IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 6, June 2021

DOI: 10.17148/IARJSET.2021.8652

Credit Card Fraud Detection Using Machine Learning Algorithms

Dr .Dinesh. D .Patil, Dr. Priti Subramanium, Evangel Denis Rodrigues

Head & Associate Professor, Computer Science, SSGBCOET, Bhusawal, India.

Assistant Professor, Computer Science, SSGBCOET, Bhusawal, India.

M. Tech Student, Computer Science, SSGBCOET, Bhusawal, India.

Abstract: The digital world is gaining popularity because of seamless, easy, convenience and wide use of e-commerce. It helps pay bills quickly at just a click. People choose online payment and e-shopping as it saves time, it is convenient, etc. As the result of huge amount of e-commerce use, there is a vast increase in credit card fraud also. Fraudsters try to misuse the card and transparency of online payments and cheat innocent people. Hence, to avoid these frauds from happening this study is necessary. The main aim is to secure credit card transactions; so people can use e-banking safely and easily without fear. To detect the credit card fraud there are various techniques which are based on Deep learning, Logistic Regression, Naive Bayesian, Support Vector Machine (SVM), Neural Network, Artificial Immune System, K Nearest Neighbour, Data Mining, Decision Tree, Fuzzy logic based System, Genetic Algorithm etc.[1]

Keywords: Frauds, Digital, E-banking, Machine Learning.

I. INTRODUCTION

Fraud credit card transaction is the unauthorized use of someone’s account without the owner being aware of it. Prevention measures need to be taken against such fraudulent practices by analysing and studying these fraud transactions to avoid similar situations in the upcoming transactions.

Briefly, Credit Card transaction Frauds can be explained as a scenario in which a fraudster use a credit card of another person for personal means without seeking permission or authorization of the owner of the credit card and the credit card issuing officials or institutions are unknown of the fraud that will be happening. In order to detect fraud, there is a need to monitor the activities of the users to avoid abnormal behaviour which include intrusion, Fraud and defaulting. Machine learning and data science are the communities that focus on problems like these since the solution has greater possibility and feasibility for being automated .Usually the legitimate transactions are more than the fraudulent ones. Furthermore, the statistical properties of transaction arrangement changes frequently over a time period .In the real world the execution of credit card fraud detection faces many more obstacles. However, instances, automatic tools scan the vast stream of transaction requests that tells which transaction to legitimise. Machine learning algorithms are used to inspect all the ligament transaction and list out transaction which are suspicious. Then the reported suspicious transaction one investigated by professionals. They contact the cardholder to identify whether the transaction was authentic or fraudulent. The automated systems are the updated by the investigators as a feedback which helps the system to further train and improve the effectiveness of the fraud detection over time .Fraud detection needs to be constantly updated to defend against emerging fraudulent strategies by the criminals.

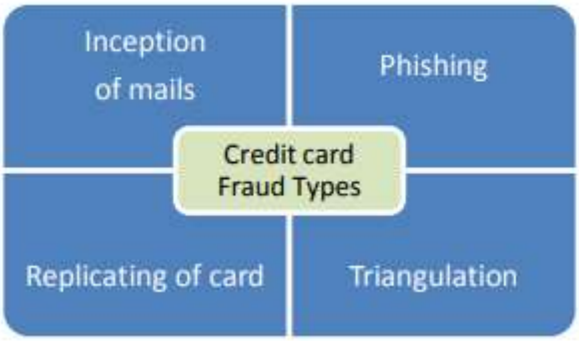

Fig 1: Credit Card Fraud Types

DOI: 10.17148/IARJSET.2021.8652

II. LITERATURE SURVEY

The illegal use of credit card or its information without the knowledge of the owner is referred to as credit card fraud. Different credit card fraud tricks belong mainly to two groups of application and behavioural fraud. Application fraud takes place when, fraudsters apply new cards from bank or issuing companies using false or other's information. Multiple applications may be submitted by one user with one set of user details (called duplication fraud) or different user with identical details (called identity fraud). Behavioural fraud, on the other hand, has four principal types: stolen/lost card, mail theft, counterfeit card and „card holder not present" fraud. Stolen/lost card fraud occurs when fraudsters steal a credit card or get access to a lost card. Mail theft fraud occurs when the fraudster get a credit card in mail or personal information from bank before reaching to actual cardholder[3]. In both counterfeit and „card holder not present" frauds, credit card details are obtained without the knowledge of card holders. In the former, remote transactions can be conducted using card details through mail, phone, or the Internet.

1. HISTORY

Traditionally businesses mainly relied on rules alone to block fraudulent activities. Today, rules are still an important part of the anti-fraud toolkit but in the past, using them on their own also caused some issues. After 1996 fraudsters started to use the Internet as a test bed for stolen credit cards. Up to this point the fraudsters were still relying on old tried-and-true techniques to get credit card information. They used skimming, phishing, dumpster diving, mail theft, actual theft of people's cards and application fraud.Fraud detection methods are continuously developed to defend criminals in adapting to their fraudulent strategies.

2. RELATED WORK

Find fraud detection need transaction dataset and for finding or classifying need some algorithms. There are plenty of algorithms for finding fraudulent transaction, so first select some better algorithms from Literature review. And Implement better algorithms in python for classifying fraudulent and non-fraudulent transaction.[6]

Researchers developed many credit card fraud detection techniques based on data mining approach. Ghosh and Rilly have proposed credit card fraud detection with a three-layer approach, feed-forward neural network (FFNN), which requires long training time. CARDWATCH: presented by Aleskerov et al. proposed that a neural network based database mining system which was a prototype for database mining system developed for credit card fraud detection application and is concerned that it requires one network per customer. Amalan Kundu et al suggested a model BLASTSSAHA Hybridization technique of credit card fraud by online detection. BLAST-SSAHA approach improves the fraud detection by combining both peculiarities as well as misuse detection techniques. Phua et al have done a major survey of existing data mining based Fraud Detection System (FDSs). Chiu et al have introduced web-services based collaborative scheme for fraud detection in the Banks. The proposed scenario supports the sharing of knowledge about fraud pattern with the participant banks in a heterogeneous and distributed environment. Abhinav srivastava et al have proposed Hidden Markov model (HMM) for credit card fraud detection which shows 80% accuracy over a large variation in the input data. Syeda et al have improved the speed by using parallel granular neural network of data mining and knowledge discovery process (KDP) for credit card fraud detection and achieve reasonable speed up to 10 processors only & more number of processors introduces load imbalance problem. Markov Model and time series are not scalable to large size data sets due to their time complexity. Fan et al recommend the application of distributed data mining in credit card fraud detection and improve the efficiency of highly distributed databases and detection system as this approach uses Boosting algorithm name Ada Cost. Ada Cost uses large number of classifiers and requires more computational resources during detection. Brause et al combine advanced data mining techniques and neural network algorithms. Stolfo et al intimate a credit card fraud detection system using various meta-learning techniques to learn models of fraudulent credit card transactions. To achieve high fraud detection along with low false alarm Elkan et al suggest Naïve Bayesian approach for credit card fraud detection. Further, Elkan and Witten presents that NB algorithm is veryeffective in many real world data sets as well as extremely capable in linear attributes. Bayesian networks were faster and accurate to train but are slower when applied to new instances/occurrence In a online system Vatsa et al. have currently proposed a game-theoretic approach to credit card fraud detection. . Wen-Fang et al have suggested a research on credit card fraud detection model which is based on outlier detection mining on distance sum, which shows that it can detect credit card fraud better than anomaly detection based on clustering. Jianyun et al have shows framework for detecting fraudulent transactions. In his paper work describes an FP tree based method to effectively create user profile for detection of fraud. But on the other hand, this technique doesn‟t recognize unusual patterns i.e. short term behavioral changes of genuine card holders. Today, some of the existing credit card fraud detection techniques which use labeled data to train the classifiers are unable to detect new kinds of frauds. Supervised learning

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

Vol. 8, Issue 6, June 2021

DOI: 10.17148/IARJSET.2021.8652

has some disadvantage, that they require human involvement to optimize parameters. On another hand, decision tree do not require any parameter setting from the user and can build faster compared to other techniques.[5]

III.PROBLEM STATEMENT

The Credit Card Fraud Detection Problem includes modeling past credit card transactions with the knowledge of the ones that turned out to be a fraud. This model is then used to identify whether a new transaction is fraudulent or not.

IV. VARIOUS METHODS USED FOR CREDIT CARD FRAUD DETECTON

There are many emerging technologies that are able to detect credit card fraud detection. Some of condign technologies that will work on some parameters and able to detect fraud earlier as well are listed below:

Bayesian Network: Baye's theorem is derived by Thomas Bayes. These are statistical classifier that predict class membership probabilities such that whether a particular given tuple belongs to a particular class.In this X is considered as "evidence" and H will be some hypothesis such that X belongs to particular class C. In this, we have two kind of probability: In this P(Fr/X) and P(X/Fr) are posterior probability conditioned on Hypothesis. And P(Fr) and P(X) are prior probability of Hypothesis. We calculate the posterior probability ,P(Fr/X),from P(Fr),P(X/Fr) and P(X) are given Baye's theorem is: P(Fr/X)= $P\ X\ Fr\ P(Fr)\ P(X)$ P (Fr/X) is the fraud probability given the observed behavior user X.This Network can model the behavior based on the assumption that whether the user is fraudulent or legitimate.

Decision Tree: Decision Tree is defined as a type of supervised learning in which we make a decision tree to reach at a particular solution. As shown in figure4 they defined that in decision tree we have some internal nodes and each node represent a test on a particular attribute and each branch in decision tree represent an outcome of test and each leaf node will represent class label means output. Decision trees are used for classification in which we give a new transaction for which class label is unknown(means it is unknown whether it is fraudulent or legitimate) and the transaction value is tested against the decision tree. A path is traced from root node to output/class label for that transaction.[7]

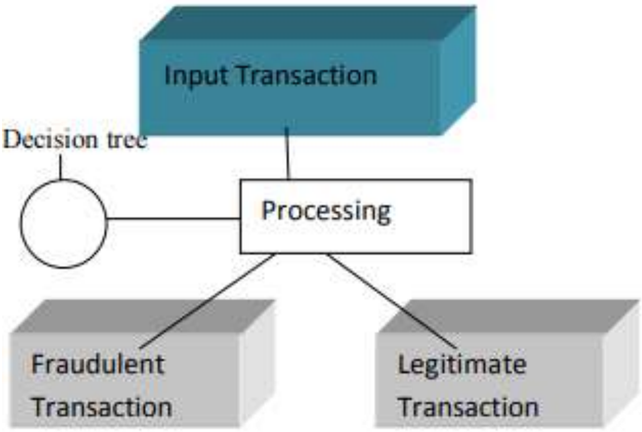

Fig 2: Decision Tree

Support Vector Machine: In supervised learning Vapnik came up with an idea of support Vector Machine.Joseph King-Fun Pun approached thati in this classification algorithm we can construct a hyper plane as a decision plane which can make distinction between fraudulent and legitimate transaction. This Hyper plane Separate the different class of data. Support Vector Machine can maximize the geometric margin and simultaneously it can minimize the empirical classification. So, it is also called Maximum Margin classifier. The separating Hyperplane is a plane that exploit the distance between the two equivalent hyper plane.[7]

V. CONCLUSION

Fraud detection using credit card is a very serious problem in financial services. The loss due to credit card fraud is increasing with the increase in e-commerce. This study deals with techniques that help to find out the credit card fraud. Various techniques like decision tree, Computational Intelligence, ANN, Modified Fisher Discriminant approach and a fusion approach using Dumpster Shafer and Bayesian Learning is also used.

DOI: 10.17148/IARJSET.2021.8652

REFERENCES

[1] A Survey on Credit Card Fraud Detection Using Machine Learning- 2nd International Conference on Trends in Electronics and Informatics (ICOEI)-11-12-May 2018

[2] Fraud Detection in Credit Card using-International Research Journal of Engineering and Technology (IRJET)-April 2020

[3] A Survey of Credit Card Fraud Detection Techniques: Data and Technique Oriented Perspective

[4] International Research Journal of Engineering and Technology (IRJET)-November 2018

[5] A review of Fraud Detection Techniques: Credit Card-International Journal of Computer Applications (0975 – 8887)-May 2019

[6] Credit card Fraud Detection based on Machine Learning Algorithms- International Journal of Computer Applications (0975 – 8887) -May 2019

[7] Survey Paper on Credit Card Fraud Detection- International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)-March 2014