A Study on Credit Card Fraud Detection using Data Mining Techniques

G.Suresh¹, R.Justin Raj²

¹Assistant Professor, PG & Research Department of Computer Applications, St.Joseph’s College of Arts & Science (Autonomous), Cuddalore
²Research Scholar, PG & Research Department of Computer Science, St.Joseph’s College of Arts & Science (Autonomous), Cuddalore

Email: sureshg2233@yahoo.co.in, justin.raj752@gmail.com

Abstract - The credit card system is the process which is initiates the global economy to develop significantly. Credit card providers issued millions of credit cards to their customers. They don’t know whether the card holder is wrong customer, if card is issued to wrong customer that can be a very crucial factor of the financial crisis. This survey represents an organized analysis of data mining methods and its applications in credit card process. Our survey mainly focuses on data mining methods implemented especially in credit card process which helps to emphasis much larger parts. Hence, this survey must be very helpful for any credit card providers to select an appropriate solution for their problem as well as for researchers to have comprehensive of the review of literature in their area.

Keywords: - data mining, credit card fraud, credit card fraud detection.

1. INTRODUCTION

In these modern days the style of payment types are changed into online transactions. There are several types of payment for civilizing online transaction which includes e-cash, card payments, internet banking, and e-services. Credit card is one of the most conventional methods of online transaction. It is classified into two types: Physical Card and Virtual Card. For making payment at a physical payment workstation a physical card is used. A digital payment tool that can be used for online transaction is said to be virtual card.

1.1 CREDIT CARD

A credit card is a card issued to consumer (cardholders) to make easy imbursement to a merchant for products and services. It is based on the consumer’s promise to the card issuer. The card issuer (usually a bank) creates an account which is usually circling and contributes a line of credit to the user. From which the user can make use of money for payment.

1.2 CREDIT CARD FRAUDS

Credit card is easy and substituted for cash, and it is also convenient method of payment. It is preapproved credit amount that can be used for purchasing products; payment of that purchase is collected later with agreed charges. The credit cards have some restrictions based upon the user perceived credit worthiness and the maximum amount lend; credit worthiness is an individual ability and readiness to pay money back. Credit card fraud is nothing but usage or removal of other person’s funds without authentication. CEO of Ripple shot Cahn, says that early detection can reduce the credit card fraud.

2. VARIOUS TECHNIQUES FOR CREDIT CARD FRAUD DETECTION

In data mining there are numerous methods for identifying the credit card fraud detection. In this Survey paper we discuss some most useful methods.

• Decision Tree
• Neural Network
• K-Means Clustering
• Hidden Markov Model
• Genetic Algorithm

2.1 Decision Tree
A Data mining induction method that recursively distributes a set of records is Decision Tree Algorithm. This is a method used for solving regression and classification problems. It used the tree representation. It contains one root node, child nodes and leaf nodes. Attribute names are used to labeled the attributes. Values of attributes are used to label the edges. For foretell a label of a class the following method is used. First, it begins from the root node then it compares the cost of the root with record node value. With this result it follows the division corresponding to that cost and travelled to the next node. This process is continued until it arrives the leaf node with predicted class value. It is simple to execute, recognize and exhibit when comparing to other classification algorithm. It is also used for tracing the mail and IP address for detecting credit card fraud. The detection depends on the location. It compares the location of preceding usage of with the present places transaction.

2.2 Neural Networks
Neural network is a technique which is also used for detecting illegal usage of credit card. It shows is valuable result in numerous problems. Neural network is a method which is based on the working principle of human brain. Like human brain, neural network also stores the existing knowledge and uses that information when needed. In detecting illegal usage of credit card, neural network split the information into different groups. It depends on the credit card holder’s earnings, career, and payment details frequency and counting of large purchases. This entire detail is going to evaluate the future transaction that is whether the transaction is scam or authentic. The following figure represents the process of neural network. It has three distinct types of layers.

• Input Layer: Input nodes are used to identify the cardholder details and by using this information it will verify the uniqueness of the transaction.
• Hidden Layer: It perform neural network operation to discover whether the transaction is authenticated or not.
• Output Layer: After analyzing the transaction, output nodes gives the result value between 0 and 1.

2.3 K-Means Clustering
Clustering refers to grouping up of data from the available dataset. In analysis of data for recognize the same pattern or group of information clustering method can be used. It facilitates the bank to draw a decision based on the importance of client and to expose similar kind of techniques used in fraud detection. K-means clustering algorithm is most used method to identify whether the transaction is fraud or legal.

In transaction we declare some variable like, transaction amount, credit card number, transaction date and id, country, merchant category id. Here credit card number is must. If we forget we can’t do the transaction. This process is done in transaction validation section. The information which we got as an input will going to store in transaction dataset. Next we assign the cluster name in which type of transaction that is and label it as, low cluster or high cluster or risky cluster. The transaction detail will take over to k-means algorithm. If the transaction is fraud or legal it displays a message.

2.4 Hidden Markov Model
A set of states connected with the probability distribution is known as Hidden Markov Model. Each and every state produces an output according to the probability distribution which is based on the specific state. In this
method only output can be visible to the user so that it is called Hidden Markov Model. In detecting dishonest transaction of credit card, HMM uses the spending manners of cardholder. Spending pattern of the authenticated card user is calculated by the previous record of transaction which has the attributes like amount that has been transferred, IP address, place of delivery and location of most recent transaction etc. The behavior of the card holder is categorized into three types. They are,
1. Low spending behavior
2. Medium spending behavior
3. High spending behavior

Cardholders who pay low amount for purchase are categorized into behavior of low spending. The cardholder who spends reasonable level of amount are said to be the behavior of medium spending. And finally the cardholder who spends huge amount is categorized into high spending behavior. The following picture demonstrates the detection of credit card fraud using HMM.

![Diagram of Credit Card Fraud Detection](image)

The first level is identification of the consumer that depends on the purchasing patterns of the cardholder. It follows two step processes to identify the illegal usage of credit card. Hidden Markov Model has been prepared by using previous history of transactions. It obtains the input and validate whether the transactions details are accepted by previous training series are not.

2.5 Genetic Algorithm
To get the improved optimal solution genetic algorithm is used. It is also used to identify the fraud transactions with the given sample data set. This method is efficient and secure. It checks whether a transaction is authenticated or not. Transaction using credit card has n number of attributes. At beginning it choose the data set that are going to be processed. Then we select the normalized data from the selected dataset that holds the entire detail about the cardholder. First it calculates the critical values using regularity usage of credit card count, present bank balance, credit card overdraft and place where they use credit card for the particular transaction and average daily spending. Then it compares the data and finally it determines whether the transaction is authenticated or not.

3. CONCLUSION

This paper represents the survey on detecting illegal usage of credit card and the strategies which concerned inside the detecting credit card fraud. Particularly, classification and prediction mission are very essential within the technique of credit card. So why information mining is selected to apply in the research for credit card fraud detection. By way of this research the credit score card presents to discover the real clients and they can lessen the cost and additionally growth profit. They carry a prepared evaluation, suitable time period of information-set and also appropriate-choice of data set.
References