Analysis on Credit Card Fraud Detection Methods

[1]Renu
HCE Sonepat
[2] Suman
HCE Sonepat

Abstract

Due to the theatrical increase of fraud which results in loss of dollars worldwide each year, several modern techniques in detecting fraud are persistently evolved and applied to many business fields. Fraud detection involves monitoring the activities of populations of users in order to estimate, perceive or avoid undesirable behavior. Undesirable behavior is a broad term including delinquency, fraud, intrusion, and account defaulting. This paper presents a survey of current techniques used in credit card fraud detection and telecommunication fraud. The goal of this paper is to provide a comprehensive review of different techniques to detect fraud.

Keywords: Fraud detection, data mining, Neural Network, anomalies.

INTRODUCTION

Credit card fraud can be defined as "Unauthorized account activity by a person for which the account was not intended. Operationally, this is an event for which action can be taken to stop the abuse in progress and incorporate risk management practices to protect against similar actions in the future". In simple terms, Credit Card Fraud is defined as when an individual uses another individual's credit card for personal reasons while the owner of the card and the card issuer are not aware of the fact that the card is being used. And the persons using the card has not at all having the connection with the cardholder or the issuer and has no intention of making the repayments for the purchase they done. Fraud detection involves identifying Fraud as quickly as possible once it has been perpetrated. Fraud detection methods are continuously developed to defend criminals in adapting to their strategies. The development of new fraud detection methods is made more difficult due to the severe limitation of the exchange of ideas in fraud detection. Data sets are not made available and results are often not disclosed to the public. The fraud cases have to be detected from the available huge data sets such as the logged data and user behavior. At present, fraud detection has been implemented by a number of methods such as data mining, statistics, and artificial intelligence. Fraud is discovered from anomalies in data and patterns. The different types of methods for committing credit card frauds are described below.

Types of Frauds

Various types of frauds in this paper include credit card frauds, telecommunication frauds, and computer intrusions, Bankruptcy fraud, Theft fraud/counterfeit fraud, Application fraud, Behavioral fraud [2].

1) Credit Card Fraud*:* Credit card fraud has been divided into two types: Offline fraud and On-line fraud. Offline fraud is committed by using a stolen physical card at call center or any other place. On-line fraud is committed via internet, phone,

shopping, web, or in absence of card holder.

2) Telecommunication Fraud*:* The use of telecommunication services to commit other forms of fraud. Consumers, businesses and communication service provider are the victims.
3) *Computer* Intrusion: Intrusion Is Defined As The act of entering without warrant or invitation; That means "potential possibility of unauthorized attempt to access Information, Manipulate Information Purposefully. Intruders may be from any environment, An outsider (Or Hacker) and an insider who knows the layout of the system [1].
4) Bankruptcy Fraud*:* This column focuses on bankruptcy fraud. Bankruptcy fraud means using a credit card while being absent. Bankruptcy fraud is one of the most complicated types of fraud to predict [1].
5) Theft Fraud/ Counterfeit Fraud: In this section, we focus on theft and counterfeit fraud, which are related to one other. Theft fraud refers using a card that is not yours. As soon as the owner give some feedback and contact the bank, the bank will take measures to check the thief as early as possible. Likewise, counterfeit fraud occurs when the credit card is used remotely; where only the credit card details are needed [2].
6) Application Fraud*:* When someone applies for a credit card with false information that is termed as application fraud. For detecting application fraud, two different situations have to be classified. When applications come from a same user with the same details, that is called duplicates, and when applications come from different individuals with similar details, that is termed as identity fraudsters. Phua et al. [3] describes application fraud as "demonstration of identity crime, occurs when application forms contain possible, and synthetic (identity fraud), or real but also stolen identity information (identity theft)."credit card fraud detection techniques.

Techniques for Fraud Detection

BAYESIAN NETWORKS

For the purpose of fraud detection, two Bayesian networks to describe the behavior of user are constructed. First, a Bayesian network is constructed to model behavior under the assumption that the user is fraudulent (F) and another model under the assumption the user is a legitimate (NF). The 'fraud net' is set up by using expert knowledge. The 'user net' is set up by using data from non fraudulent users. During operation user net is adapted to a specific user based on emerging data. By inserting evidence in these networks and propagating it through the network, the probability of the measurement x less than two above mentioned hypotheses is obtained. This means, it gives judgments to what degree observed user behavior meets typical fraudulent or non fraudulent behavior. These quantities we call p(X | NF) and p (X | F).
By postulating the probability of fraud P (F) and P (NF) =1-P(F) in general and by applying Bayes' rule, it gives the probability of fraud, given the measurement x,

$$P(F|X) = \frac{p(F)P(X|F)}{P(F|X)} \quad \text{----- - (1)}$$

Where the denominator p(x) can be calculated as

P(x) = P (F) p (X | F) + P (NF) p (X | NF) --------- (2)

The fraud probability P (F | X) given the observed user behavior x can be used as an alarm level. On the one hand, Bayesian networks allow the integration of expert knowledge, which we used to initially set up the models [4]. On the other hand, the user model is retrained in an unsupervised way using data. Thus our Bayesian approach incorporates both, expert knowledge and learning.

HIDDEN MARKOV MODEL

A Hidden Markov Model is a double embedded stochastic process with used to model much more complicated stochastic processes as compared to a traditional Markov model. If an incoming credit card transaction is not accepted by the trained Hidden Markov Model with sufficiently high probability, it is considered to be fraudulent transactions. HMM[5], Baum Welch algorithm is used for training purpose and K-means algorithm for clustering.HMM sores data in the form of clusters depending on three price value ranges low, medium and high[6]. The probabilities of initial set of transaction have chosen and FDS checks whether transaction is genuine or fraudulent. Since HMM maintains a log for transactions it reduces tedious work of employee but produces high false alarm as well as high false positive[7].In this fig1 shows state diagram of HMM,Which shows different stages and iterations.

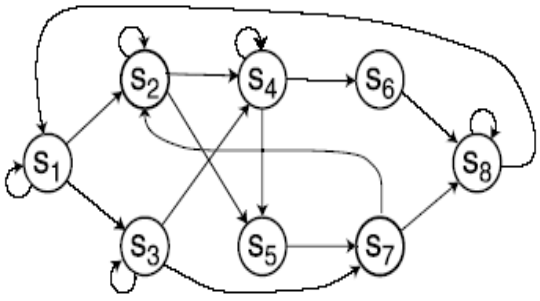


Fig1: State Diagram

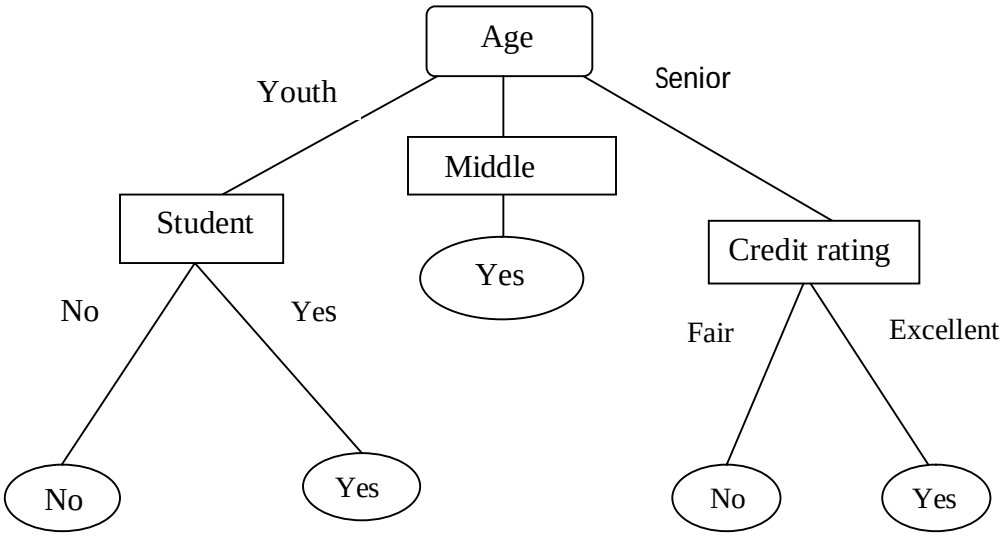
GENETIC ALGORITHM

Genetic algorithms, inspired from natural evolution were first introduced by Holland (1975). Genetic algorithms are evolutionary algorithms which aim at obtaining better solutions as time progresses**.** Fraud detection problem is classification problem, in which some of statistical methods many data mining algorithms have proposed to solve it. Among decision trees are more popular. Fraud detection has been usually in domain of E-commerce, data mining [8].

GA is used in data mining mainly for variable selection [9] and is mostly coupled with other DM algorithms [10]. And their combination with other techniques has a very good performance. GA has been used in credit card fraud detection for minimizing the wrongly classified number of transactions [10]. And is easy accessible for computer programming language implementation, thus, make it strong in credit card fraud detection. But this method has high performance and is quite expensive.

DECISION TREE

Decision trees are statistical data mining technique that express independent attributes and a dependent attributes logically AND in a tree shaped structure. Classification rules, extracted from decision trees, are IF-THEN expressions and all the tests have to succeed if each rule is to be generated [11]. Decision tree usually separates the complex problem into many simple ones and resolves the sub problems



through repeatedly using [11][12]. Decision trees are predictive decision support tools that create mapping from observations to possible consequences. There are number of popular classifiers construct decision trees to generate class models. . The most well-know algorithm in the literature for building decision trees is the C4.5. This is an extension of ID3 algorithm. Decision trees can be translated into a set of rules by creating a separate rule for each path from the root to a leaf in the tree.

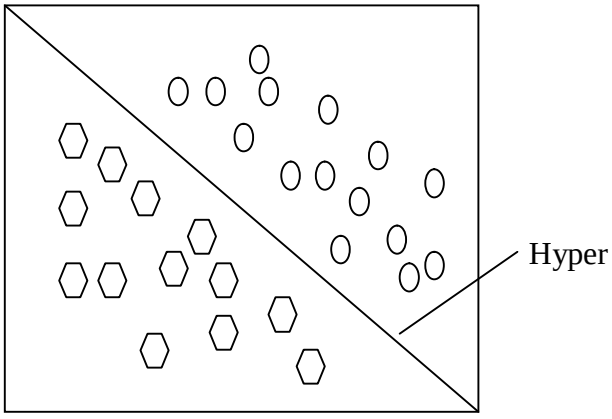
Each internal (non leaf) node represents a test on an attribute. Each leaf node represents a class (either buys computer = yes or buys computer = no).

Decision tree methods C5.0,C&RT and CHAID. The work demonstrates the advantages of applying the data mining techniques including decision trees and SVMs to the credit card fraud detection problem for the purpose of reducing the bank's risk. The results show that the proposed classifiers of C&RT and other decision tree approaches outperform SVM approaches in solving the problem under investigation.

SUPPORT VECTOR MACHINE

The basic idea of SVM classification algorithm is to construct a hyper plane as the decision plane which making the distance between the positive and negative mode maximum [17]. The strength of SVMs comes from two important properties they possess - kernel representation and margin

optimization. Kernels, such as radial basis function (RBF) kernel, can be used to learn complex regions. A kernel function represents the dot product of projections of two data points in a high dimensional feature space. In SVMs, the classification function is a hyper-plane separating the different classes of data. The basic technique finds the smallest hypersphere in the kernel space that contains all training instances, and then determines on which side of hypersphere a test instance lies. If a test instance lies outside the hypersphere, it is confirmed to be suspicion. SVM can have better prediction performance than BPN(Back propagation network) in predicting the future data.



This shows the hyperplane which classify the data from one class to another class

NEURAL NETWORK

Neural networks have been widely used in fraud detection. Neural network is a set of Connected input/output units and each connection has a weight present with it. During the learning phase, network learns by adjusting weights to predict the correct class labels .

Fraud detection methods based on neural network are the most popular ones. An artificial neural network [13][14] consists of an interconnected group of artificial neurons .The principle of neural network is motivated by the functions of the brain especially pattern recognition and associative memory [15]. The neural network recognizes similar patterns, predicts future values or events based upon the associative memory of the patterns it was learned. It is widely applied in classification and clustering. The advantages of neural networks over other techniques are that these models are able to learn from the past and thus, improve results as time passes. They can also extract rules and predict future activity based on the current situation. By employing neural networks, effectively, banks can detect fraudulent use of a card, faster and more efficiently. Among the reported credit card fraud studies most have focused on using neural networks. In more practical terms neural networks are non-linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data.

There are two phases in neural network [16] training and recognition. Learning in a neural network is called training. There are two types of NN training methods supervised and unsupervised. In supervised training, samples of both fraudulent and non fraudulent records are used to create models. In contrast, unsupervised training simply seeks those transactions, which are most dissimilar from the norm. On other hand, the unsupervised techniques do not need the previous knowledge of fraudulent and non fraudulent transactions in database. NNs can produce best result for only large transaction dataset. And they need a long training dataset. In this Fig shows Single layer feedforward model,which shows weight, input and output.

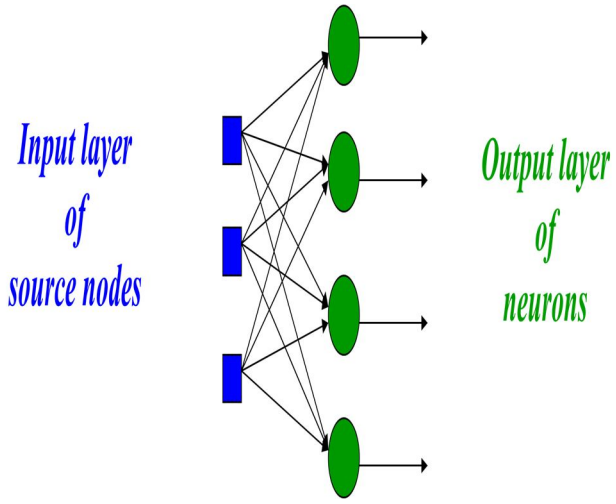


Fig : Single Layer Feed Forward Model

CONCLUSION

Credit card fraud has become more and more rampant in recent years. To improve merchants' risk management level in an automatic and effective way, building an accurate and easy handling credit card risk monitoring system is one of the key tasks for the merchant banks. One aim of this study is to identify the user model that best identifies fraud cases. There are many ways of detection of credit card fraud. If one of these or combination of algorithm is applied into bank credit card fraud detection system, the probability of fraud transactions can be predicted soon after credit card transactions by the banks. And a series of anti-fraud strategies can be adopted to prevent banks from great losses before and reduce risks. This paper gives contribution towards the effective ways of credit card fraudulent detection.

References

[1]. Linda Delamaire (UK), Hussein Abdou (UK), John Pointon (UK), "Credit card fraud and detection techniques: a review", Banks and Bank Systems, Volume 4, Issue 2, 2009 .

[2]. Khyati Chaudhary, Jyoti Yadav, Bhawna Mallick, "A review of Fraud Detection Techniques: Credit Card", International Journal of Computer Applications (0975 – 8887) Volume 45– No.1, May 2012 .

[3] .Vladimir Zaslavsky and Anna Strizhak," credit card fraud detection using selforganizing maps", information & security. An International Journal, Vol.18,2006.

[4] L. Mukhanov, "Using bayesian belief networks for credit card fraud detection," in Proc. of the IASTED International conference on Artificial Intelligence and Applications, Insbruck, Austria, Feb. 2008, pp. 221– 225.

[5] Abhinav Srivastava, Amlan Kundu, Shamik Sural and Arun K. Majumdar, "CreditCard Fraud Detection Using Hidden Markov Model" IEEE, Transactions On Dependable And Secure Computing, Vol. 5, No 1. , January-March 2008

[6] V. Bhusari, and S. Patil, "Study of Hidden Markov Model in Credit Card Fraudulent Detection", International Journal of Computer Applications (0975 – 8887) Volume 20– No.5, April 2011

[7] V.Bhusari ,S.Patil ," Study of Hidden Markov Model in Credit Card Fraudulent Detection ",International Journal of Computer Applications (0975 - 8887) Volume 20- No.5, April 2011

[8] K.RamaKalyani, D.UmaDevi, "Fraud Detection of Credit Card Payment System by Genetic Algorithm", International Journal of Scientific & Engineering Research Volume 3, Issue 7, July-2012

[9] Bidgoli, B. M., Kashy, D., Kortemeyer, G. & Punch, W. F "Predicting student performance: An Application of data mining methods with the educational web-based system LON-CAPA". In Proceedings of

ASEE/IEEE frontiers in education conference. . (2003).
[10] Ekrem Duman, M. Hamdi Ozcelik "Detecting credit card fraud by genetic algorithm and scatter search". Elsevier, Expert Systems with Applications, (2011). 38; (13057–13063).
[11] S. Benson Edwin Raj, A. Annie Portia, "Analysis on Credit Card Fraud Detection Methods", International Conference on Computer, Communication and Electrical Technology – ICCCET2011, 18th & 19th March, 2011
[12] Y. Sahin and E. Duman, "Detecting Credit Card Fraud by Decision Trees and Support Vector Machines", International Multiconference of Engineers and computer scientists March, 2011.
[13] S. Benson Edwin Raj, A. Annie Portia "Analysis on Credit Card Fraud Detection Methods". IEEE-International Conference on Computer, Communication and Electrical Technology; (2011). (152-156).
[14] Ray-I Chang, Liang-Bin Lai, Wen-De Su, Jen-Chieh Wang, Jen-Shiang Kouh "Intrusion Detection by Backpropagation Neural Networks with Sample-Query and Attribute-Query". Research India Publications; (2006). (6-10).
[15] Raghavendra Patidar, Lokesh Sharma "Credit Card Fraud Detection Using Neural Network". International Journal of Soft Computing and Engineering (IJSCE), (2011). Volume-1, Issue; (32-38).
[16] Tao Guo, Gui-Yang Li "Neural Data Mining For Credit Card Fraud Detection". IEEE, Proceedings of the Seventh International Conference on Machine Learning and Cybernetics; (2008). (3630-3634).
[17] Joseph King-Fung Pun,"Improving Credit Card Fraud Detection using a Meta-Learning Strategy", Chemical Engineering and Applied Chemistry University of Toronto 2011