Literature Study – Data Mining Techniques on Detecting Fradulent Activities in Credit Card

S. K. Saravanan¹, Dr. G. N. K. Suresh Babu²

¹Research Scholar - Research and Development Centre, Bharathiar University, Coimbatore, India
²Associate Professor - Dept. of Computer Applications, Acharya Institute of Technology, Bangalore, India

Abstract:
In contemporary days the more secured data transfer occurs almost through internet. At same duration the risk also augments in secure data transfer. Having the rise and also light progressiveness in e–commerce, the usage of credit card (CC) online transactions has been also dramatically augmenting. The CC (credit card) usage for a safety balance transfer has been a time requirement. Credit-card fraud finding is the most significant thing like fraudsters that are augmenting every day. The intention of this survey has been assaying regarding the issues associated with credit card deception behavior utilizing data-mining methodologies. Data mining has been a clear procedure which takes data like input and also proffers throughput in the models forms or patterns forms. This investigation is very beneficial for any credit card supplier for choosing a suitable solution for their issue and for the researchers for having a comprehensive assessment of the literature in this field.

Keywords: Data Mining, Credit Card, Fraud Detection, Skimming

I. INTRODUCTION

Credit Card is beneficial for everyday life. The CC (credit card) has been a tiny plastic card, which is issued as a payment system to the user. Today, several transactions are implemented with the CC (credit card)s usage associated with different operations in online and also offline. Hence, security to such kinds of transactions must be given for a great extent. [1] Credit card suppliers have been supplied multifarious credit cards for their customers. Whilst providing credit cards to any incorrect customers which may be a salient component of the financial crisis. [2] These days with the broadening of credit cards with online transactions, it has been a vital issue for financial institutions in their effort to restrict Credit Card Treachery behavior. [3] The information kept in on the CC (credit card) may be read by the Automatic Teller Machines (ATM’s), banks, store readers, and also utilized in online Internet Banking System. Having rapid development in its transactions of credit card number and divergent applications, every treacherous activity is augmented.

1.1. Credit Card Treachery

Credit Card Treachery has been signified as, while an individual utilizes another individual credit card (CC) their personal usage whilst the card owner and the card supplier have not been attentive of that someone is utilizing his card. [4] In the CC (credit card) trickery, only little salient information as per a card (card number, expiration date, secure code,) is required for doing the payment. Such purchases are usually executed on the telephone or by internet. [5] On account of the rapid progressiveness of transactions of credit card straight to a salient in treacherous behavior. Credit Card Treachery has been a general term to the theft and also fraud devoted utilizing credit card like a deception source of funds in the provided transactions. [6] Credit Card Treachery have been acknowledged in the below methods,

- Thievery of actual cards.
- Falsification of account or else personal information,
- Illegal or else unauthorized usage of account to personal gain.

Credit Card Treachery happens online and offline.

- Whilst unofficial users uses credit card having the PIN is termed online fraud. Utilizing physical card to transactions for instance. restaurants, exchanging electronic goods, etc.,
Whilst an illegal user uses credit card with no PIN is termed offline fraud transaction, for instance, by shopping websites, phone transactions etc.,

The below figure 1 illustrates the divergent sorts of Credit Card Treachery. [7] The intent of the fraud finding systems has been verifying each transaction to the probability of being treacherous despite of the obstacle mechanisms, and for recognizing fake ones as speed as believable after the fraudster has commenced for executing a fake replacement. The most notorious fraud sorts are treacherous transactions in the CC (credit card) systems and also e-commerce systems, money laundering in financial systems, intrusions for computer systems, treacherous calls or else service utilizations in telecommunication systems, and also treacherous claims in the health and also auto insurance systems.

1.2. Methods utilized by fraudsters
a) Merchant Based Frauds:
   Merchant based frauds are commenced either by the merchant owners’ establishment or else their employees. The sorts of frauds commenced by merchants are illustrated below:
   i. Merchant Collusion: This sort of fraud happens while merchant owners or else their employees work together for committing fraud utilizing the cardholder accounts or else by utilizing the personal information. Then they transfer the knowledge regarding cardholders for fraudsters.
   ii. Triangulation: Triangulation has been a kind of fraud that is accomplished and operates by a web site. Here, the customer whilst browsing the site and also if he wishes the product he put the online information like name, address and also legitimate credit card explanations to the site. Whilst the fraudsters get such information, they arrange goods from a rightful site utilizing snatched credit card information. Then the fraudsters by utilizing the CC (credit card) facts buy the products.

b) Lost/ Stolen Cards:
   Whilst one person misses his card or else a card is pilfered by someone or else whilst an authentic account holder gets a card and misses it or else someone steals it for illegal intentions. It is the hardest customary credit card form fraud for tackling.
c) Account Takeover:
This sort of fraud happens whilst the suitable customer’s personal information has been derived by the fraudsters. The fraudster controls a lawful account through either proffering the customer’s account number or else the card number. Here, the fraudster then links the card supplier, like the authentic cardholder, to appeal the mail for redirecting to an innovative address. The fraudster tells that card lost and appeals for a substitution for being sent.

d) Fake and also Counterfeit Cards:
This is other sort of fraud in which counterfeit cards’ the creation, combined with lost or else stolen cards proffers greatest hazard in Credit Card Frauds. The fraudsters are incessantly finding more innovative and new innovative ways for creating counterfeit cards.

e) Skimming:
Skimming has been rapidly evolving as the most well known Credit Card Treachery form. Several situations of Counterfeit fraud associates skimming. It is a process where the real data on a card’s magnetic stripe has been electronically copied into another. Here, Fraudsters are discovered found to have pocket skimming devices, which are a battery-operated electronic magnetic stripe reader, that they may swipe customer’s cards for having customer s card data.

Thus by assaying the above methods, Credit card associated fraud may be generally separated as three categories

✓ Fraud perpetrated with physical card.
✓ Fraud inflicted on merchant. This sort of fraud is perpetrated either by merchant owners (named merchant collusion) or else the merchant employees (triangulation). Here, in merchant collusion, merchant owners run into with every other for stealing money from their customers. Stolen details of card holders are derived and provided to fraudsters.
✓ Fraud perpetrated by the internet: fraudsters illegitimately derive card holder details and usage it for purchasing items online. [8]

Generally, the statistical techniques and the data mining algorithms is utilized for solving this fraud finding issue. Fraud finding methodologies are constantly developed for defining offenders in familiarizes their policies. Data mining denotes for deriving or else mining knowledge from enormous amount of data. Data mining has been a procedure that utilizes an assortment of data evaluation tools for finding patterns and also relationships in data which is utilized for making a applicable forecast. [9]

II. METHODOLOGIES FOR FINDING FRAUDS
Credit Card Treachery finding has been a stratification issue. Divergent data mining methodologies for fraud finding is below listed:

A. Logistic Regression
Logistic Regression is a universal linear model Logistic regression which is beneficial for conditions where we can assume the existence or else the absence in a behavior or else outcome values of a series of the predictor variables. It has been same for a linear regression model yet is apt for models where the dependent variable is dichotomous. Here, Logistic regression coefficients are utilized for estimating odds ratios to every single variable in the model and also it is suitable for a broader range in research situations than the discriminant evaluations. Qualitative response designs are suitable when dependent variable is definite. Here, our dependent variable fraud has been binary, and also logistic regression is a broadly utilized technique in these issues. [13] Logistic regression is utilized to assume the outcome of a dependent variable according to one or else more predictor variable. Also, Predictor variable may be either categorical or else numerical.

Advantages:
It produces an easy probability formula to stratification. It functions well having linear data to Credit Card Treachery finding.

Disadvantages:
It doesn’t function well with non-linear data to Credit Card Treachery detection.
B. Decision tree:

Decision trees have been predictive decision supporting tools which form mapping from explanation to feasible consequences. Also, Decision tree is a flow related structure. Moreover, the decision tree is a structure which comprises root node, leaf node and branch. Every internal node signifies an experiment on the attribute, the consequence of test signifies every branch and also the class label keeps by every leaf node. Furthermore, the root node is the upmost node in the tree. The test on the attribute denotes every internal node. Every leaf node signifies a class. Figure 3 presents the example to decision tree. [14]

Advantages:

- It can manage non-linear and also interactive impacts of input variable.

Dis-Advantages:

- It contains complicated algorithm. The tree structure will modify, while there has been a small modification in the data observed. Selecting splitting criteria is also tough.

![Decision tree example](image)

C. K-Nearest neighbour algorithm:

K-Nearest neighbour algorithm has been one of best data mining method in fraud finding issue. Here in adjacent neighbour algorithm every transaction is stratified concerning closeness i.e., distance. Through computing the distance or else closeness to another transaction, each new incoming transaction is classified. Whilst they are near, then it will be signified as a triumphant transaction else the transaction is denoted as a fraud. K-nearest neighbor (KNN) classifiers are regarding learning by analogy.

Advantages:

- It does not need any predictive modeling prior to stratification.

Dis-Advantages:

- Accuracy is greatly dependent on the distance measure. [15]

D. Naive Bayes algorithm

Naive Bayes is utilized like simple Probabilistic classifier according to Bayes conditional probability law. Naive Bayes classifier assists a restricted independence assumption in which an attribute value impact of a provided class is self-regulating of other characteristics. The benefit is that it only proffers a theoretical rationalization to the reality but does not utilize bayes theorem. Limitation: In actual practice the dependences prevails betwixt the variable. [16]

E. Artificial Neural Network

Neural network is signified as a set of interrelated nodes intended for representing the human brain functioning. Every terminal has a weighted correlation to many other associated nodes in nearby layers. Every node take the input derived from connected nodes and utilize the weights of the linked nodes combined with simple function to compute the output values. As described in the figure 4, the neural network have been train on information concerning to divergent categories regarding the card holder occupation, income, occupation...
may fall in one type, whilst in other category information concerning the enormous amount of purchasing are placed. Such information comprise the number of huge purchase, frequencies in huge purchase, location in which such sort of purchase occurs etc. In a fixed duration, while credit card is being utilized by unlawful user the neural network related fraud finding system examine for the pattern utilized by the fraudster and also compares with the actual card holder pattern where the neural network is qualified, where the pattern compares the neural network announce the transaction ok. [17]

![Neural network layers in credit card](image)

Figure 4: Neural network layers in credit card

*Advantages:*
Initially, this is adaptive; next, it can create powerful models; and third, the stratification procedure is changed if novel training weights have been set. Neural networks are applied chiefly to credit card, automobile insurance and also corporate fraud.

**F. Hidden Markov Model**

Hidden Markov Model (HMM) has been a double embedded stochastic method that is utilized to model more intricacy stochastic methods as matched with a conventional Markov model. This is a statistical model having a finite series of states; each one is connected with a probability distribution. While transaction of an incoming credit card is not acknowledged by the qualified HMM with adequately great possibility, it is recognized for being treacherous transactions. Furthermore, A HMM is first qualified with the general characteristic of a cardholder. It functions on the user spending profiles that is separated into three categorizations of Lower profile, Middle profile, and higher profile. For each credit card, the spending profile is divergent; hence it may figure out a variation of user profile and also strives for finding treacherous transaction. [18] The benefit of utilizing HMM related model is diminishing false positive transaction predict like fraud though they are authentic customer.

Table 1: Comparison on assorted classification data mining methods

<table>
<thead>
<tr>
<th>METHODS</th>
<th>SPEED OF DETECTION</th>
<th>ACCURACY</th>
<th>COST</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMM(HMM)</td>
<td>Fast</td>
<td>Low</td>
<td>High Expensive</td>
</tr>
<tr>
<td>Neural Networks</td>
<td>Fast</td>
<td>Medium</td>
<td>Expensive</td>
</tr>
<tr>
<td>Fuzzy Neural Networks</td>
<td>Very Fast</td>
<td>Good</td>
<td>Expensive</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>Very Fast</td>
<td>High</td>
<td>Expensive</td>
</tr>
<tr>
<td>SVM(Support Vector Machine)</td>
<td>Low</td>
<td>Medium</td>
<td>Expensive</td>
</tr>
</tbody>
</table>

**III. REVIEW OF RESEARCH ON CREDIT-CARD FRAUD FINDING**

Credit-card fraud finding is the most salient and involved research field. Numerous data mining techniques are there for solving such treacherous issue.

Nuno Carneiro *et.al* [19] have intended a data mining related system to credit-card fraud finding in e-tail. They provided the integration of manual and also automatic stratification, which proffers insights for the whole development procedure and matches divergent machine learning techniques. Thus the paper assist
researchers with practitioners for designing and executing data mining related systems to fraud finding or else similar issues. Much contribution is given by them with a mechanized system, and with insights for the fraud analysts to enhance their manual revision procedure, that results a complete superior performance.

D. Sanchez et al. [20] have intended a series of rules to Credit Card Treachery detection utilizing data mining methods. The aim of data mining has been deriving beneficial, non-explicit facts from database. In the data mining context, one of the most examined knowledge extraction models has been association rule that presumes that the fundamental object of attention is an item. The piece of that information seems in the existence of an item set named transaction. Moreover, they have enhanced the data mining procedure for extricating association strategies utilizing fuzzy logic. Such data mining technique contains the below three stages: 1. The set of linguistic labels establishment that assist the test transaction set for being defuzzified, utilizing Serrano’s software tools 2. The linguistic labels’ Incorporation and their membership degrees like items in the client and also the transaction tables. 3. Application in the Fuzzy Query 2+ software tool. Such tool allows evaluation to be executed with divergent levels of confidence, support, certainty factor and also number of items for being recognized in the procedure. 4. Assessment of the consequences, choosing from the fuzzy association rule set that rules having a certainty component in a particular threshold and that are not very clear or which are general sense. 5. Iterate from Step 2, changeable the assistance, confidence and also certainty factor, for optimizing this procedure. The utilized methodology overcomes the intricacies of minimum assistance and also confidence, maximizes the implementation times, diminishes the unnecessary creation of rules, and assists make the outcomes more intuitive, thence enabling the fraud analysts task.

Sanjeev Jha et.al [23] have intended a policy to find credit-card frauds through utilizing transaction information. They applied transaction aggregation strategy for detecting Credit Card Treachery. They also aggregated transactions for capturing consumer buying attitude before every transaction and also utilized such aggregations to model evaluation for recognizing treacherous transactions. We utilize real-life data of transactions of credit card by an international credit card operation to transaction aggregation and also model evaluation.

Suvasini Panigrahi et.al [24] have suggested a novel method to Credit Card Treachery detection, that integrates evidences from recent and past attitude. The fraud finding system (FDS) comprises four features, called, rule-related filter, transaction history, Dempster–Shafer adder, database and also Bayesian learner. In the rule-related component, we resolve the suspicion stage of every incoming transaction regarding the degree of its divergence by good pattern. Here, Dempster–Shafer’s theory is utilized for combining several such proofs and an original belief is calculated. The transaction is stratified as normal, abnormal or else suspicious relying upon this original conviction. Once a transaction is discovered to be doubtful, belief is strengthened further or weakened based on its similarity having treacherous or else true transaction history utilizing Bayesian learning. Widespread simulation having stochastic models presents that different evidences fusion has a great positive result on the performance in a Credit Card Treachery detection system as matched with other techniques. They have utilized stochastic models for generating synthetic transactions to assay the system performance.

Jon T.S. Quah et.al [25] have been suggested an computational intelligence method for actual time Credit Card Treachery detection. Their task aimed on real-time fraud finding and also suggested an innovative and novel approach in comprehending spending patterns for decipher potential fraud cases. It uses self-organization map for deciphering, filtering and analyzing customer attitude for fraud finding. SOM has been a neural network which uses unmanaged learning for configuring its neurons concerning the topological formation of the input data. Such process, termed as self-organization, has been an iterative tuning in the weights of neurons hence for approximating the input data. The approach uniqueness lied in utilizing the clustering and also filtering capabilities in SOM to fraud finding. Here, Clustering assists in recognizing novel concealed patterns in the input data, which otherwise may not be recognized by conventional statistical techniques. Thus, the key for correct fraud finding and also customer profiling lied in enhancing a powerful system was designed.

Abhinav Srivastava et.al [26] have been designed a Credit Card Treachery detection systems utilizing HMM. They designed the set of operations in the CC (credit card) transaction processing utilizing a HMM (HMM) and presented how it may be utilized for the frauds detection. An HMM was originally trained with usual attitude of a cardholder. While an incoming credit card transaction had not been acknowledged by the qualified HMM having sufficiently great probability, it may be recognized as fraudulent. Meanwhile, they have tried for ensuring that true transactions were not discarded. They have utilized the transaction amount ranges

© www.ermt.net All Rights Reserved
like the observation symbols, where the kinds of item had been identified to be the HMM states. They have recommended a technique for detecting the expenses profile in cardholders, and application of this information in determining the observation symbols value and original estimate in the model parameters. Also, It had been demonstrated how the HMM could find if an incoming transaction had been treacherous or else not. Investigational outcomes presented the effectiveness and performance of their system and illustrate the usage of analyzing the expenses profile in the cardholders. Proportional analysis present that the Accurateness of the system had been near to 80 percent through a broad variation of the input data. Also, the system had been measurable to manage enormous volumes of transactions.

William N. Robinson et.al [33] have been suggested an sequential fraud finding system utilizing HMM divergence. . Their task flow presented the information regarding how a transaction design may be dynamically generated and also updated, and fraud might be automatically found for prepaid cards. Moreover, they have utilized the store terminals models which were formed through a card processing company. The utilized technique that creates automatically, updates, and also matches HMMs (HMM) of the merchant terminals. Also, they have suggested fraud finding and also evaluations on actual transactional data, that was presenting the efficacy and quality of the method. In fraud test cases, obtained from recognized fraud cases, the methods contain a good F-score.

P. Ravisankar et.al [35] have been suggested a finding of financial fraud statement system utilizing data mining methods. For recognizing companies which resort to the financial statement fraud, they applied data mining techniques of Multilayer Feed Forward Neural Network (MLFF), Genetic Programming (GP), Support Vector Machines (SVM), Logistic Regression (LR),Group Method of Data Handling (GMDH), and Probabilistic Neural Network (PNN). Every method was verified on a dataset connecting 202 Chinese companies and matched with and with no feature selection. PNN outperformed every technique with no feature choice, GP and also PNN performed others having feature selection with marginally parallel accuracies.

Table 2: Techniques and limitations of reviewed articles

<table>
<thead>
<tr>
<th>Authors</th>
<th>Techniques used</th>
<th>Plus</th>
<th>Minus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carminati et al. [27]</td>
<td>A semi-supervised and also unsupervised decision support system to manage fraud and also anomaly detection had been suggested.</td>
<td>Suggested technique is according to actual word data and requisitions.</td>
<td>Clustering phase of suggested technique have huge storage space</td>
</tr>
<tr>
<td>Van Vlasselaer et al. [28]</td>
<td>A new, dynamic and also correct model was suggested</td>
<td>New technique for fraud propagation by the network commencing from a restricted series of labeled edges (for instance, treacherous transactions) and inferring a score to ever network components (i.e., credit card holders, merchants and also transactions).</td>
<td>Excessive Level of Data imbalance.</td>
</tr>
<tr>
<td>Modi et al. [29]</td>
<td>New and also enhanced Credit Card Treachery detection solution had been suggested</td>
<td>Fraud transactions were assumed soon after the transactions of credit card.</td>
<td>ANN training time is unhurried</td>
</tr>
<tr>
<td>Mhamane and Lobo [30]</td>
<td>Novel HMM was utilized</td>
<td>Removals of genuine transactions are prohibited by using one time password that is created by server and then</td>
<td>Explanations regarding the result is not announced</td>
</tr>
</tbody>
</table>
### IV. CONCLUSION

Credit Card Treachery has been a criminal dishonesty act. Credit card finding has been an enthralling domain. Several techniques are available for finding Credit Card Frauds. The salient attention of this analysis is recognizing the best data mining algorithms to Credit Card Treachery detection. While one of these or else integration of algorithm is utilized into bank Credit Card Treachery detection system, the possibility of fraud transactions may be assumed once the transactions of credit card are done by the banks. Moreover, Anti-fraud policies may be adopted for restricting banks from huge losses before and then trim down risks.

### REFERENCES


[34] S. Bekireva, V. V. Klimova, M. V. Kuzinb, and B. A. Shchukina “Payment Card Fraud finding Using Neural Network Committee and Clustering”, Optical Memory and Neural Networks Vol. 124, No. 3, pp. 193-200, 2015.
