

# An Overview of Credit Card Fraud Detection Using Data Mining Techniques

Arpita Mantri<sup>1</sup>, Chelsi Sen<sup>2</sup>, Dr. Sunil Kumar<sup>3</sup>

<sup>1,2</sup> Dept of Computer Science

<sup>3</sup> Assistant professor, Dept of Computer Science

<sup>1,2,3</sup> Sangam University, Bhilwara, Rajasthan

**Abstract-** Along with the increase of credit card transactions, credit card fraud has become more and more rampant in recent years. Therefore, the main aim of fraud detection systems is to detect fraud accurately and before fraud is committed. The goal is to detect and accurate false fraud detection. There are several methodologies for detecting credit card fraud, like neural networks, Genetic Algorithms, k-means clustering. The objectives of the current study area are first to detect various MasterCard and electronic business fraud and at the moment investigate the methods used for the aim of detection.

**Keywords-** Credit card fraud, data mining, credit card fraud detection.

## I. INTRODUCTION

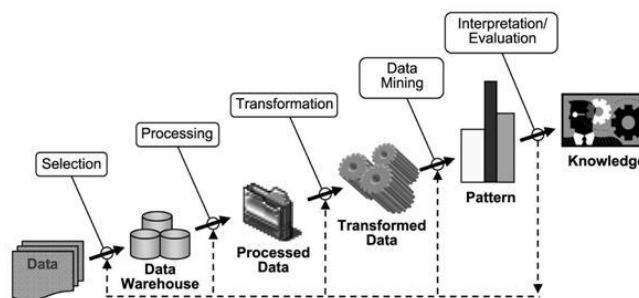
The advancement and innovation in technology has open several new doors for committing fraudulent acts. These acts force genuine risk to organizations on the financial, operational and mental dimensions. Notwithstanding the fiscal misfortunes, fraud can staggeringly affect the organizations reputation, altruism and client relations. Hence, organizations attempt to execute an assortment of techniques to detect and prevent fraud. Among those techniques is data mining

Credit Card Fraud is one of the biggest dangers to business associations today. Be that as it may, to overpower the fraud effectively, it is imperative to initially comprehend the mechanisms of executing a fraud for example we have to comprehend the techniques of cyber credit card frauds. Since prior the fraud is detected only when the billing for credit card is done, it is difficult to forestall fraudulent transactions. In this way the need to guarantee unexposed transactions for credit-card proprietors when utilizing their credit cards to make electronic payments for products and enterprises gave on the internet is a paradigm. This research paper explores a portion of the data mining techniques utilized for credit card fraud detection. Before going into the details, a short depiction of fraud and data mining is acquaint to clear the way.

## II. WHAT IS DATA MINING?

Data mining refers to removing valuable information from tremendous measures of data. Many individuals treat data mining as an equivalent word for another prominently utilized term, knowledge discovery from data, or KDD, while others see data mining as just an essential step during the process of knowledge discovery [1].

In Data mining, Knowledge extraction or discovery can be done in seven sequential steps as in Fig:[2]



1. Data cleaning: This is the initial step to remove noise data and unessential data from gathered raw data.
2. Data integration: At this step, different data sources are consolidated into important and valuable data.
3. Data Selection: Here, data relevant to the analysis are recovered from different sources.
4. Data transformation: In this progression, data is changed over or combined into required forms for mining by performing diverse tasks, for example, smoothing, normalization or aggregation.
5. Data Mining: At this progression, different shrewd techniques and tools are connected so as to extricate data pattern or principles.
6. Pattern evaluation: At this progression, Attractive patterns representing knowledge are distinguished dependent on given measures.
7. Knowledge representation: This is the last stage in which, perception and knowledge representation procedures are utilized to assist users to understand and translate the data mining knowledge or result.

The goal of knowledge discovery and data mining process is to discover the patterns that are unknown among the huge set of data and interpret useful knowledge and information.

### III. KINDS OF CREDIT CARD FRAUD

**Credit Card:** A credit card is a card issued to shopper (cardholders) to make simple imbursement to a trader for items and administrations. It depends on the shopper's guarantee to the card guarantor. The cardboard institution (usually a bank) creates an account that is typically circling and contributes a line of credit to the user. From that the user will create use of cash for payment. [4]

**Types of frauds:** The credit-card fraud is isolated into two sorts:

- I. The online credit card fraud (or no card present misrepresentation) and
- II. The offline credit card fraud (card present misrepresentation)

The online credit-card fraud (otherwise called digital credit card fraud) is submitted with no presence of a credit-card however rather, the utilization of a credit-card data to make electronic buy for merchandise and enterprises on the web. The offline credit-card fraud is submitted with the presence of a credit-card which as a rule has been stolen or fake and accordingly utilized at a nearby store or a physical area for the buy or a few products or services.

There are many cyber credit card fraudsters. Some of them are:[3]

- (i) **Credit-card information buyers:** these are the fraudsters who either have little or no professional computer skills like computer programming, networking etc. They buy stolen or hacked credit card information on an illegal credit card sales website, with the intension of buying goods and products online.
- (ii) **Physical credit-card stealers:** these are the fraudsters who physically steal credit cards may be by pick pocketing and use the information on it for making e-payment on internet for shopping.
- (iii) **Black hat hackers:** "Black hat hackers" which are otherwise called a cracker are the individuals who damage PC security with pernicious expectations or for individual gains. They pick their objectives

utilizing a two dimensional procedure known as the "pre-hacking stage"; which incorporates Targeting, Research and Information Gathering lastly completing the Attack. These hackers are profoundly talented in Computer Programming and Computer Networking and with such skills they can burst in a network of computers. The fundamental reason for their act of interruption or hacking is to steal individual or private data, for example, credit-card data, bank-account data, and so on for their very own personal gain.

### IV. VARIOUS TECHNIQUES FOR CREDIT CARD FRAUD DETECTION

In data mining there are various methods for distinguishing the credit card fraud detection. In this Survey paper we talk about some most helpful methods. [4][5]

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

#### Decision Tree

A Data mining acceptance technique that recursively conveys a lot of records is Decision Tree Algorithm. This is a technique utilized for taking care of regression and classification issues. It utilized 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 utilized to mark the edges. For predict a mark of a class the following technique is utilized. To start with, it starts from the root node then it compares the cost of the root and record node value. With this outcome it pursues the division corresponding to that cost and traveled to the following node. This procedure is preceded until it arrives the leaf node with anticipated class value. It is easy to execute, recognize and display when comparing with other classification algorithm. It is likewise utilized for tracing the mail and IP address for detecting credit card fraud. The detection relies upon the location. It compares the location of going before use of with the present places transaction.

#### Neural Network

Fraud detection utilizing neural network is completely founded on the human mind working principal. Neural network technology has made a computer equipped to think. As human mind learn through past involvement and

utilize its information or experience in settling on the choice in everyday life issue a similar technique is applied with the credit card fraud detection technology. At the point when a specific customer utilizes its credit card, There is a fix pattern of credit card use, made by the manner in which customer utilizes its credit card. When credit card is being utilized by unauthorized client, the neural network based fraud detection framework check for the pattern utilized by the fraudster and matches with the pattern of the authorized card holder on which the neural network has been prepared, if the pattern matches the neural network proclaim the authorize transaction. At the point when a transaction arrives for authorization, it is characterized by a stream of authorization data fields that convey data distinguishing the cardholder (account number) and characteristics of the transaction (e.g., amount, merchant code). There are extra data fields that can be taken in a feed from the authorization framework (e.g., time of day). The neural network is design to produce output in real values somewhere in the range of 0 and 1 .If the neural network produce output that is underneath .6 or .7 then the transaction is alright and in the event that the output is over .7 then the possibility of being a transaction illegal increase. In the structure of neural network-based pattern recognition Systems, there is dependably a procedure of business History descriptors contain highlights portraying the utilization of the card. For transactions, the payments made to the account over Some promptly earlier time interim. Other a few descriptors can Include such factors as the date of issue (or latest issue) of the credit card. This is critical for the detection of NRI (non-receipt of issue) fraud.

### **K-Means Clustering**

Clustering refers to gathering up of data from the accessible dataset. In examination of data for recognize a similar pattern or gathering of data clustering technique can be utilized. It facilitates the bank to draw a decisionbased on the significance of client and to uncover comparative kind of techniques utilized in fraud detection. K-means clustering algorithm is most utilized technique to identify whether the transaction is fraud or lawful.In transaction we proclaim some variable like, transaction amount, credit card number, transaction date and id, country, merchant category id. Here credit card number is must. In the event that we overlook we can't do the transaction. This procedure is done in transaction validationsection. The data which we got as an input will going to store in transaction dataset. Next we assign the cluster name in which sort of transaction that is and mark it as, low cluster or high cluster or risky cluster. The transaction detail will take to k-means algorithm. On the off chance that the transaction is fraud or legal it shows a message.

### **Hidden Markov Model**

A set of states associated with the probability distribution is known as Hidden Markov Model. Every single state creates an output as per the probability distribution which depends on the particular state. In this technique output can be visible to the client only which is why it is called as Hidden Markov Model. In detecting exploitative transaction of credit card, HMM utilizes the spending habits of cardholder. Spending pattern of the authenticated card client is determined by the past record of transaction which has the characteristics like amount that has been exchanged, IP address, place of delivery and location of latest transaction and so on. The behavior of the card holder is classified into three types. They are,

1. Low spending behavior
2. Medium spending behavior
3. High spending behavior

Cardholders who pay low amount for buy are classified into behavior of low spending. The cardholder who spends sensible dimension of amount are said to be the behavior of medium spending. Lastly the cardholder who spends tremendous amount is classified into high spending behavior.

The primary level is distinguishing proof of the consumer that relies upon the purchasing patterns of the cardholder. It pursues two step procedures to identify the illicit utilization of credit card. Hidden Markov Model has been set up by utilizing past history of transactions. It gets the input and validate whether the transactions details are acknowledged by past preparing series are most certainly not.

### **Genetic Algorithm**

To get the improved ideal arrangement genetic algorithm is utilized. It is likewise used to identify the fraud transactions with the given sample data sets. This strategy is proficient and secure. It checks whether a transaction is authenticated or unauthenticated. Transaction utilizing credit card has n number of attributes. At starting it pick the data set that will be prepared. At that point we select the standardized data from the selected dataset that holds the whole insight concerning the cardholder. First it calculates the critical values utilizing consistency use of credit card count, present bank balance, credit card overdraft and place where they use credit card for the specific transaction and normal every day spending. At last, it compares the data and then determines whether the transaction is authenticated or not.

## V. CONCLUSION

This paper starts with an overview of the concepts of data mining and fraud detection, trailed by a discussion of different techniques for credit card fraud detection. One aim of this study is to spot the user model that best identifies fraud cases. There are numerous methods for detection of credit card fraud. If one in all these or combination of rule is applied into bank credit card fraud detection system, then the probability of fraud transactions can be predicted soon after credit card transactions by the banks. A series of anti-fraud ways is adopted to forestall the banks from nice losses sooner and scale back the risks.

## REFERENCES

- [1] Jiawei Han, Micheline Kamber, Data Mining: Concepts and Techniques, London: Academic Press, 5, 2001.
- [2] Deepashri, K M. "Survey on Techniques of Data Mining and its Applications.", International Journal of Emerging Research in Management & Technology, Vol 6 No. 2, February 2017.
- [3] Sethi, Neha and Anju Gera. "A Revived Survey of Various Credit Card Fraud Detection Techniques", International Journal of Computer Science and Mobile Computing, Vol 3 No. 4, pp. 780-791, April 2014.
- [4] G. Suresh and R. Justin Raj, "A study on credit card fraud detection using data mining techniques", International journal of Data Mining Techniques and Applications, Vol 7 No. 1, pp. 2278-2419, June 2018.
- [5] Dinesh L Talekar and K.P. Adhiya, "Credit Card Fraud Detection System: A Survey", International Journal Of Modern Engineering Research (IJMER).
- [6] Ogwueleka, F. N. (2008). Credit card fraud detection using data mining techniques. Ph.D. Dissertation. Department of Computer Science. Nnamdi Azikiwe University, Awka, Nigeria.
- [7] Maes, S.; Tuyls, K.; Vanschoenwinkel, B.; and Manderick, B. (2002). Credit card fraud detection using Bayesian and neural networks. Proceeding at International NAISO Congress on Neuro Fuzzy Technologies.
- [8] Bhatla, T.P.; Prabhu, V.; and Dua, A. (2003). Understanding credit card frauds. Crads Business Review# 2003-1, Tata Consultancy Services.
- [9] Edelstien, H.A. (1999). Introduction to data mining and knowledge discovery. (2nd Ed.), Two Crows Corporation.
- [10] Haykin, S. (1999). Neural networks: A comprehensive foundation. Prentice Hall. New York.
- [11] Kohonen, T. (1997). The self-organizing map. Proceedings of the IEEE, 78(9), 1464-1480.