IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

ONLINE TRANSACTION FRAUD DETECTION SYSTEM

[1]Vijith P R, [2]Vidya Vinayan, [3]Vysakh K M, [4] Aleena Benny, [5]Rasneena V S

[1]Assistant Professor, Department of Electronics and Communication Engineering, Universal Engineering College Vallivattom, Thrissur, India.

[2345]B.Tech Student, Department of Electronics and Communication Engineering, Universal Engineering College Vallivattom, Thrissur, India.

Abstract: The growth in internet and e-commerce appears to involve the use of online credit/debit card transactions. The increase in the use of credit / debit cards is causing an increase in fraud. The frauds can be detected through various approaches, yet they lag in their accuracy and its own specific drawbacks. In this work, the behaviour-based approach to classification is used to improve its accuracy. If there are any changes in the conduct of the transaction, the frauds are predicted and taken for further process. Due to large amount of data credit / debit card fraud detection problem is rectified by the proposed method. The credit card frauds can be detected by evaluating the CC purchasing patterns using the historical data in order to detect the frauds. This data evaluation can help the banks or other organizations offering credit cards to minimize their losses due to the credit card frauds. The historical data evaluation with the current purchasing patterns requires the statistical modelling, which can automatically evaluate the fraudulent patterns and alarm the banks about the transactions. This helps the banks for early detection of the frauds, where they can easily eliminate the CC frauds by declining the suspected transactions

Keywords: Credit Card, Machine Learning, Cyber Security

I. INTRODUCTION

Fraud, such as phone fraud, insurance fraud and credit card fraud, causes severe problems for government and business. However, detecting such a fraud has always been challenging. With the rapid development of the e-commerce and e-payment, the problem of online transaction fraud has become increasingly prominent. Compared with traditional areas, online transaction is facing a considerably larger volume of fund transfer. Misrepresentation alludes to getting products/administrations and cash by illicit way. Extortion manages occasions which include criminal intentions that, for the most part, are hard to recognize. Charge cards are a standout amongst the most mainstream goal of extortion yet not alone. Charge card extortion, a colossal term for burglary and misrepresentation submitted or any comparable instalment instrument as a fake asset of assets in an exchange. Charge card misrepresentation has been growing issue in the Master card business. Recognizing Visa misrepresentation is a troublesome undertaking when utilizing typical process, so the advancement of the charge card extortion identification models has happened to significance whether in the scholastic or business associations right now. Besides, part of extortion has been changed abruptly amid the most recent couple of decades alongside headway of technologies. Card fraud begins either with the theft of the physical card or with the important data associated with the account, including the card account number or other information that necessarily be available to a merchant during a permissible transaction. Card numbers, generally the Primary Account Number (PAN) are often reprinted on the card, and a magnetic stripe on the back contains the data in machine-readable format. It contains the following Fields: Name of card holder Card number Expiration date Verification/CVV code Type of card There are more methods to commit credit card fraud. The proposed techniques are used in this paper, for detecting the frauds in the credit card system. we can make use of different machine learning algorithms such as Decision Trees, Random Forest, to determine which algorithm suits best and can be adapted by credit card merchants for identifying fraud transactions.

II. Related works

Here are some papers from which the proposed system takes references from.

In this paper [1], the authors propose a Specific algorithm based on artificial intelligence and neural networks are also being proposed and implemented to predict the credit card frauds with increased accuracy. The distribution of the datasets used for fraud detection is highly imbalanced. So, to overcome this obstacle, under- sampling and oversampling techniques are being designed to obtain comparatively balanced data. Data mining techniques are also being implemented in order to create a more efficient Fraud Detection System [9]. Another important area of development is the emergence of new hybrid models. These are derived from pre-existing supervised as well as unsupervised machine learning techniques. Hybrid Models may be able to produce a more accurate result as they enrapture the capabilities of both supervised as well as unsupervised machine learning.

The paper [2] proposes performance of all machine learning datasets is hindered due to the skewness of available data sets which are usually unbalanced. To overcome this problem, the unbalanced datasets are to be converted to balanced ones. This can be done by mainly two ways which are Intrinsic Method and Network based Method. In Intrinsic Feature Method, a pattern in the customer Activity is observed whereas in Network -based features Method, the network of users and the card merchants is exploited. These techniques may significantly improve the functioning of certain Models as they work on a more Balanced Dataset.

The paper [3] proposes a system A comprehensive understanding of fraud detection technologies can be helpful for us to solve the problem of credit card fraud. The work in provides a comprehensive discussion on the challenges and problems of fraud detection research. Mohammad et.al. review the most popular types of credit card fraud and the existing nature-inspired detection methods that are used in detection methods. Basically, there are two types of credit card fraud: application fraud and behaviour fraud. Application fraud is that criminals get new credit cards from issuing companies by forging false information or using other legitimate cardholders' information. Behaviour fraud is that criminals steal the account and password of a card from the genuine cardholder and use them to spend.

The paper [4] Introduces the Cashless payments becomes very convenient to the user to make payment without carrying cash in hand. High possibilities are there to stole individual information moving towards digital payments. Using imbalanced data set checked with different supervised machine learning algorithms, decision tree is the best suitable algorithm for detecting the fraud.

The paper [5] Presents the growth in the credit card transactions has led to rise in the fraudulent activities. Customer or Customer card details are mandatory to do the transactions. The merchant also can't ably find whether authentic cardholder or not. Random Forest is the proposed model to improve the sensitivity, accuracy, precision and specificity in fraud detection

III. Methodology

Data collected from the credit card users having 23 attributes like, Amount of the given credit, Gender, Age, History of past payment etc. These attributes are fed into a machine learning algorithm (Random Forest). In some cases, we cannot get all the details. In such cases, we need to skip that record and gather details only if there are all these 23 attributes present. Out of the whole data collected, we are taking 4/5 of the data for training the model and the rest 1/5 of the data is used for testing purposes to confirm its accuracy. It is further gone through the processing step where we can classify the details in its attributes and have an extra 1 group where we have the Label where the status of that patient is given. The training set data is moved through a Machine Learning Algorithm after the processing step is completed to get a 'Trained Model'. This Trained Model is now able to analyse the credit card status by providing the 23 attributes

IV. Models used

4.1 Decision Tree

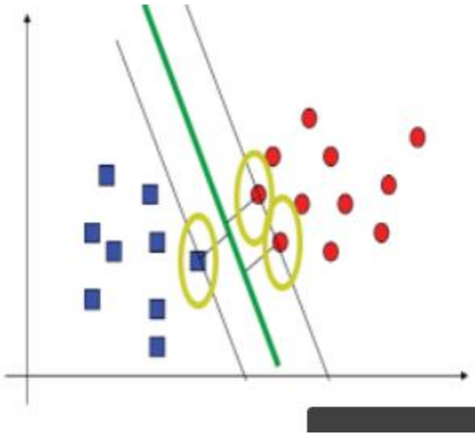
This is one of the most widely used predictive modelling approaches. As per the name of the model, this is built in the form of a tree like structure. This model maybe used in case of a multi-dimensional analysis where there are multiple classes present. The past data also known as the past vector is used to create a model that can be used to predict the value of the output based on the input being provided. There are multiple nodes in a tree and each node corresponds to one or the other vector. The tree terminates at a leaf node where each such node represents a possible outcome or output.

4.2 Random Forest

This model is basically an ensemble classifier, i.e., a combining classifier that uses and combines many decision tree classifiers. The main agenda behind using multiple trees is to be able to train the trees enough, such that, contribution from each of them comes in the form of a model. After the generation of the tree, the output is combined through majority. It uses multiple decision trees so that, the dependence of each of them is on a particular dataset possessing similar distribution throughout the tree. This particular model has the quality of efficiently balancing errors in a class population of unbalanced data sets. It can be used to solve both classification as well as regression problems.

4.3 Support Vector Machine (SVM)

VM may also be an algorithm for supervised machine learning that can be used for both classification and regression problems. It's often used in classification issues, however. In this algorithm, we plot each data item in a "N" dimensional space to some degree, with the price of each function being the price of a chosen coordinate showing the hyper-plane differentiating the 2 groups from the hyper-plane. A supervised machine learning algorithm may also be a support vector machine (SVM), which can be used for both classifications. During this algorithm, we take the shopping data item of each consumer to some degree in a 'n' dimensional space (where n is the number of features you have) with the price of each function being the price of a chosen coordinate.



100%.

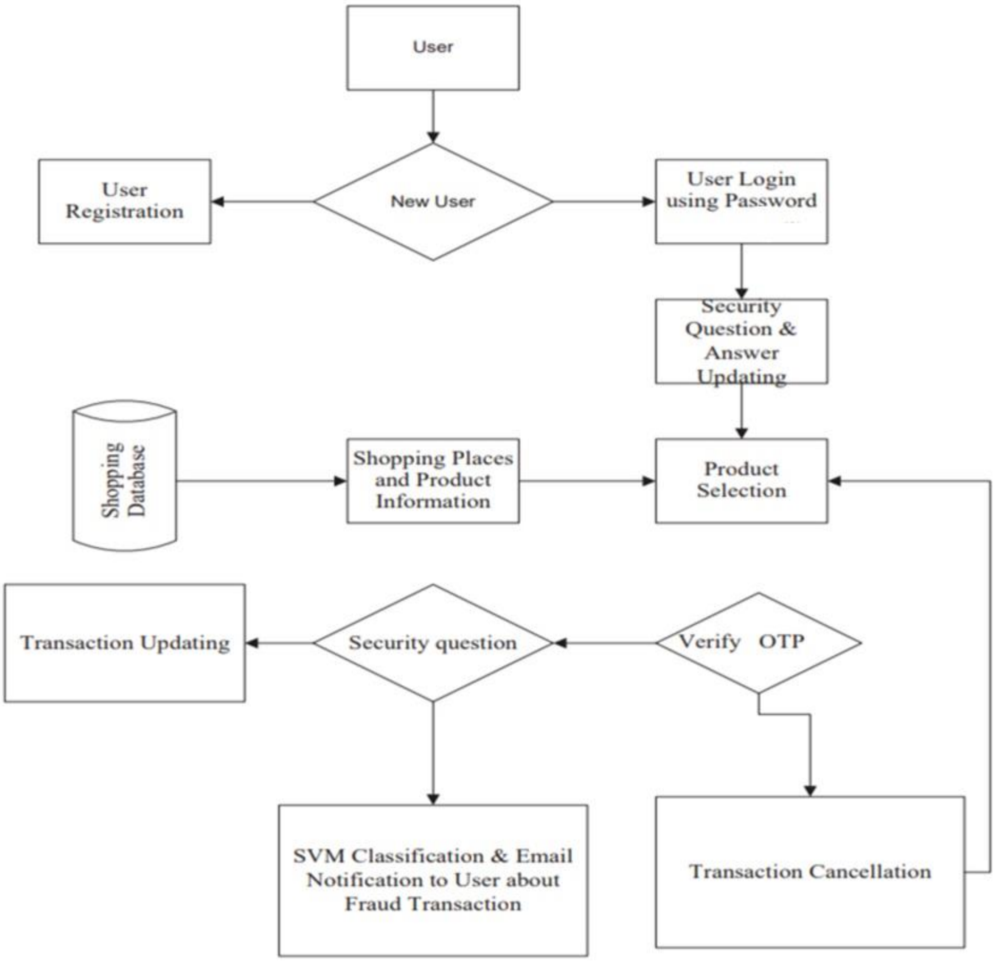
[redacted]

SVM classification

V. SYSTEM OVERVIEW

5.1 Architecture

The basic system architecture of system is shown in figure below


User
User Registration
New User
User Login using Password
Security Question & Answer Updating
Shopping Database
Shopping Places and Product Information
Product Selection
Transaction Updating
Security question
Verify OTP
SVM Classification & Email Notification to User about Fraud Transaction
Transaction Cancellation

system architecture

5.2 Architecture Descriptions

5.2.1 New User Registration with fingerprint image

Here, we develop an integrated online shopping module for user. The new user registration form consists of username, mobile number, address, city, E-Mail ID, etc. The secondary security clearance level which deals with the fingerprint scanning. It is a standalone fingerprint identification device with many excellent features such as high identification performance. user's fingerprints collected and unique hash code value generated using one-way hash function and it's stored into the bank server's database**.**

5.2.2 User Login authentication using fingerprint Image

The user is requested to enter their login details such as username, password.

5.2.3 Security question Customization and Verification module

To enrol for the 'Security Questions' verification, the Login user selects several questions and supplies confidential answers that only the user knows. The banking server system provides a set of default questions to users. Security questions are stored into database by securely on a user and can only be answered by you during online shopping's checkout verification. There is a provision available for individual user to read or modify questions and/or answers.

5.2.4 One-Time Password (OTP) Verification

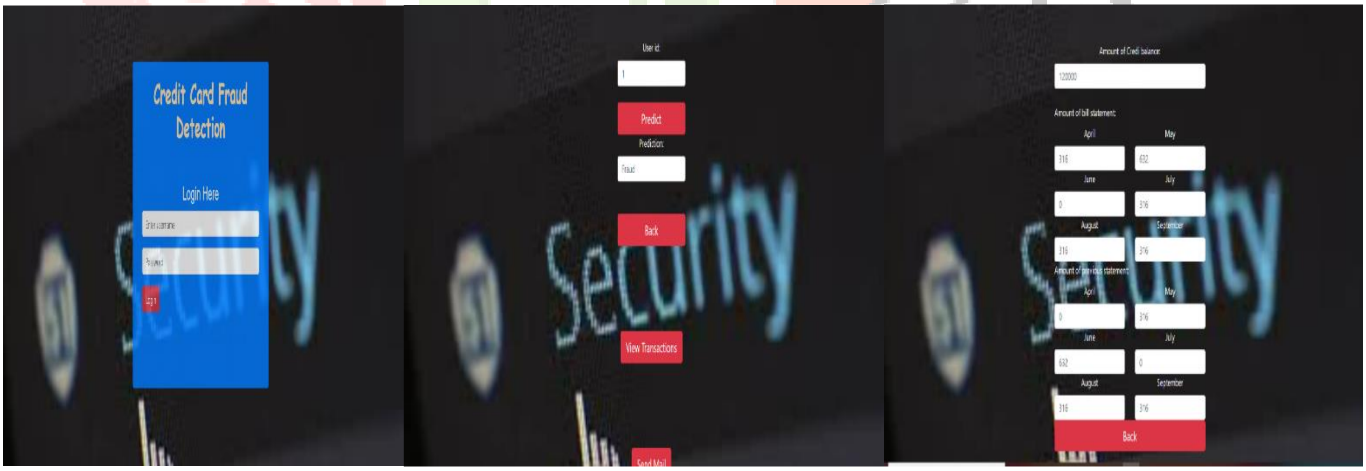
The aim of authentication is to prove that the shopping user is that the authentic user or fraud user for suspicious shopping transaction identification. One-Time Passwords are utilized as a supplemental think about multi-factor sanction/authentication applications. they're only valid for precisely one sanction or authentication request. To evade password lists, a convenient thanks to provide the user with an OTP is to send it through email. the e-mail id of the user must be registered for the accommodation that gives email OTPs for authentication or sanction. OTPs are quite popular as a supplemental sanction or authentication think about web predicated accommodations. These passwords are often utilized to authenticate a user, i.e., the user needs a legitimate OTP to prove his identity to authenticate in to an application or to access the network. Email OTPs are withal utilized for account verification, e. g., Google Mail. After successful OTP verification process, Online shopping's banking server will through Security questions which may be used Asan additional level of security when your clients contact you so as to verify their identify. User can then prompt them for the solution to the question they chose during the registration phase.

5.2.5 Online Fraud Transaction Classification using SVM

In order to predict or classify patterns into two categorises may be a classifier: fraudulent or non-fraudulent. During this algorithm, we take the shopping data item of each user to some extent in a 'n' dimensional space with the value of each function being the value of a particular coordinate. Within the linearly separable case, the hyper-plane differentiating, there is one or more hyperplane that separates the two classes represented by training data with 100 percent. If a shopping transaction is classified as fraud, payment deduction and checkout process will be cancelled by server and fraudulent user's transaction details are sent to that specific user thru email else shopping is completed and transaction stored into database.

VI. RESULT

The system presents classification of online credit / opened credit the challenges faced by cardholder also because the card issuer, verity of fraud implemented by the persons who commit that fraud, some latest news regarding master card fraudster and provide some prevention techniques that need to be followed by the cardholder against the fraudulent activity. Master card s has become the most common mode of payment in recent times and if master card transactions increase, frauds also do so. The great news is that in recent years, technology for preventing online fraud has also improved and reducing computer costs helps to implement complex systems that can analyse fraudulent behaviour within a fraction of a second duration. Hereby, it concludes that each system faces its own problems while dealing with dataset description. Even in early system of random forest model which deals with pre-processing datasets faces low accuracy rate. So, the proposed system come with the Random Forest model of real databases which helps in acquiring the maximum of 83.34% accuracy.



User id: 1

Prediction: Fraud

Amount of Credi balance: 120000

Amount of bill statement:

April: 316
May: 632
June: 0
July: 316
August: 316
September: 316

Amount of previous statement:

April: 0
May: 316
June: 632
July: 0
August: 316
September: 316

credit card fraud detection

VII. CONCLUSION

A centralized fraud management platform is the need of the hour to facilitate a shared fraud prevention approach. Proposal is to develop an operating framework and model to bring organizations across the globe to leverage this framework through adhering to its standards and hence share and leverage fraud patterns to proactively alert and be alerted on fraudulent transactions there by building a strong layer of protection for their applications.

REFERENCES

[1]. S. Mittal and S. Tyagi, "Performance Evaluation of Machine Learning Algorithms for Credit Card Fraud Detection", 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence),2019.

[2]. S. Dhankhad, E. Mohammed and B. Far, ’ Supervised Machine Learning Algorithms for Credit Card Fraudulent Transaction Detection: A Comparative Study,” 2018 IEEE International Conference on Information Reuse and Integration (IRI), Salt Lake City, UT, 2018, pp. 122-125.

[3]. Behdad, M., Barone, L., Bennamoun, M., and French, T. (2012). Nature inspired techniques in the context of fraud detection. IEEE Transactions on Systems Man and Cybernetics Part C, 42(6), 1273-1290.

[4]. Samidha Khatri.,Aishwarya Arora., and Arun Prakash Agarwal., “Supervised Machine Learining Algorithms for Credit Card Fraud Detection: A Comparion”.,IEEE.,2020

[5]. Devi Meenakshi B.,Janani B.,Gayathri S., and Indira N., “Credit Card Fraud Detection Using Random Forest”., International Research Journal of Engineering and Technology (IR JET). ISSN:2395- 0056,Volume 06,Issue 03,March 2019