International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 04 | Apr 2020 www.irjet.net p-ISSN: 2395-0072

Detecting Credit Card Fraud using Selected Machine Learning Algorithm

Neha Kamat[1], Swapnali Kakade[2], Rishikesh Gawand[3]

[1-3]*Department of Computer Engineering, Pillai HOC College of Engineering and Technology, Rasayani, India*

ABSTRACT – *In this paper we mainly focus on credit card fraud detection in real world. Here the credit card fraud detection is based on fraudulent transactions. Generally credit card fraud activities can happen in both online and offline. But in today's world online fraud transaction activities are increasing day by day. So in order to find the online fraud transactions various methods have been used in existing system. In proposed system we use Random Forest Algorithm (RFA) for finding the fraudulent transactions and the accuracy of those transactions. This algorithm is based on supervised learning algorithm where it uses decision trees for classification of the dataset. After classification of dataset a confusion matrix is obtained. The performance of Random Forest Algorithm is evaluated based on the confusion matrix.*

Key Words: Credit Card Fraud Detection, Transactions, Classification technique, Random Forest Algorithm.

1. INTRODUCTION

In credit or debit card based purchase, the cardholder presents card to a merchant for making payment. To commit fraud in this kind of acquisition, the fraudster has to steal the credit card. If the legitimate user does not understand the loss of card, it can lead to important financial loss to the credit card company and also to the user. In online payment mode, attackers need only little information for false transaction, for example, secure code, expiration date, card number and many other factors. In this purchase method, many transactions will be done through Internet or telephone. To obligate fraud in these types of purchases, an impostor simply needs to know the card details. Most of the time, the honest cardholder is not aware that someone else has seen or stolen his card information. The only way to detect this kind of fraud is to analyze the spending patterns on every card and to figure out any irregularity with respect to the "usual" spending patterns. The examination of existing purchase data of cardholder is a likely way to reduce the rate of positive credit card frauds. Since humans tend to display specific behaviorist profiles, every cardholder can be characterized by a set of patterns comprising information about the distinctive purchase category, the time since the last buying, the amount of money spent, and other things. Nonconformity from such patterns is reflected as fraud.

Credit card frauds are increasing day by day as the use of credit card is increasing [1]. Occurrence of credit card fraud has increased dramatically both online and offline. Credit card based purchase can be done in two ways: (i) physical card (ii) virtual card. In physical card purchase, the cardholder presents his card physically to the merchant for making payment. For this type of fraud the attacker has to steal the credit card. In virtual card purchase only the information about the card is stolen or gathered like card number, secure code etc. Such purchases are done over the Internet. For this type of fraud the attacker needs only the card details so the only way to detect this type of fraud is to analyze the spending pattern of the card holder. When one's credit card or credit card information is stolen and used to make unauthorized purchases on e-commerce systems on the Internet, one becomes a victim of internet credit card fraud or no card present fraud. This is nothing new and there is nothing unusual about this because identity theft and credit-card fraud are present-day happenings that affect many people and involve substantial monetary losses. Fraud is a million dollar business and increasing every year.

Credit card is refers to a method of selling goods or services without the buyer having cash in hand [2]. A credit card transaction involves four entities. The first entity is the consumer; that is the person who owns the card and who carries out the legitimate transactions. The second entity is the credit card issuer; that is usually the consumer's bank also known as issuing bank which provides the credit services to the consumer. The credit card issuer sends the bill to the consumer in order to request a payment for their credit card transactions. The third entity is the merchant who sells goods or services to the consumer by charging consumer's credit card. This charge is achieved through merchant's bank the forth entity which sends the request for the transaction to the issuing bank. The issuing bank will check whether the amount of the transaction does not reach the credit card's limit before authorizing that transaction. If the transaction is valid, the issuing bank will block the requested amount from consumer's credit card account and send an authorization response to merchant bank. As soon as the authorization response is received by the merchant's bank, the merchant is notified; the transaction is marked as completed and the consumer can take the goods. The blocked amount on consumer's credit card account will be transferred into merchant's bank account.

2. LITERATURE SURVEY

[3] Stated that today it is easy to do banking transaction digitally, both on a computer or by using a mobile phone. As the banking-services increase and get implemented to multi-platforms, it makes it easier for a fraudster to commit financial fraud. In their research, they discovered the need to focus on investigating log-files from a mobile money system that makes it possible to do banking transactions with a mobile phone. They developed a system whose main objective is to evaluate if it is possible to combine two statistical methods, Benford's law together with statistical quantiles, to find a statistical way to find fraudsters within a Mobile Money system. To achieve this, rules were extracted from a case study with focus on a Mobile Money system and limits were calculated by using quantiles. A fraud detector was implemented that uses these rules together with limits and Benford's law in order to detect fraud. The fraud detector used the methods both independently and combined.

Finally, the results obtained showed that it is possible to use the Benford's law and statistical quantiles within the studied Mobile Money system. It is also shown that there is only a very small difference when the two methods are combined or not both in detection rate and accuracy/ precision. Meanwhile, [3] concluded that by combining the chosen methods it is possible to get a medium-high true positive rates and very low false positive rates. The most effective method to find fraudsters is by only using quantiles.

[4] Proposed credit card fraud detection model using Hidden Markov Model. Hidden Markov Models (HMMs) which is a statistical tool and extremely powerful method used for modeling generative sequences characterized by a set of observable sequences. Hidden Markov Model is probably the simplest and easiest model which can be used to model sequential data, i.e. data samples which are dependent on each other. An HMM is a double embedded random process with two different levels, one is hidden and other is open to all. The Hidden Markov Model is a finite set of states, each of which is associated with a probability distribution. Transitions among the states are governed by a set of probabilities called transition probabilities. In a particular state, an outcome or observation can be generated according to the associated probability distribution. It is only the outcome, not the state visible to an external observer and therefore states are "hidden" to the outside; hence, the name Hidden Markov Model [5]. HMM has been successfully applied to many applications such as speech recognition, robotics, bio- informatics, data mining etc.

Achieved their aim by storing all the information about credit card (Like Credit card number, credit card CVV number, credit card Expiry month and year, name on credit card etc.) in the credit card database. If details entered by User into the database are correct then it will ask for Personal Identity number (PIN). After matching of Personal Identity number (PIN) with database and account balance of user's credit card is more than the purchase amount, the fraud checking module will be activated. The verification of all data will be checked before the first page load of credit card fraud detection system. If user credit card has less than 10 transactions then it will directly ask to provide personal information to do the transaction. Once database of 10 transactions is developed, then fraud detection system will start to work. Observation probabilistic in an HMM Based system is initially studied, spending profile of the cardholder and followed by checking an incoming transaction against spending behavior of the cardholder one can show clustering model is used to classify the legal and fraudulent transaction using data conglomeration of regions of parameter, HMM based credit card fraud detection during credit card transaction.

3. EXISTING SYSTEM

In existing System, a research about a case study involving credit card fraud detection, where data normalization is applied before Naïve Bayer's and Cluster Analysis and with results obtained from the use of these methods on fraud detection has shown that by clustering attributes neuronal inputs can be minimized and promising results can be obtained by using normalized data. This research was based on unsupervised learning. Significance of this paper was to find new methods for fraud detection and to increase the accuracy of results. The data set for this paper is based on real life transactional data by a large European company and personal details in data is kept confidential. Accuracy of an algorithm is around 50%. Thus the accuracy of the results obtained from these methods are less when compared with the proposed system.

4. PROPOSED SYSTEM

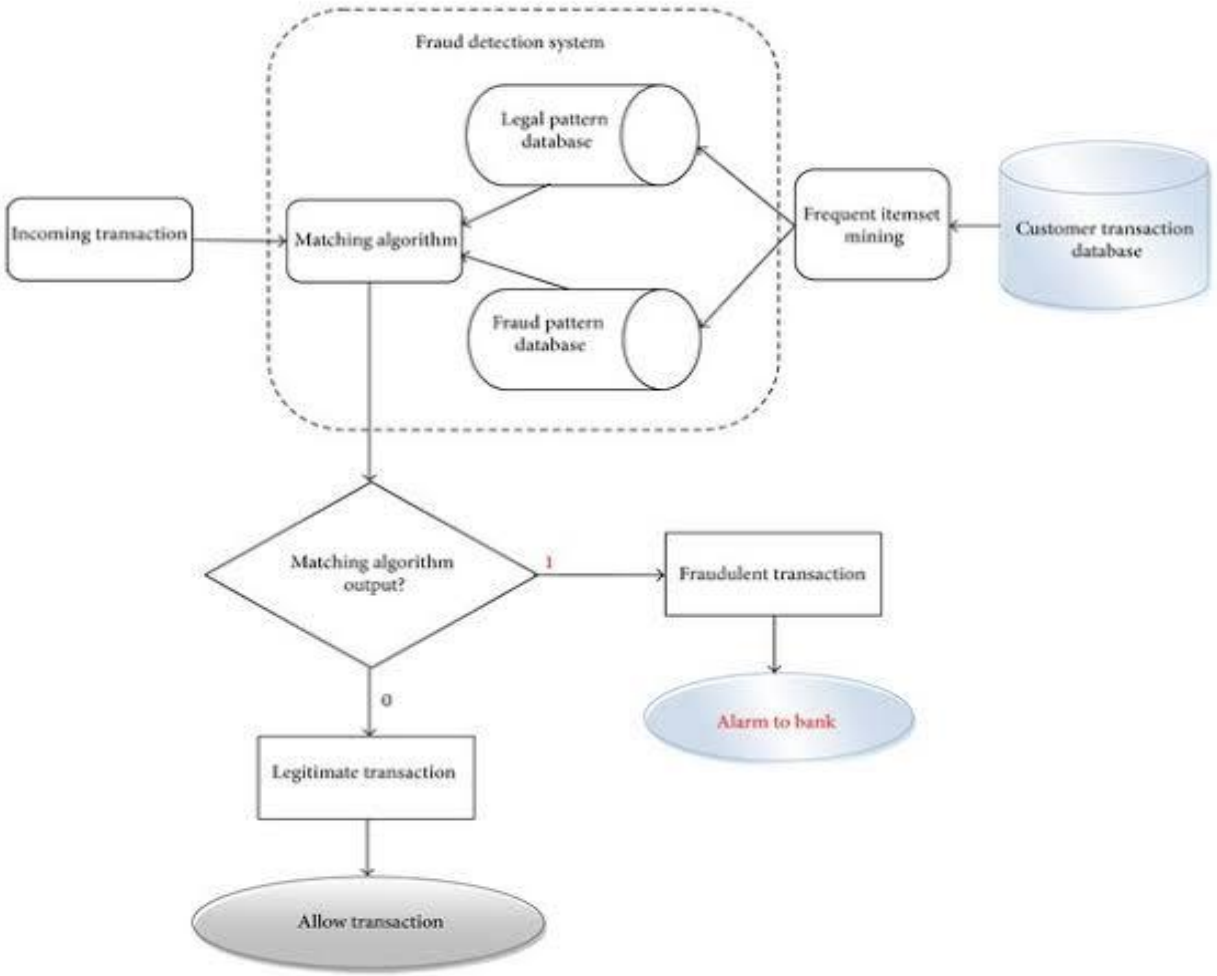

Fig1. System Architecture

In Proposed system we use Random Forest Algorithm and Neural Networks for classification and regression of dataset. First we will collect the Credit Card dataset and analysis will be done on the collected dataset. After the analysis of dataset then cleaning of dataset is required. Generally in any dataset there will be many duplicate and null values will be present, so to remove all those duplicate and null values cleaning process is required. Then we have to split the dataset into two categories as Trained dataset and Testing dataset for comparing and analyzing the dataset. After dividing the dataset we have to apply the Random Forest Algorithm where this algorithm will gives us the better accuracy about the credit card fraud transactions. By applying the Random Forest Algorithm the dataset will be classified into four categories which will be obtained in the form of confusion matrix. Based on the above classification of data performance analysis will be done. In this analysis the accuracy of credit card fraud transactions can be obtained which will be finally represented in the form of graphical representation.

[A] RANDOM FOREST ALGORITHM

Random Forest is also called as Random Decision Forest (RFA) which is used for Classification, Regression and other tasks that are performed by constructing multiple decision trees. This Random Forest Algorithm is based on supervised learning and the major advantage of this algorithm is that it can be used for both Classification and Regression. Random Forest Algorithm gives you better accuracy when compared with all other existing systems and this is most commonly used algorithm. In this paper the use of Random forest algorithm in credit card fraud detection can give you accuracy of about 90 to 95%.

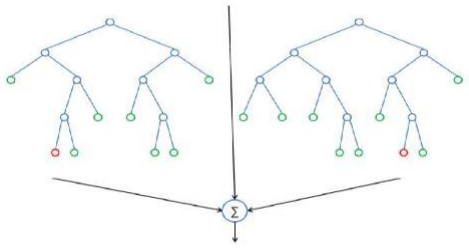

Fig2. Decision Tree

5. CONCLUSION

The work developed a new object oriented approach of solving bank frauds problems especially in the area of credit card fraud. A conceptual framework for a system based on credit card fraud (CCF) process was developed. Various classes of object diagrams were proposed to provide a set of functionalities for CCF in electronic environment for banks.

6. REFERENCES

1. Patel, Twinkle, & Ompriya, Kale. (2014). A Secured Approach to Credit Card Fraud Detection using Hidden Markov Model. International Journal of Advanced Research in computer Engineering and technology, 3(5), 1576.

2. Delawaire, L., Hussein, A.,& John P (2009). Credit Card Fraud and Detection Techniques: A Review, International of Journal of Technology and, 4(2), 57-68.

3. Kappelin, F. and Rudvall, J. (2015): "Fraud Detection within Mobile Money: A mathematical statistics approach" MSc Thesis submitted to the Dept. Computer Science &Engineering Blekinge Institute of Technology SE–371 79 Karlskrona, Sweden.

4. Khan, A. P., Mahajan, V. S., Shaikh, S. H and Koli, A. B. (2013):"Credit Card Fraud Detection System through Observation Probability Using Hidden Markov Model" International Journal of Thesis Projects and Dissertations(IJTPD) Vol. 1, Issue 1, PP: (7-16), Month: October-December 2013, Available At: www.researchpublish.com

5. Ghosh and Reilly (2014). Credit Card Fraud Detection with a Neural Network. IEEE Proceedings of the Twenty-Seventh Annual Hawaii International Conference on System Sciences, 3, 621-630.