Scam Detection for Online Shopping using Deep Learning

K. Anupriya
PG Scholar
University College of Engineering and Technology
(BIT Campus)
Thiruchirappali

Mrs. C. Kanimozhi
Assistant Professor
University College of Engineering and Technology
(BIT Campus)
Thiruchirappali

***Abstract*—Internet transactions have recently raised big concerns. The problem with making business through the internet lies in the fact that neither the card nor the cardholder needs to be present during sales. Credit card fraud event cause loses is a fact of life for all merchants, who accept credit card payment as a part of their business operations. It is impossible for the retailers to check whether the customer is the genuine cardholder or not. Fraud is an adaptive crime, so it needs special methods of intelligent data analysis to detect and prevent it. Fraud detection is one of the earliest industrial applications of data mining and machine learning. Typical machine learning tasks are concept learning or predictive modeling, finding predictive patterns. This paper shows how advanced Machine Learning algorithm and neural network algorithm combined successfully to obtain a high fraud coverage and also with a low false alarm rate. Artificial neural network when trained properly can work as a human brain, depend for their working on the neuron, which is the small functional unit in brain as well as ANN. Deep learning, advanced machine learning algorithm is used for fraudulent detection based on the user's behavior from their transactional record.**

***Keywords—Fraud detection; Credit card fraud; Artificial Neural Network; Machine Learning; Deep learning*.**

I. INTRODUCTION

In recent years, the development of new technologies has also provided yet further ways in which criminals may commit fraud. Traditional forms of fraudulent behavior such as money laundering have become easier to perpetrate. Fraud Detection involves identifying fraud as quickly as possible once it has been perpetrated. Fraud detection comes into play once fraud prevention has failed. In practice, of course fraud detection must be used continuously, as one will typically be unaware that fraud prevention has failed. To prevent credit card fraud by guarding our cards assiduously, but if nevertheless the card's details are stolen, then it can be able to detect, as soon as possible, that fraud is being perpetrated.

Fraud detection is a continuously evolving discipline. Whenever it becomes known that one detection method is in place, criminals will adapt their strategies and try others. Of course, new criminals are also constantly entering the field. Many of them will not be aware of the fraud detection methods which have been successful in the past and will adopt strategies which lead to identifiable frauds. This means that the earlier detection tools need to be applied as well as the latest developments. Techniques used for fraud detection fall into two primary classes such as statistical techniques and Artificial Intelligence. Early data analysis techniques were oriented toward extracting quantitative and statistical data characteristics. useful data interpretations and can help to get better insights into the processes behind the data.

The machine learning and Artificial Intelligence solutions may be classified into Supervised and Unsupervised Learning. These methods seek for accounts, customers, suppliers, etc. that behave unusually in order to output suspicion scores, rules or visual anomalies, depending on the method. Whether supervised or unsupervised methods are used, note that the output gives us only an indication of fraud likelihood. Fraud transaction can be happen on internet, the user will need some important information about a credit card such as a credit card number, validity, name of card holder, CVV number. Fraud detection can be detected on analyzing of previous transactions data which helps to form spending profile of the card holder. Every card holder having unique pattern contains information about amount of transactions, details of purchased items, merchant information, date of transaction, etc. It will be the most effective method to counter fraud transaction through internet.

The backpropagation algorithm for learning multiple layers of non-linear features but it did not seen to be able to make good use of multiple hidden layers and not work well in recurrent networks. Due to this reason, people moving towards DeepLearning. Deep Learning algorithm is a set of algorithms that attempts to model High-level abstractions in data by using architectures composed of multiple no-linear transformations and in order to model complex relationships among data .

II. LITERATURE SURVEY

Overview of big data initiatives, technologies and research in industries and discuss challenges and potential solutions in term of storage management and analytics. Data management, high performance analytics, high performance data visualization and flexible deployment options[6]. Need for predictions about future or unknown outcomes in the statistical or machine learning methods by predictive analytics. Predictive analytics needs when data generated by sensors, surveillance, transactions to run faster, more accurately and using larger heterogeneous information sources of varying data

quality and complexity [3]. Predictive analytics has investigated the use of social media data, but not limited to generated content and user behavior on social media sites[2].

By using Machine Learning approaches to keep an up- to- date on web indexed pages by representing each document as a bag of words[12], each web pages is denoted by a limited number of content and related features

A literature survey and system tutorial for big data analytics platforms and also prevalent Hadoop Framework for addressing big data challenges [1]. Big data research remains on typical big data applications can generate profit for business, improve the efficiency of government sectors and promote the development of human science and technology is also required to accelerate by big data progress.

Neural networks have become ubiquitous in applications including Computer Vision, Speech Recognition and Natural Language Processing focus on the CNN which has beat traditional algorithms in computer vision tasks. This leads to pruning and Quantizing Neural Networks. Traditionally the use of ANN for fraud detection is done using the generated network as a classifier. With this approach, the network is trained with examples of fraudulent and non-fraudulent actions.

Once trained, the ANN is able to classify new data as fraudulent or non-fraudulent activities[5]. However, some difficulties can appear. First a considerable imbalance between the non-fraudulent and the fraudulent samples should exist. Otherwise, this should be compensate be some preprocessing [11].

To overcome the problems by implementing Deep Neural Network , two hidden layers would, in principle[4], be able to preserve information into the pooling, high level features as second layer and the low level features as a first layer. After preserving, Fine-tune all the parameters of this deep architecture with respect to a supervised training criterion. Since the Supervised gradient is only non-null for the weights and the hidden layer biases of each layer.

The performance of neural network method is still inferior to supervised methods. To overcome this problem by using a Maximum Likelihood (ML) classifier[9] improves classification accuracy and also to make more essential use of ML methods.

An ANN is composed of neurons. These neurons are grouped in layers, where the last one is called the output layer, and the previous ones are called hidden layer. The connection of the neurons can be done in different ways, originating different kinds of ANN. Feed forward networks, which is the most widely used for time-series prediction.

A very simple Deep Learning network, which comprises only the very basic data processing components: Cascaded Principal Component Analysis (PCA), binary hashing and block-wise Histograms for indexing and pooling[10]. A framework recently introduced for achieving robust information representation is the DeepSpatio Temporal Inference Network (DeSTIN) model. This populates the entire hierarchy, and each of these nodes operates independently and in parallel to all other nodes[7].

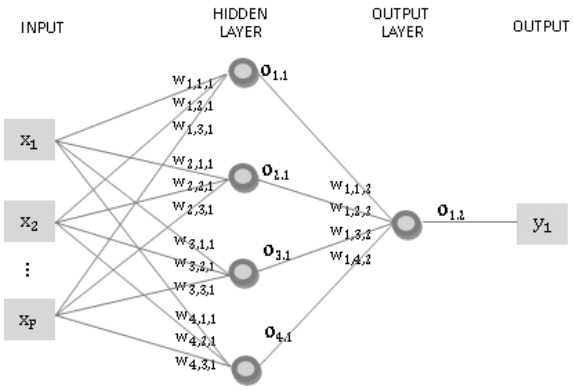

Figure.2. Artificial neural network.

III. CREDIT CARD FRAUD DETECTION METHODOLOGIES

A. Learning Algorithm (Feed Forward Back Propagation)

The back propagation learning rule is a standard learning technique. It performs a gradient descent in the error/ weights space. To improve the efficiency, a momentum term is introduced, which moves the correction of the weights in the direction compliant with the last weight correction. It is a multi- layer feed forward network that is trained by supervised learning. A standard back propagation network consists of 3 layers, an input, an output and a hidden layer. The processing elements of both input and output layer are fully connected with the processing elements of the hidden layer, as shown in figure. The fact that it is feed forward means that there are no recurrent loops in the network. The output of a node never returns at the same node, because cycles are not allowed in the network. In standard back propagation this can never happen because the input for each processing elemental ways comes from the previous layer (except the input layer, of course). This, again, is a large simplification compared with the real brain because the brain itself appears to contain many recurrent loops.

Supervised learning means that the network is repeatedly presented with input/output pairs (I,O) provided by a supervisor, where O is the output the network should produce when presented with input I. These input/output pairs specify the activation patterns of the input and output layer. The network has to find an internal representation that results in the wanted input/output behavior. To achieve this, backpropagation uses a two-phase propagate-adapt cycle.

i. First Phase: In the first phase the input is presented to the network and the activation of each of the nodes (processing elements) of the input layer is propagated to the hidden layer, where each node sums its input and propagates its calculated output to the next layer. The nodes in the output layer calculate their activations in the same way as the nodes in the hidden layer.

ii. Second Phase: In the second phase, the output of the network is compared with the desired output given by the supervisor and for each output node the error is calculated. Then the error signals are transmitted to the hidden layer where for each node its contribution to the total error is calculated. Based on the error signals received, connection weights are then a d a p t ed by each node to cause the network to converge toward a state that allows all the training patterns (input/output pairs) to be encoded.

b. NEURAL NETWORK

Neural network is the need as a set of interconnected nodes designed to represent functioning of the human nodes designed to represent functioning of the human nodes designed to represent functioning of the human brain. Each node has a weighted connection to several other linked nodes in adjacent layers. Single node take input received from linked nodes and use the weights of the connected nodes together with easy function for computation of output values. Neural networks can be created for supervised and/or unsupervised learning. The user specifies the number of hidden layers along with the number of nodes within a specific hidden layer. Artificial Neural Network classifier for fraud detection finds two important difficulties: the low number of samples and the increasing tendancy of the time series. The use of ANN prediction algorithms applied to fraud detection in time series data. Our ANN makes the prediction and the real results are compared with the prediction. If the results presents a great difference with the prediction it will mean that this sample should be deeply investigated.

c. Deep Learning Algorithm

Deep learning approach has found effective usage in pattern- recognition tasks such as stock market prediction and fraud detection. The deep learning use artificial neural network algorithms, these systems effectively gather data insights and recognize analysis and more. Now these systems are turning out to be effective in recognizing the patterns and characteristics of cybercrime and online fraud as well.

d. Artificial Neural Network

An Artificial Neural Network (ANN) , a computational model that is loosely based on the neuron cell structure of the biological nervous system. In most cases, an ANN is an adaptive system that changes its structure based on external or internal information flowing through the network during the learning phase. ANN has a potential for intelligent systems because they can learn and adapt, they can approximate nonlinear functions, and they naturally model multivariate systems.

IV. ARCHITECTURE DIAGRAM

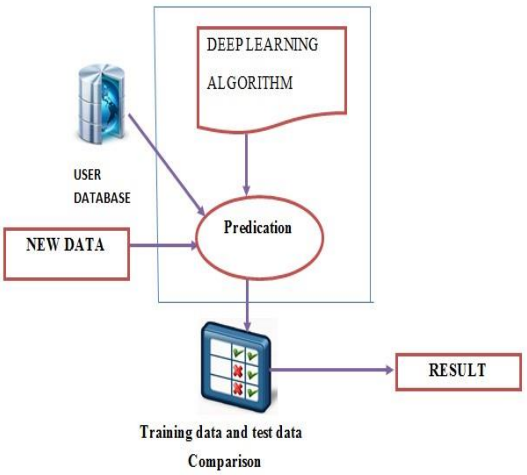

Fig. 4.1. System Architecture

Fig 4.1 shows Architecture of system, which explains that data from the user database which contains historical data of the user, passed into the framework in which prediction is performed with the help of Deep Learning algorithm. Prediction checks if there is any new data enters into the database then it compared with the user database. If the data is relevant to the user's historical database then it allows otherwise data is considered to be as false.

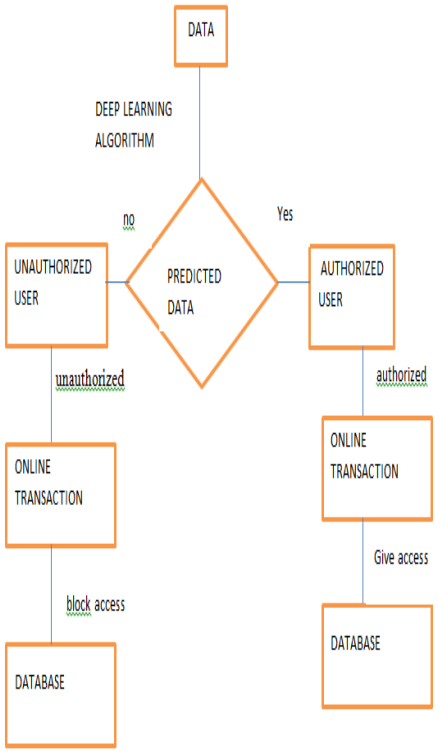

Fig. 4.2. Workflow of Credit card fraud detection.

In this figure Fig 4.2 shows the workflow of credit card fraud detection, transaction data which is not similar to that corresponding user's historical database then it is considered as unauthorized transaction.

V. DEEP LEARNING ALONG WITH NEURAL NETWORK

Classification of data by giving training to hierarchial network on a large set of observation and later data extracting is from this network to a relatively simple classification engine by Deep learning. Deep Learning can pose the tasks like vision, Speech Recognition as mapping concrete inputs such as image pixels to abstracts outputs like the identification.

The popularity of Deep Learning today are the extremely increased chip processing abilities eg., General-Purpose Graphical Processing Units (GPGPUs). The significant increased size of data used for training and the recent advances in Machine Learning and Signal or Information Processing researches.Deep Learning algorithm ability to expose classification in a diverse range.

With the help of Unsupervised Classification, once the data are classified the analyst attempt a posteriori to assign their natural class to the information classes of interest. By this way, users allow computer to select the class mean and covariance matrices to be used in the classification

Deep Learning is to address technical challenges with new way of thinking and transformative solutions. Deep learning has been very effective in integration of data from different sources. Deep Learning algorithm focused on learning from massive amounts of data and refer the algorithms to stacked Restricted Boltzmann Machine's by including Deep Belief Network's for massively parallelizing unsupervised learning[13].

The summary of main stream deep machine learning approaches provides a brief comparison in a variety of application domains, which is shown in below Table 1. Furthermore, Deep Learning platforms can also benefit from engineered features while learning more complex representations which engineered system typically lack.

Deep Learning, one of the most far flung borders of ML research utilizing neural net architecture along Unsupervised model development. It makes sense to use multiple models than a single one on its own even though that single model is superior

TABLE I. SUMMARIZES THE CURRENT PROGRESS IN LARGE- SCALE DEEP LEARNING

Methods	Computing Power	Number of examples and free parameters	Average running time
DBN [41]	NVIDIA GTX 280 GPU with 1 GB mwmory	One million images and 100 minllion parameters	~ 1 day
CNN [55]	Two GTX 580 GPUs, each with 3GB memory	1.2 million high resolution (256 x 256) images and 60 million parameters	~ 5-6 days
DisBelief [56]	1,000 CPUs with Downpour SGD with Adagrad	1.1 billion audio examples and 42 million model parameters	~16 hours
Sparse autoencoder [50]	1,000 CPUs with 16,000 cores	10 million 200 x 200 pixel images and one billion parameters	~3 days
COTS HPC [58]	64 NVIDIA GTX 680 GPUs, each with 4GB memory	10 million 200 x 200 images and 11 billion parameters	~3 days

To provide a solution for fraud detection is to deploy an ensemble of models, then it aggregate the majority votes of them for a single binary decision output(fraud/no-fraud). Even state-of-the-art random-forest or convolutional deep neural network on their own could still produce decision that performed under one of an ensemble models. It makes sense to use multiple models than a single one on its own even though that single model is superior (on one to one performance comparison to those that comprised of an ensemble).

When credit card is being used by unauthorized user the neural network based fraud detection system check for the pattern used by the fraudster and matches the pattern of card holder on which the neural network has been trained , if the pattern matches the neural network declare the transaction proceed[8] .

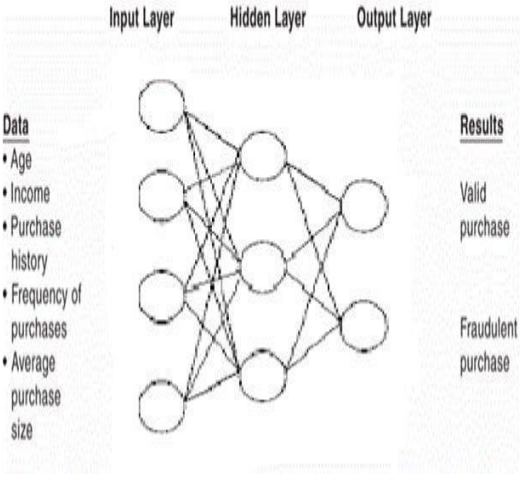

Figure 5. Feed Forward Neural Network

Figure 5, it is a network of computers all working to learn behaviors and patterns so that it will recognize those patterns and deduce a conclusion by utilizing backpropagation algorithm which needs to estimate activities by using sigmoid function.

$$y_j = \frac{1}{1+e^{-x_i}}$$

Once the activities of all output units have been determined, the network computes the error E, which is defined by the expression:

$$E = \frac{1}{2}\sum_i \left(y_i - d_i\right)^2$$

Where, yj is the activity level of the jth unit in the top layer and dj is the desired output of the jth unit.

The back-propagation algorithm used to compute how fast the error changes as the activity of an output unit is changed. Below given formulas, to determine error changes and error qualities,

$$\mathrm{EI}_j = -\mathrm{EA}_j = \frac{\partial E}{\partial y_j} = y_j - d_j \quad \mathrm{A}_j y_j\left(1-y_j\right)$$

$$\mathrm{E}W_{ij} = \frac{\partial E}{\partial W_{ij}} = \frac{\partial E}{\partial x_j} \times \frac{\partial x_j}{\partial W_{ij}} = \mathrm{EI}_j y_i$$

Deep learning is use to keep your clear of fraud. One of the problems with fraudsters is that they will try different things to get to you private data. Thus, Deep learning algorithm has created a network of computers patterned after neural networks in our own brains. Just like our own neural networks, DL network is constantly looking for patterns or correlations to user's behaviors. By using DL, if transaction seems there is something unusual, flag it as a concern and then either have one of their own experts or the account owner themselves take a look and see if this is approved behavior or not.

VI. CONCLUSION

Fighting cyber-criminals is not an easy task, and staying ahead of them is becoming increasingly more difficult. Using traditional methodologies is no longer the ideal strategy. In this paper by using different technique that is being used to execute credit card fraud how credit card fraud impact on the financial institution as well as merchant and customer. Neural network is a latest technique that is being used in different areas due to its powerful capabilities of learning and predicting. By using this capability of neural network in the area of credit card fraud detection as we know that Back propagation algorithm is the most popular learning algorithm to train the neural network and then in order to choose those parameter such as weight, network type, number of layer, number of node, etc. The "Deep Learning" approach to fight against criminals who attempt to exploit the online payment platform. This plays an important role to perform as accurately as possible, using combination of neural network to detect the credit card fraud successfully.

VII. FUTURE WORK

Fraud detection system efforts towards utilizing Deep Learning systems can be seen actively in its current anti-fraud systems. The Deep Learning systems have been effective in analyzing factors such as timelines, location, etc. as part of payment transactions. By utilizing this Deep Learning methodology to determine which fraud-detection models will be implemented. In future, to take this Deep Learning approach forward and one day also generate data driven insights in real time to curb fraud.

REFERENCES

[1] M.Chau and H. Chen, " A Machine Learning approach Webpages filtering using content and structured Analysis", Decisio Support Systems, Vol. 44, No. 2, pp. 482-494, 2008.

[2] Daniel Zeng, Brian Kirkm Jeeny Stout , " Crystal balls, Statistics, Big data and Psychohistory: Predictive Analytics and Beyond", IEEE Intelligent System, Vol. 30, No. 3, pp.114-122, 2015.

[3] Donald E. Brown, Ahmed Abbasi, Rayond Y.K. Lau, " Predictive Analytics: Predictive modelling at the micro level", IEEE Intelligent System, Vol. 30, No. 3 , pp. 1541-1672,2015.

[4] Gao Huang, Shiji Song,Jatinder N.D. Gupta and Cheng Wu, " SemiSupervised and Unsupervised Extreme Learning Machines", IEEE Transition on Cybernetics, Vol. 44, No. 12, pp. 2405-2417, 2014.

[5] Hau Hu, Yonggang Wen, Tat-Seng Chau and Xuelong Li, "Toward Scalable Systems for Big data Analytics: A technology Tutorial", Vol. 2, No. 3, pp. 1556-1603, 2014.

[6] Hua Fang, Zhaoyang Zhang, Chanpaul Jin Wang, Daneshmand, Chonggang Wang, Honggang Wang, " A survey of Big data research", IEEE Networking, Vol. 29, No. 5, pp.6-9,2015.

[7] Itamar Arel, Dereck C. Rose and Thomas P. Karnowski, " Deep Machine Learning : A new frontier in Artificial Intelligence Research", IEEE Transactions on Cybernetics, Vol. 5, No. 2, pp. 1556-1562, 2010.

[8] Raghavendra Patidar, Lokesh Sharma, " Credit Card Fraud Detection Using Neural Network ," International Journal of soft computing and Engineering (IJSCE), Vol.1, pp. 32-38, June 2011.

[9] Ruslan Salakhutdinov, Joshua B. Tenenbaum and Antonia Torraiba, " Learning with Hierarchial Deep models", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 37, No. 3, pp. 01162- 0169, 2013.

[10] Tsung-Han Chan, Kai Jia, Shenghua Gau, Jiwen Lu, Zinan Zeng and Yi Ma, " PCANET : A simple Deep Learning Baseline for Image Classification?", Vol. 11, No. 3, pp. 145-151, 2015.

[11] Yoshihisa Hara, Robert G. Atkins, Robert T. Shin, Jin Au Kong, Simon H. Yuch and Ronald Kwok, "Applcation of neural networks for sea ice classification in polarimetric SAR images", IEEE Geoscience and Remote Sensing, Vol. 33, No. 3, pp. 0196-2892, 1995.

[12] Yu Deng, Dahl G.E, Acero, "Context-dependent trained deep neural networks for large vocabulary speech recognition", IEEE Transactions on Audio, Speech and Language Processing, Vol. 20, No. 2, pp. 33-42, 2012.

[13] Xue-Wen-Chen and Xiaotong Lin, " Big Data Deep Learning : Challenges and Perspectives," IEEE Access, Vol.2, No. 2, pp. 2169- 3536, 2014.