International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

IJIEMR Transactions, online available on 12th Nov 2018. Link :http://www.ijiemr.org/downloads.php?vol=Volume-07&issue=ISSUE-12

Title: **REVIEW ON VARIOUS METHODS FOR FRAUD TRANSACTION DETECTION IN CREDIT CARDS:**

Volume 07, Issue 12, Pages: 75–81.

Paper Authors

HARDIK MANEK,NIKHIL KATARIA,SUJAI JAIN,PROF.CHITRA BHOLE
K.J.S.I.E.I.T

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

REVIEW ON VARIOUS METHODS FOR FRAUD TRANSACTION DETECTION IN CREDIT CARDS:

[1]HARDIK MANEK, [2]NIKHIL KATARIA, [3]SUJAI JAIN, [4]PROF. CHITRA BHOLE

Department of Computer Engineering, K.J.S.I.E.I.T

[1]h.manek@somaiya.edu, [2]nikhil.kataria@somaiya.edu, [3]sujai.jain@somaiya.edu,
[4]cbhole@somaiya.edu

Abstract— with the advent of new technologies and the Internet, cashless transactions have become simpler. But for online transactions, you do not need to be personally present in a certain place where the transaction occurs, making it vulnerable to fraudulent attacks. Several ways in which you can pretend to be the user and make a fraudulent transaction. If a transaction is fraudulent or not, it can be determined by analyzing previous transactions and comparing them with the current one. If the difference in nature of previous transactions and the current transaction is large, there is a possibility that the current transaction is a fraudulent transaction. Bank and credit card companies use different methods to detect frauds such as the genetic algorithm, the neural network, the nearest neighborhoods algorithm, etc. the methods analyze the client's past behavior and make decisions. In this document a comparative analysis of different algorithms will be performed**.**

Keywords:-credit card fraud, cashless transactions, Genetic Algorithm, Neural Network

I. INTRODUCTION

The payment card industry has grown rapidly in recent years. Regardless of location, consumers can do the same shopping as before on the desk. The evolution is a big step forward for efficiency, accessibility and the point of view of profitability, but also has some drawbacks. Evolution is accompanied by greater vulnerability to threats. The problem of doing business through the Internet lies in the fact that neither the card nor the cardholder. You must be present at the point of sale. So it is impossible for the trader to check if the customer. It is the actual holder of the card or not. Financial institutions have attention focused on recent computational methods for managing the problem of credit card fraud. Credit card fraud detection is the process of identifying those transactions that are fraudulent in two types of legitimate (genuine) and fraudulent transactions. Credit card fraud detection is based on the analysis of a card's spending behavior. The features are extracted and transformed from raw data as it is given to train model. This document compares different techniques. Features are extracted and transformed from raw data while giving it to train model [2]. Many techniques have been applied to credit card fraud detection, machine learning algorithms[1], artificial neural network , genetic algorithm, support vector machine , frequent item set mining , decision tree , migratingBirds optimization algorithm. Bayesian performance and the neural network are evaluated in data on credit card fraud. Decision tree, neural networks and logistic regression have demonstrated its applicability in fraud detection.

II. LITERATURE SURVEY

Ghosh and Reilly [3] used three-layer neural networks to detect fraud in 1994. The neural network was trained on examples of fraud containing stolen cards, application fraud, counterfeit fraud, non-received fraud problems (NRIs) and orders for post fraud.

Abhinav and Amlan [4] proposed a hidden Markov model to detect credit card fraud. The proposed model does not require fraudulent signatures and can still detect frauds considering the cardholder’s spending habits.

Y. Sahin and E. Duman [5] proposed an approach to identify credit card fraud using the decision tree and the Support Vector Machine. The different methods of tree decision- making performance intensifier classification models (C5.0, C and RT and CHAID) and several other SVM methods (SVM with polynomial, sigmoidal, linear and RBF kernels) are compared in this study.An approach is proposed towards fraud detection in banking transactions in [6] using fuzzy clustering and neural network. In this approach, fraud detection is performed in three phases. The first step is initial user authentication and verification of card details. After completion of this operation, a blurry half clustering algorithm is performed to discover the behavior of normal user usage based on past transactions. If it turns out that a new transaction is uncertain at this stage, based on a neural network to determine whether it was in fact a fraudulent transaction or the mechanism applies.

Kang Fu, Dawei Cheng, Yi Tu, and Liqing Zhang at [7] proposed a convivial based neural network approach (CNN) to find fraudulent transactions. Convolutional neural network is a part of deep learning and is a type of advanced neural network composed of more than one hidden layer. In this document, to find more complex fraud models and improve the accuracy of classification, a new feature of commercial entropy is proposed. In this document for the first time, CNN is used to detect fraud.Different outlier techniques [8] can also use to differentiate fraudulent transaction as outlier dataThe description of the various techniques are as follows:-

1.) Credit Card Fraud Detection with a Neural Network by Sushmito Ghosh (IEEE 1994) [3] with the techniques, feed forward artificial Neural Network and data sets from Mallon bank 450000 transactions to train model.

2.) Credit Card Fraud Detection using Bayesian and Neural Networks by Sam Maes (International Naiso Congress on Neuro-Fuzzy Technology, 2002)[11] with the technique, Bayesian Neural networks and data sets from Euro pay International.

3.) Credit Card Fraud Detection using Hidden Markov model [4] by Avhinav Srivastava (IEEE Dep and Sec Comp, 2002) with the technique, Hidden Markov Model and data sets from completely simulated and simplified data.

4.) Detecting Credit Card by Decision Trees and Support Vector Machines [5] by Y. Sahin (Proc Int. MultiConf of Eng and Comp Sci 2011) with the techniques Decision Tree (C5.0, C and RT and CHAID) SVM (polynomial sigmoid, linear and RBF kernel functions) and data sets from National bank Credit card data warehouse 978 fraud, 22 million normal transactions.

5.) Credit Card Fraud Detection using Convolutional Neural Network [6] by Tanmay kumar Behera (IEEE ComputerSociety, 2015) with the technique, Fuzzy Clustering and Neural Network using Synthetic data.

6.) Credit Card Fraud Detection using Convolutional Neural Network by Kang Fu [7] (Springer, 2016 Convolutional Neural Network, Cost-Based Sampling for imbalance data) with the technique Commercial Bank and data sets from 260 million transactions and 4000 fraud.

III. ISSUES WITH CREDIT CARD FRAUD DETECTION

The main issue with creating a Mastercard fraud detection system is getting information for coaching. It is difficult to induce real information as a result of this type of information is sensitive and personal. In several techniques [7],[3],[5],[11], the researches have trained with real-world information by arrival with banks. But otherwise, synthetic information are often generated and is accessible for coaching.

Second issue is to contend with the distinction between varieties the amount the quantity of legitimate and therefore the number of deceitful transactions. Synthetic minoring over-sampling ways square measure accustomed increase range of low incidence information in information set that generate artificial deceitful transactions connected with original dataset. In [7], value primarily based sampling is employed to get artificial deceitful transactions to balance information set.

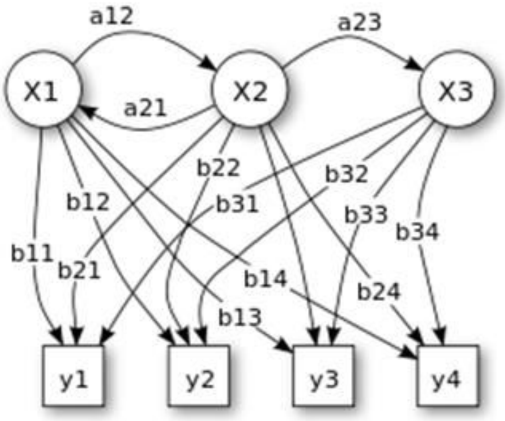
Overlapping of information is a new drawback as a number of dealings seem like deceitful transaction, once truly they're legitimate transactions it's conjointly doable that deceitful transactions seem to be traditional transactions.

IV. VARIOUS TECHNIQUES FOR FRAUD DETECTION

A. HIDDEN MARKOV MODEL

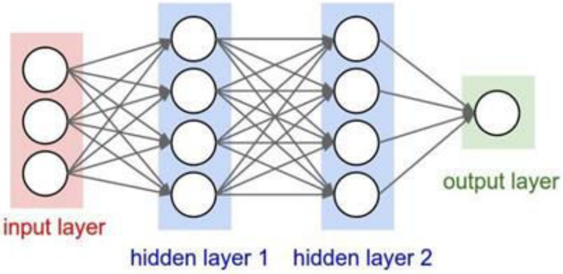
A hidden Markov model of science model (HMM) is also a maths Markov model inside that the system being sculptural is assumed to be a Markov chain with hidden states academic degree HMM is also a double embedded likelihood distribution technique with hierarchy levels. Fraud detection Approach victimization HMM is projected. They need thought-about three price ranges low, medium and high asset of potential observation, as an example, let l=(0,200USD],m=(USD250,USD600],h=(U

SD700,credit card limit). If a user makes a bunch action of USD 400, then resultant observation image are medium. Every human action amount generally depends on the equivalent kind of purchase. The set of all potential sorts of purchase and also the set of all potential lines of business of merchants forms the set of hidden states of the HMM. The projected approach in [7], Hidden Andre Markoff Model (HMM) - based mastercard FDS does not require fraud signatures and still it can detect frauds by considering a users spending pattern



B. ARTIFICIAL NEURAL NETWORKS

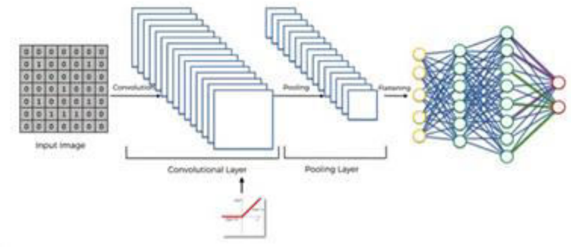
Artificial Neural Network (ANN) is one of the most powerful classifiers to look out hidden patterns among completely totally different attributes. ANN works same as human brain. ANN consists of varied layers throughout that first layer is input layer and last layer is output layer. It ought to have type of hidden layer or no hidden layer. If Neural network embrace quite one hidden layer, then it's deep learning each layer has completely totally different neurons, and every somatic cell is connected with weighted edges. Output of each somatic cell could also be a performance of its unit. This performance is called activation perform. Example of varied activation functions used square measure sigmoid performs, step performs, function, linear performs etc.



Most used to perform is Sigmoid perform among all output layer has the same type of neurons as classification label, each vegetative cell of output layer offers likelihood of being that category. In the figure, four neurons square measure in input layer that creates five picks regarding to importance of input options. Neurons of second layer connected to output layers neurons. Neural network makes an alternative of weights on edges from data given thereto for employment and regulate weights.

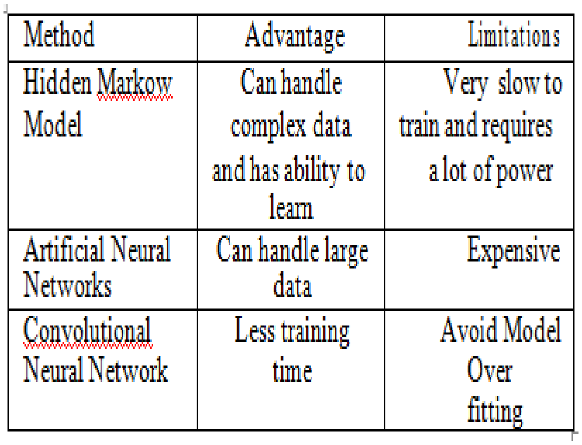
C. CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Network (CNN) is also a section of deep learning. Mapping of input into hidden layer represents one feature map each feature map represents one characteristic method of press neurons into feature map is termed convolution as shown in the figure below. Sub- sampling reduces parameters of feature map fully connected layer is same as neural network[11].



CNN is with success applied in face recognition, character recognition, image classification, etc. Kang Fu, Dawei Cheng, Yi Tu, and Liqing Zhang[7] at projected a convolutional neural network primarily based approach to find fallacious transactions in MasterCard. Input options square measure reworked into feature matrices so reborn into pictures. For finding additional complicated fraud patterns and to enhance classification accuracy, replacement feature commerce entropy is projected to alleviate the matter of the unbalanced dataset; they used price primarily based sampling technique to get completely different range of artificial frauds to coach the model. They applied CNN model as a result of it's appropriate for coaching giant size of information and CNN has mechanism to avoid over fitting.

COMPARISON TABLE



Method	Advantage	Limitations
Hidden Markow Model	Can handle complex data and has ability to learn	Very slow to train and requires a lot of power
Artificial Neural Networks	Can handle large data	Expensive
Convolutional Neural Network	Less training time	Avoid Model Over fitting

V. PROPOSED SYSTEM

A. AUTO ENCODER NEURAL NETWORK

Autoencoders are a type of Neural Network which is used to learn data codings in an unsupervised manner. The main functionality of the encoder model is to find an appropriate method to encode the data such that decoded data will be close to the input data. Autoencoder model consists of 3 main layers which are the input layer, an output layer and one or more hidden layer connecting them, although the number of nodes in the output layer are same as that of input layer.

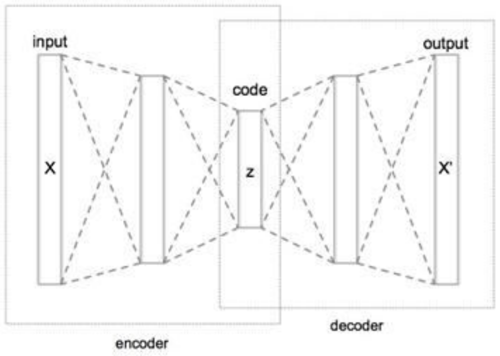
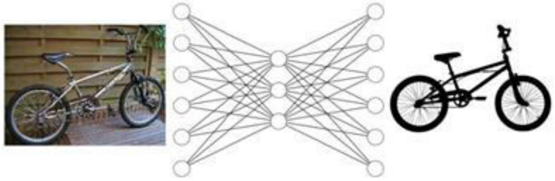


Figure Structure of an Autoencoder Model with 3 connected hidden layers which encodes input x and fives output x(Z)f W.b(x) = x The Reconstruction error is minimized by using traditional squared error given by: $L(x,x) = ||xx||^2$



A Simple demonstration of Autoencoder model is given in above figure in which first part that is input layer and hidden layer encodes a bicycle and subsequently, the 2nd part i.e. outer layer and hidden layer decodes to achieve similar bicycle as the output. Yimin Yang, Q.M.Jonathan Wu, Yaonan Wang [10] have proposed Autoencoder model for dimension reduction and image reconstruction

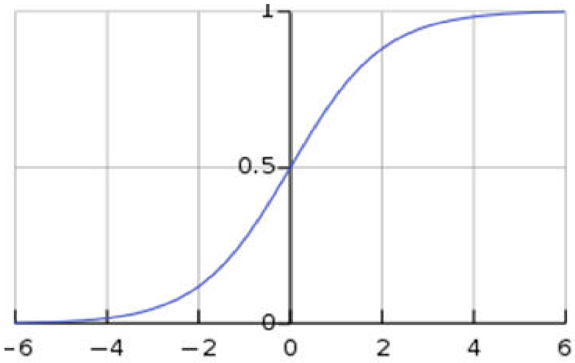
B. LOGISTIC REGRESSION

Logistic Regression is the most renowned machine learning calculation after Linear Regression. From multiple pointsof view, Linear Regression and Logistic Regression are comparative. Be that as it may,the greatest contrast lies in what they are utilized for. Linear Regression calculations are utilized to anticipate conjecture esteems however Logistic Regression is utilized for grouping assignments.There are numerous arrangement assignments done routinely by individuals. For instance, arranging whether an email is a spam or not, characterizing whether a tumor is harmful or kindhearted, ordering whether a site is fake or not, and so on. These are run of the mill precedents where machine learning calculations can make our lives significantly simpler. An extremely basic, basic and valuable calculation for characterization is the Logistic Regression calculation.Sigmoid Function (Logistic Function) Calculated relapse calculation likewise utilizes a direct condition with free indicators to anticipate an esteem. The anticipated esteem can be anyplace between negative interminability to positive vastness. We require the yield of the calculation to be class variable, i.e. 0-no, 1-yes. Along these lines, we are squashing the yield of the straight condition into a scope of [0,1]. To squash the anticipated an incentive somewhere in therangeof0and1, we utilize the sigmoid capacity.

$$z = \theta_0 + \theta_1.x_1 + \theta_2.x_2 + \ldots.$$

$$h = g(z) = 1/(1 + e^{-z})$$

We take the output (z) of the straight condition and provide for the capacity g(x) which restores a squashed esteem h, the esteem h will lie in the scope of 0 to 1. To see how sigmoid capacity squashes the qualities inside the range, how about we imagine the chart of the sigmoid capacity.



VI. CONCLUSION

In this review paper, we have compared various Machine Learning models for fraud detection in the banking transactions. According to our research Autoencoder model gives suitable results, even though other methods can also be employed for fraud detection.

REFERENCES

[1] Credit card fraud detection using Machine Learning Techniques, John O. Awoyemi, Adebayo O. Adetunmbi , Samuel A.Oluwadare, Akure, Nigeria

[2] Alejandro Correa Bahnsen, Djamila Aouada, Aleksandar Stojanovic, Bju″rn Ottersten, Feature engineering strategies for credit card fraud detection, 0957-4174/ 2016 Elsevier.

[3] Ghosh, S., Reilly, D.L.: Credit card fraud detection with neural-network. In Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, 1994, vol. 3,IEEE(1994)

[4] Abhinav Srivastava, Amlan Kundu, Shamik Sural, Senior Member, IEEE, and Arun K. Majumdar, Senior Member, IEE, Credit Card Fraud Detection Using Hidden Markov Model , IEEE transactions on dependable and secure computing, vol. 5, no. 1, January-March 2008.

[5] Y. Sahin and E. Duman, Detecting Credit Card Fraud by Decision Trees and Support Vector Machines, IMECS vol 1, 2011.

[6] Tanmay Kumar Behera, Suvasini Panigrahi, Credit Card Fraud Detection: A Hybrid Approach Using Fuzzy Clustering and Neural Network, IEEE Computer Society, 2015

[7] Kang Fu, Dawei Cheng, Yi Tu, and Liqing Zhang, Credit Card Fraud Detection Using Convolutional Neural Networks, Springer International Publishing AG 2016.

[8] Krishna Modi, Bhavesh Oza, Outlier Analysis Approaches in Data Mining, IJIRT vol 3 issue 7.

[9] Sam Maes, Karl Tuyls, Bram Vanschoenwinkel, Bernard Manderick, Credit card fraud detection using bayesian and neural network, International Naiso Congress on Neuro Fuzzy Technology, 2002.

[10] "Autoencoder With Invertible Functions for Dimension Reduction and Image Reconstruction", Yimin Yang, Q. M. Jonathan Wu, Yaonan Wang, IEEE 2018

[11] Michael Nielsen(2017,March 15), Deep learning available,https://neuralnetworksanddeeplearning.com/chap6.html

[12] Abhinav Shrivastava,Amlan Kundu,Shamik Sural,Senior member IEEE and Arun k Majumdaar, senior member IEEE"Credit card Fraud detection using Hidden Markov Model",IEEE transactions on depend- able and secure computing,vol 5, no 1, January-March-2008

[13] "Improved Fuzzy Multicategory Support Vector Machines Classifier", Xi-zhao Wang, Shu-xia Lu, 2006 International Conference on Machine Learning and Cybernetics