Sowmyaa M S[3], Vijayakumar A[4]

[1]Department of Information Security and Digital Forensics, Dr. M.G.R. Educational and Research Institute, Chennai - 600095, Tamil Nadu, India.
dhanmigasrgd@gmail.com
[2]Department of Information Security and Digital Forensics, Dr. M.G.R. Educational and Research Institute, Chennai - 600095, Tamil Nadu, India.
raga.m2909@gmail.com
[3]Department of Information Security and Digital Forensics, Dr. M.G.R. Educational and Research Institute, Chennai - 600095, Tamil Nadu, India.
sowmyaasrinivasan@gmail.com
[4]Center of Excellence in Digital Forensics, Chennai - 600096, Tamil Nadu, India.
vijisbi@gmail.com

ABSTRACT

Credit card fraud is currently the most common problem in the modern world. This is because internet transactions and e-commerce sites are on the rise. Credit card fraud occurs when a credit card is stolen and used for unauthorized reasons, or when a fraudster exploits the credit card information for his own interests. In today's environment, we're dealing with a bunch of credit or debit card issues. Despite the fact that the financial industry is filled with criminal activity, credit card theft is the most common and concerning to online clients. A credit card fraud detection system was implemented to detect fraudulent actions. The major goal of this initiative is to focus on fraud transaction prevention. The ML algorithm determines whether or not the transaction is valid. This is accomplished by using the provided dataset to learn. Random forest is the algorithm that were used. We built a website where you can register your details for login and submit card information for registration, for further security. The algorithms will then examine and identify the transactions. To prevent fraud by exploiting the authorized person's photo, the facial recognition detection uses Haar features and a cascade classifier. It would be ideal if the system could discriminate between known and unknown faces, allowing only authorized individuals access.

INTRODUCTION

Online buying is becoming increasingly popular. One-tenth of the world's population shops online from 2005. Germany and the United Kingdom have the most online shopping, and credit cards are the most preferred payment method (59 percent). Barclaycard, the largest financial services business in the United Kingdom, was said to be processing around 350 million transactions each year. Wal-Mart, for example, often handles a significantly greater volume of credit card transactions, including both online and in-store purchases. As the number of people who use credit cards grows around the world, so do the potential for attackers to obtain credit card information and commit fraud.Purchases made with a credit card can be divided into two categories:

1. Physical cards.
2. Virtual cards.

In a physical based credit card transaction, the cardholder physically presents his card to the merchant for payment. An attacker must steal the credit or debit card in order to carry any fraudulent activity in this type of purchase. If the cardholders does not notice the theft of the card, the credit card firm may suffer a significant financial loss. Only a few key details about a credit card information (card number, expiration date, and security code) are necessary to complete the transaction. Typically, such purchases are made over the phone or on the Internet. A scammer only has to have the card details to conduct fraud in such types of transactions. Almost all of the times, the actual cardholder is unaware that his or her card information has been seen or stolen by someone else. But one way to identify this fraudulent activity is to examine each card's spending history and look for any anomalies from "normal" purchasing patterns. Detection of fraud based on an examination of a cardholder's existing purchase data is a promising method of reducing the percentage of successful fraudulent activity. Each card holder can really be characterized by a sequence of pattern including information about the average purchase category, the duration since the last transaction, the amount of cash spent, and so on, because persons have distinct behavioural profiles. Since people have different behavioural profiles, each card holder can be identified by a series of patterns, such as the average purchase category, the time from the last transaction activity, the quantity of money spent, and so on. Deviation from these patterns poses a risk to the system.

In recent years, several strategies for detecting credit card fraud have been proposed.

RELATED WORK

1.Andhavarapu Bhanusri et al., 2020; have outlined the theory of credit card frauds and implemented various supervised ML algorithms on an unbalanced dataset, including random forest, logistic regression, naive bayes, and ensemble classifiers utilising the boosting technique. The study's result reveals that using supervised techniques to train and assess the best classifier delivers a better answer. (1)

2.Munira Ansari et al.,2021; utilised publicly available card information to discover fraudulent transactions and calculate model advantages using ML approach. The study's final findings reveal that the simple majority mode of voting achieves higher precision rates in identifying fraudulent transactions. (2)

3.Lakshmi SVSS, 2018; compares the features of random forest, decision tree, and logistic regression algorithms for detecting and preventing fraudulent activity. And those are the 3 major strategies used for the databases, and the work is done in the R programming language. The accuracy, sensitivity, error rate, and specificity of the approaches are all examined for different factors.The accuracy of the random forest, logistic regression analysis, and decision tree classifiers is 95.5, 90.0, and 94.3, respectively. The Random Forest outperforms the logistic regression analysis and decision tree procedures, according to the final report. (3)

4.Senthamizh Selvi.R, 2019; employed the Haar-Cascade classifier technique to recognise human faces in OpenCV, which was arranged using the Python computer language. Only 200 features from 6,000 are used in this system, which results in a recognition rate of 85 to 95 percent. When compared to conventional techniques, the Haar cascade classifiers algorithm provides a high detection accuracy also with fluctuating facial reactions and low false positive characteristics. (4)

5.Chang et al.Proposed a replacement learning methodology in the field of novel intrusion detection system with back propagation neural networks(BPN) using sample query and attribute query. A combination of classification with a query based learning methodology and data reduction is used in this paper as it consumes less time. Experiment has revealed that the proposed method takes 1447 seconds to be trained on the other hand the BPN takes 21746 seconds for training future works is to explore in the field of BPN to enhance learning techniques for real world application.(5)

6.Srivastava et al. has worked on a model to exhibit the order of credit card transaction process and gives the experimental result which explains how effective the system is and the use of learning the spending habit of the card holders. The systems accuracy is around 80% over a large difference in the input data. The system can be used to deal large amount of transaction(5)

7.Subashini and Chitra has constructed d classifier models named C5.0 ,CART from the five classification methods namely the decision tree, SVMs using SMO algorithm with kernels of polynomial functions, logistic regression and bayes net to detect fraud using the credit card fraud data set in the banking sector. The genuine user is marked good and the fraud user is marked bad. The logistic regression method gives the success rate of 73.1% and the highest success rate 74.1% is produced by the CART.(5)

8.Mishra et al. has provided sufficient theory for credit card transaction fraud detection using hidden Markov model the and how the model is used in fraud detection. If HMM is not accepting a transaction then it is high likely considered as fraudulent. Meanwhile HMM makes sure that legitimate transaction is not rejected. HMM states are based on the types of items and observation symbol is based on various the ranges of transactions amount. Also a technique to find the cardholders pending habit. It is suggested to apply this data to decide the observation symbol. More than 85% of transaction is genuine in the proposed system and also it has low rate of false alarms around 8%.(5)

EXISTING SYSTEM

In the current approach, fraud is recognized after the fraud has been committed, that is, after the holder has filed a complaint. As a result, the cardholder experienced a lot of difficulties before the investigation was completed. Also, because all transactions are kept in a log, we need to save a lot of data, and because a lot of purchases are made online these days, we don't know who is using the card online, so we only collect their IP address for verification. As a result, cybercrime assistance is required to investigate the scam. To eliminate all of the abovementioned drawbacks, we suggest a system for quickly detecting fraud. There is no mechanism to prevent fraudulent transactions.

PROPOSED SYSTEM

This suggested method is designed to evaluate whether a certain transaction is genuine or not. The random forest ML algorithm is used to detect the fraudulent transactions. The efficiency, accuracy, recall, and F1-score of the two techniques are used to compare their outcomes. The confusion matrix is used to plot the ROC curve. When the Random Forest method is evaluated, this algorithm is determined to be the best for detecting fraud.The Hidden Markov model is based on the user's spending habits. User is assigned to one of three categories: low, medium, or high. Facial recognition is now available in Python and OpenCV. This adds another layer of security to the user's credentials. The detection of fraudulent card use is more faster than the current system. Every transaction includes face authentication for the original cardholder. As a result, the log is kept for fraud detection. The log, which will be kept, will also serve as proof of the transaction to the bank. Using this method, we may get a most precise detection. This cuts down on a bank employee's boring duties.

ARCHITECTURE DIAGRAM

First, the user registers their credentials and also its detect the face for face registration. It collects the banking details of the user and these are stored in encrypted form in the database. When the transaction begins its goes under ML algorithm and detects whether the transaction is fraudulent or legitimate

transaction. Fraud transactions are stored in separate databases and finally it asks for face authentication, if the user passes, the transaction takes place or the transaction fails.

MACHINE LEARNING

Machine learning is the field of study which gives computers the ability to learn without being entirely programmed. Machine learning is one of the most interesting technologies in existence. As the name suggests it makes the computer more similar to humans that is it provides the ability to learn. Today machine learning is used all around the world. Machine learning has a wide range of applications, including photo tagging, search engine, spam detection, weather prediction, speech and facial recognition software, and customer service, among others.

Machine learning is divided into three main divisions they are

1. supervised learning
2. unsupervised learning
3. reinforcement learning

Based on the nature of the signal or response of the system. Datas which are too complex to be solved with conventional programming can be e used to feed/train machine learning algorithms.

RANDOM FOREST ALGORITHM

The type of supervised learning that Random Forest belongs to is supervised machine learning. It is simple and adaptable to use, resulting in increased and improved accuracy. Random forest has a wide range of applications in classification and regression. It achieves the best results by combining decision trees and reducing error due to biased and variation.

Train the data set and apply the classification technique based on it in the proposed model. Outliers & null values were eliminated from the dataset after it was collected. The data preprocessing was completed as a result of this. The dataset is separated into two sections after pre-processing, one for train and the other for testing. The dataset is divided into three parts: 33 percent for training and 77 percent for testing. The accuracy % was anticipated by analysing the method with training and testing sets. The Random Forest Algorithm is tested using Google Colab and the Python programming language.

FACE AUTHENTICATION

Facial recognition technology has a significant and positive impact on businesses and society. Facial recognition software is now being used to protect personal data from cyber-attacks, reduce mistaken arrests, and even detect patients with hereditary diseases. Face detection and image or video recognition are two popular biometrics research topics. We use

Python, OpenCV, Haar Cascade classifier to create a webcam real-time facial recognition system and develop an algorithm.

Haar Cascade Classifier Detection is a robust face detection technology that has been around for a long time. It has existed for a long time, long before Machine Learning got popular. Haar Features were employed to recognize faces as well as eyes, noses, lips, licence number plates, and other features. OpenCV methods can be used to access the models. Facial recognition is a powerful vision technology. Face recognition/detection is the process of locating and seeing human faces in digital images. Face detection technology is useful in a variety of industries, including marketing and security.

CASCADE CLASSIFIERS AND HAAR FEATURES

Haar Cascade Classifiers are used to detect objects. It is a machine learning approach in which we use a large number of photos to train a cascade function. There are two types of images: positive images with the goal object & negative images without the target object.

Cascade classifiers come in a various types of shapes and sizes, depending on the target item.

To recognise the human face as the target image in the project, we will utilise a classifier that analyses the human face. The goal of the Haar Feature Selection approach is to extract human face traits. Haar features work in a similar way to convolution kernels. Different combinations of white & black squares make up these features. We find the total of pixel under white & black squares in each feature computation. OpenCV includes pre-trained models using Haar features & cascade classifiers for Haar-cascade detection. These models can be found in the OpenCV installation directory.

RESULTS

The random forest method has a higher accuracy of 99.6 percent in detecting anomalies in credit card transactions than the logistic regression approach, which has a 92.6 percent accuracy. The results clearly suggest that the Random forest method outperformed the logistic regression approach in terms of significance. The random forest algorithm has a precision of 23.35 percent, recall of 73.40 percent, accuracy of 99.6 percent, and an F1-score of 35.43 percent. Finally, the proposed classifier was shown to be 99.6% accurate. As a result, the model can effectively detect fraud in credit card transactions. The mean, standard error, and standard deviation mean of random forest and logistic regression algorithm based credit card payment fraud detection, which demonstrates that the logistic regression has a mean accuracy of 97.25 percent, Standard deviation of .15128 with a sample size of N=10 and an efficiency mean of 99.3 percent for the Random Forest. With a standard deviation of.26717 for a sample size of N=10, the statistical significance of the

Random Forest Algorithm is excellent. The mean, standard deviation, and significant difference of random forest and logistic regression algorithm based credit card fraud detection, which reveals a significant difference between the two groups since p0.018 (Independent Sample T Test)

CONCLUSION

For all of the reasons stated, this is the ideal location for a card holder to handle their cards in a safe manner and handle transactions regularly. Because of the efficient and user-friendly security component given on this site, fraud may be easily discovered and cardholders can experience free to access their credit card. This project's procedure is thoroughly defined and examined, allowing visitors to use the site gradually and without hesitation.

REFERENCES

1.Andhavarapu Bhanusri, K.RatnaSreeValli , P.Jyothi , G.Varun Sai , R.Rohith Sai Subash (2021). Credit card fraud detection using Machine learning algorithms. Journal of Research in Humanities and Social Science Volume 8 ~ Issue 2 (2020)pp.: 04-11 ISSN(Online):23219467

2.Munira Ansari, Hashim Malik, SiddheshJadhav, Zaiyyan Khan (2021). Credit Card Fraud Detection. International Journal of Engineering Research & Technology (IJERT), Special Issue – 2021, ISSN: 2278-0181

3.Lakshmi S V S S, SelvaniDeepthiKavila (2018). Machine Learning For Credit Card Fraud Detection System. International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 24 (2018) pp. 16819-16824

4.Senthamizh Selvi.R , D.Sivakumar, Sandhya.J.S , Siva Sowmiya.S, Ramya.S , KanagaSuba Raja.S(2019). Face Recognition Using Haar - Cascade Classifier for Criminal Identification. International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6S5, April 2019

P. Jayant, Vaishali, D. Sharma, Survey on Credit Card Fraud Detection Techniques, International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 3 Issue 3, March – 2014

Sadineni, Praveen Kumar. 2020. "Detection of Fraudulent Transactions in Credit Card Using Machine Learning Algorithms." 2020 Fourth International Conference on I-SMAC (IoT in Social,Mobile, Analytics and Cloud) (I-SMAC).

Gawas, Bhakti G. 2019. "Fraud Detection in Mobile Payment System Using Machine Learning: AComprehensive Survey." International Journal for Research in Applied Science and Engineering Technology (IJRASET)ISSN: 2321-9653;Volume 7 Issue IV, Apr 2019

8.Sailusha, Ruttala; Gnaneswar, V.; Ramesh, R.; Rao, G. Ramakoteswara (2020). [IEEE 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)

Madurai, India (2020.5.13-2020.5.15)] 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS) - Credit Card Fraud Detection Using Machine Learning.

9. M. Khan, S. Chakraborty, R. Astya and S. Khepra, "Face Detection and Recognition Using OpenCV," 2019 International Conference on Computing, Communication, and Intelligent

Systems (ICCCIS), 2019, pp. 116-119International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056.

10.Dileep, M. R., Navaneeth, A. V., & Abhishek, M. (2021). A Novel Approach for Credit Card Fraud Detection using Decision Tree and Random Forest Algorithms. 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV).

11.Tanouz, D., Subramanian, R. R., Eswar, D., Reddy, G. V. P., Kumar, A. R., &Praneeth, C. V. N. M. (2021). Credit Card Fraud Detection Using Machine Learning. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS).