Available Online at www.ijecse.org **ISSN- 2277-1956**

Exploration of Data mining techniques in Fraud Detection: Credit Card

Khyati Chaudhary [1], Bhawna Mallick [2],
[1,2]Galgotias College of Engg. & Technology, Greater Noida
E-mail id: khyati1907@gmail.com

Abstract- Data mining has been increasing as one of the chief key features of many security initiatives. Often, used as a means for detection of fraud, assessing risk as well. Data mining involves the use of data analysis tools to discover unknown, valid patterns as well as relationships in large data sets. Decades have seen a massive growth in the use of credit cards as a transactional medium. Data mining become even more common in both the private and public sectors. Data mining has been used widely in industries such as Banking, Insurance, Medicine and Retailing to reduce costs, enhance Research and increase Sales. Credit cards are much safer from theft than is cash and also a promising area for buying and sales. Credit Cards are growing as a popular medium of transaction. Therefore, Fraud Detection involves monitoring the behavior of users/customers in order to estimate, detect or avoid undesirable behavior in future. In this paper, we investigated the factors and various techniques involved in credit card fraud detection during/after transaction as well.

KEY TERMS: Data Mining, Credit Card. Fraud, Detection Tools

1. Introduction

Data mining uses Data Analysis tools to discover unknown, hidden and valid patterns as well as relationships in large data sets. Data mining tools include mathematical algorithms, statistical models, and machine learning methods such as algorithms which improve performance automatically through learning such as Neural Networks and Decision Trees. Data Mining consists of collection and management, analysis and prediction of corresponding data sets. Data mining can be performed on data sets represented in quantitative, textual or multimedia forms. On the other hand, Data mining applications can use a range of parameters to observe the data. Data mining applications include association rules, sequence or path analysis, classification methods, clustering and forecasting as well. Credit Card Fraud (CCF) is a typical task when using normal procedures, so the development of the credit card fraud detection model has become of significance whether in the academic or business community recently. These models are mostly statistics-driven or artificial intelligent-based which have the hypothetical advantages in not imposing random assumptions on the input variables [1]. Timely information on fraudulent activities information is a main goal and a good strategy for banks and industries as well. As banks have many and huge databases. Then sometimes, it is difficult to gain access to databases. Valuable business information can be extracted from data stores where data has been stored for time being. Credit card fraud detection is the process of identifying those transactions that are fraudulent and partitioned these database into two classes of legitimate (genuine) and fraudulent transactions. Credit card frauds can be further broadly classified into three categories, that is, traditional card related frauds (fake, application, stolen, account takeover and counterfeit), merchant related frauds and Internet frauds (site cloning, credit card generators and false merchant sites) [2].

Credit estimation is one of the essential and complex tasks for credit card companies, mortgage companies, banks and other financial institutes as well. Credit cards also offer a number of ancillary benefits unavailable from cash or checks. False credit judge ment causes huge financial losses. Credit cards also allows consumers to carry interest-free balances for approximately two months as the cardholder can carry the balance interest-free not only during the credit cycle but also even for a “grace period” of twenty or more days after the credit period ends [3].

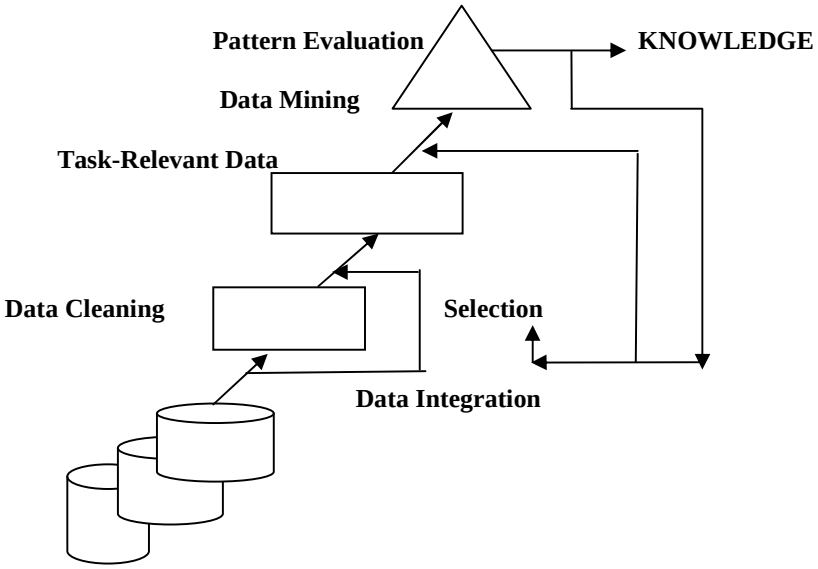

Fig. Data Mining (Knowledge Discovery Process)

Data Mining Techniques used in Credit Card Fraud Detection

I. Clustering

Among various data mining techniques, Clustering is a data mining technique that makes significant or useful cluster of object(s) that have similar characteristic using automatic technique. Apart from classification, clustering technique also defines the classes and put objects in them, although in classification, object(s) are assigned into predefined classes [4]. For example: In a library, books have ample variety of topics available. Challenging task is how to keep those books systematically that readers can take numerous books on a particular topic without disturbance. Hereby, by using clustering technique, we can keep books that have some similarity in one cluster or in one shelf and label them with a meaningful name. If in case readers want to take books on a topic, he/she would only go to that shelf instead of looking the complete in the whole library. Clustering is the method by which like records are grouped (cluster) together. Generally it is accomplished by giving the end user a high level view of what is going on in the database. Clustering is sometimes used to be alike as segmentation, in which most marketing people would tell you is more useful for coming up with a birds eye view of the business.

Bolton & Hand (2002) suggest two clustering techniques for behavioral fraud detection. Peer Group Analysis (PGA) is a system that allows identification of accounts that are behaving in a different way from others at one moment in time whereas they were behaving the same previously [5]. Those accounts are then flagged in suspicious activity.

Fraud Analysts (FA) have then to investigate those cases. The approach of the Peer Group Analysis (PGA) is that if accounts behave the same for a certain period of time and then one account is behaving considerably differently, this account must have to be notified. Another analysis technique as, Breakpoint analysis uses a different hypothesis. The approach is that if a change of card usage is notified on an individual basis, the account has to be investigated. In other words, we can say that, based on the transactions of a single card, the break-point analysis can identify suspicious behavior. Signals of suspicious behavior are indication of sudden transaction for a high amount and a high frequency of usage [6]. Clustering helps in grouping the data into similar clusters that helps in easy recovery of data. Cluster analysis is a method for breaking data down into connected components in such a way that patterns and order becomes observable.

Alike classification, clustering is the association of data in classes. Yet, in different classification and clustering, class labels are unknown and it depend on the clustering algorithm to determine acceptable classes. Clustering is also known as unsupervised classification because the classification is not dictated by given class labels. There are many clustering approaches that are based on the principle of maximizing the likeness between objects in a same

class (also known as intra-class similarity) and minimizing the similarity between objects of different classes (also known as inter-class similarity).
Conversely, there are some problems occurring in clustering are:
(a) Outline handling is difficult; the elements do not naturally lie into any cluster.
(b) Dynamic data in the database means that cluster membership may change over time.
(c) Interpreting the semantic meaning of each cluster can be difficult.

I.1. Classification of clustering algorithms

Classification may refer as gathering of different types of clustering algorithms. Clustering algorithms may also vary based on whether they produce overlapping or non-overlapping clusters. Non-overlapping clusters can be viewed as Extrinsic Clusters or Intrinsic Clusters.
Extrinsic technique/algorithms categorize the items to support in the classification process. Clustering algorithms are the traditional classification supervised learning algorithms that uses a special input training set. On the other side, intrinsic algorithms/techniques do not use ay priori category labels but depend only on the adjacency matrix containing the distance objects.

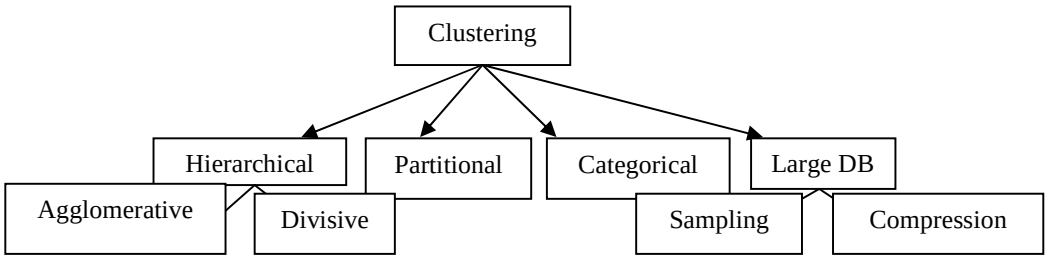

Fig: Classification of clustering algorithms

I.1.2. Clustering with Neural Networks

Neural Networks (NNs) that use unsupervised learning attempt to find features in the data that characterize the desired output. They look for clusters of like data. These types of NNs are often called Self-Organizing Neural Networks (SONN). There are two types of unsupervised learning: noncompetitive and competitive.
With the noncompetitive learning, the weight between two nodes is changed to be proportional to both output values. That is,
$\Delta w = \eta y^1 y^2$ [7]
With competitive learning, nodes are allowed to compete. This approach usually assumes a two-layer NN in which all nodes from one layer are connected to all nodes in the other layer. Thus, this provides a grouping of tuples together into a cluster.

I.1.3. CLUSTERING LARGE DATABASES

Clustering techniques should be able to adapt as the database changes. A clustering algorithm should have:
(i) Require no more than one scan of the database.
(ii) It should have the ability to provide status. This is sometimes referred to as the ability to be online.
(iii) It should be suspend able, stoppable and resume able.
(iv) It should process each tuple only once.

II. Neural Networks

A Neural Network (NN) is a collection of "processing nodes" transferring activity to each other via connections. Neural Networks (NN) have been successfully applied in a broad range of supervised and unsupervised learning applications. Neural Network (NN) learning algorithms that are capable to form logical models and that do not require extreme training times. Neural Networks (NN) topologies/architectures, has been formed by organizing nodes into layers and associate these layers of neurons with modifiable weighted interconnections. However, with a

nonlinear mapping relation from the input space to output space, Neural Networks (NN) can learn from the given cases and then summarize the internal principles of data even without knowing the potential data principles [8]. On the other side, Neural Network (NN) can easily familiarize its own behavior to the new environment with the corresponding results of formation of general capability of evolution from present situation to the new environment. In this approach, NN use multi-layer neural network model and Back Propagation (BP) algorithm) runs on the network. Back Propagation learns by iteratively processing a data set of training tuples comparing the network's prediction for each tuple with the real known target value. Weights are modified so as to minimize the mean squared error between the network's prediction and the actual target value for every training tuple. These modifications are made in the backwards direction, that is, from the output layer through each hidden layer down to the first hidden layer. ANN (Artificial Neural Network) refers to a group of non-linear, statistical modeling techniques derived from the structure of the human brain. ANN can be used in modeling of any complex transactional pattern, such that they are well suited to the credit card fraud detection problem [9]. Basic element of a Neural Network (NN) is a neuron which accepts many inputs, sums them, applies a transfer function and then generates corresponding result either as a model prediction or as input to other neurons. A Neural Network is a structure of many neurons connected in a regular way. The most well-known Neural Networks used are Feed-Forward Neural Networks (FFNN), which is also known as multilayer perceptrons.

Neural networks with supervised learning

The Feed-Forward Neural Networks (FFNN) can be used to represent a random non-linear mapping, being provided that we have data exemplifying mapping **as** Input-Output (I/O) pairs. Main problem of supervised learning is to adapt the neural network weights so that the input-mapping corresponds to the Input-Output samples which are being provided. For estimation, the density of past behavior in batch mode, we should retrieve the data from the last x days and adapt the mixing proportions to maximize the probability of past behavior. This could be done for each subscriber individually [10]. Though this approach seems first suited for the work being assigned, this requires too much interaction with the billing system to be used in practice.

Neural Network as a Classifier

However, Neural Network (NN) when seen as a classifier has some weaknesses and strength as well. Neural Network (NN) weakness involves:
(a) Long training time
(b) Require a number of parameters typically best determined empirically, that is the network topology.
(c) Poor interpretability: Difficult to understand the symbolic meaning behind the learned weights and of ``hidden units'' exists in the network.
NN has its own strengths:
(a) High tolerance to noisy data
(b)Ability to classify untrained patterns
(c)Well-suited for continuous-valued inputs and outputs
(d)Successful on a wide array of real-world data
(e)Algorithms are inherently parallel
(f)Techniques have recently been developed for the extraction of rules from trained neural networks

ISSN 2277-1956/V1N3-1765-1771
International Journal of Electronics and Computer Science Engineering
WWW.IJECSE.ORG

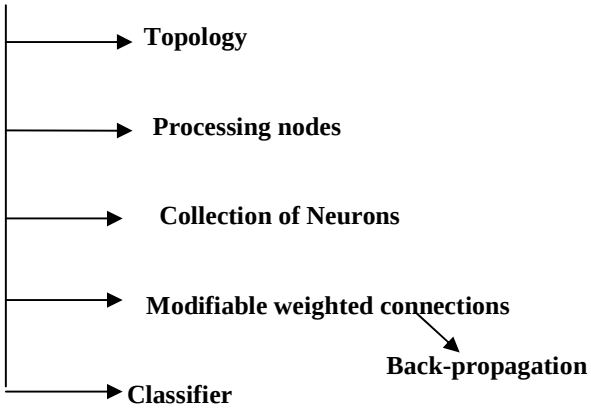

Fig: - Framework for Neural Network(s)

III. Bayesian Classification:

Another Data Mining technique used for identification of suspicious activity in between large datasets is Bayesian Classification. As there are no such deterministic rule which allows us to identify a subscriber as a fraudster, Bayesian networks can be used as an expert system [11]. This refers that an expert of the problem domain draws a graph according to assumed causal impacts between variables. The resultant conditional distributions can then be injected by the expert **as** well. Once a Bayesian network is set up, we can conclude probabilities for unknown variables by inserting evidence in the network and propagating evidence through the network using propagation rules. While, a statistical classifier performs probabilistic prediction that means classifier predicts class membership probabilities. There is Baye's Theorem which interprets Bayesian networks and classifier as well. Bayes classifier has some benefits and characteristics as well:

(i) **Performance:** Simple Bayesian classifier, Naive Bayesian classifier has comparable performance with decision tree and selected neural network classifiers

(ii) Incremental: Each training example can incrementally increase/decrease the probability that the particular hypothesis is correct prior to knowledge of past observations.

(iii) Standard: Even when Bayesian methods are computationally inflexible, Bayesian Classifiers can provide a standard of optimal decision making against which other methods can be measured. Bayesian belief network allows a subset of the variables for being conditionally independent.

IV. Fuzzy Darwinian Detection of Credit Card Fraud

Nowadays, Fraud is a big problem today. Genetic programming evolves fuzzy logic rules capable of classifying credit card transactions into "suspicious" and "non-suspicious" classes. When took notice of credit card transactions alone, with million(s) of purchases every month, it is simply not possible to check every one individually. Whenever many purchases are made with stolen credit cards, this unavoidably results in losses of significant sums. Through the multimodal and multi criteria, search space is guided by fitness functions. These fitness functions use the results formed by the Rule Parser [12]. Fuzzy expert system that takes more than one rules and interpret their meaning when they are applied to each of the previously fuzzified data items in turn. This system should be capable of two different types of fuzzy logic rule interpretation: traditional fuzzy logic and membership-preserving fuzzy logic.. Depending on the method of interpretation that has been selected by the user the meaning of the operators within rules and the method of defuzzification is different.

V. RELATED WORK

Fraud detection tools

Fraud detection can be classified as 'supervised' or 'unsupervised'. Supervised methods uses database of known fraudulent/legitimate cases from which model has been constructed which yields a suspicion score for new cases when some different score for which past behavior has been analyzed. (Hand, 1981; McLachlan, 1992) demonstrated traditional statistical classification methods such as linear discriminant analysis and logistic discrimination that have proved to be effective tools for many applications but more powerful tools (Ripley, 1996; Hand, 1997; Webb, 1999) such as neural networks that have also been extensively applied.

Supervised learning algorithms are Rule-based methods that produce classifiers using rules of the form:

If {certain conditions},

Then {a consequent}

Some of the examples of such algorithms include BAYES (Clark and Niblett, 1989), FOIL (Quinlan 1990) and RIPPER (Cohen 1995), Tree-based algorithms such as CART (Brieman *et al*, 1984) and C4.5 (Quinlan 1993) that produce classifiers of a similar form. Combinations of some or all of these algorithms can be used as meta-learning algorithms which improve prediction in fraud detection that is, Chan *et al* (1999). When building a supervised tool for fraud detection major consideration includes those of uneven class sizes and different costs for different types of misclassification [13]. On the other side, there must be some consideration for the costs of investigating observations and corresponding benefits of identifying fraud.

Unsupervised methods, are used when there are no prior sets of legitimate and fraudulent observations available. Some techniques that are being employed are usually a combination of both profiling and outlier detection methods. There exists model in which baseline distribution that represents normal behavior, then attempt to detect remark that show greatest different behavior from this existing norm. Benford's law (Hill 1996) said that the distribution of the first significant digits of numbers drawn from a vast range of arbitrary distributions would have some certain form. Until currently, this law was regarded as merely a mathematical curiosity with no noticeable useful application. Nevertheless, Nigrini and Mittermaier (1997) and Nigrini (1999) show that Benford's law can be used to detect fraud in accounting data. The assertion behind fraud detection using tools such as Benford's law is that fabricating data which is conventional to Benford's law is difficult [14]. Fraudster(s) familiarize new prevention and detection measures as well so that fraud detection needs to be adaptive over time. Still, legitimate account users may regularly change their behavior over a longer period of time and it is significant to avoid false alarms.

VI. CONCLUSION

Efficient and well-organized credit card fraud detection system is an greatest requirement for any card issuing bank. Credit card fraud detection has drawn quite a lot of interest from the research community and a number of techniques have been proposed to counter/identify credit card fraud. The Fuzzy Darwinian fraud detection systems improve the system accuracy, while neural network improve the method time to detect particular fraud termed as suspicious activity. Since the Fraud detection rate of Fuzzy Darwinian fraud detection systems in terms of true positive is 100% and shows good results in detecting fraudulent transactions on the other side, the neural network based CARDWATCH shows good accuracy in fraud detection and Processing Speed is also high but it is limited to one-network per customer. The Fraud detection rate of using Clustering is very compare to other methods. As usage of credit cards become more and more popular in every field of the daily life, credit card fraud has become much more rampant. Therefore, there is a need for improving security of the financial transaction systems in an automatic and effective way, by building an accurate and efficient credit card fraud detection system. As, it is the key task for the financial institutions. In this study, we gather various methods that were used to build fraud detecting models. Currently, due to the security issues, only a few approaches for credit card detection are available in public. In between them, neural networks approach is a very popular tool. Though, it is difficult to implement because of lack of available data set.

REFERENCES

[1] A. Shen, R. Tong, and Y. Deng, "Application of classification models on credit card fraud detection," June 2007.

[2] Brause, R., Langsdorf, T. and Hepp, M. (1999). Neural data mining for credit card fraud detection *Proceedings 11th IEEE International Conference on Tools with Artificial Intelligence*.TAO GUO, GUI-YANG LI, NEURAL DATA MINING FOR CREDIT CARD FRAUD DETECTION 978-1-4244-2096-4/08 ©2008 IEEE, 3630, July 2008.

[3] Chen, R.-C., Luo, S.-T., Liang, X., Lee, V. C. S.: Personalized approach based on SVM and ANN for detecting credit card fraud. In: Proceedings of the IEEE International Conference on Neural Networks and Brain, Beijing, China (2005).

[4] Dahl, J.: Card Fraud. In: Credit Union Magazine (2006).

[5] Dorronsoro, Ginel, Sgnchez and Cruz. Neural fraud detection in credit card operations. Neural Networks, *IEEE Transactions.* Volume: 8, Issue: 4: 827-834, 1997.

[6] Mirjana Pejic-Bach, Profiling intelligent systems applications in fraud detection and prevention: survey of research articles, 2010 International Conference on Intelligent Systems, Modeling and Simulation.

[7] Prabin Kumar Panigrahi, A Framework for Discovering Internal Financial Fraud using Analytics, International Conference on Communication Systems and Network Technologies 2011

[8] Raghavendra Patidar, Lokesh Sharma, "Credit Card Fraud Detection Using Neural Network", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1, Issue-NCAI2011, June 2011.

[9] [1]S. Benson Edwin Raj, [2]A. Annie Portia Analysis on Credit Card Fraud Detection Methods ICCCET2011, 18th & 19th March, 2011 978-1-4244-9394-4/11/$26.00 ©2011 IEEE] 152

[10] Sahin, Y., Duman, E.: An overview of business domains where fraud can take place, and a survey of various fraud detection techniques. In: Proceedings of the 1st International Symposium on Computing in Science and Engineering, Aydin, Turkey (2010).

[11] Tej Paul Bhatla, Vikram Prabhu & Amit Dua "Understanding Credit Card Frauds," 2003.

[12] V. Filippov L. Mukhanov B. Shchukin Credit Card Fraud Detection System.

[13] Y. Sahin, E. Duman "Detecting Credit Card Fraud by ANN and Logistic Regression" 2011