# Analyzing Techniques for Fraud Detection in Financial Transactions

**Rohan Bhimnrao Rangari, Manoj B. Chandak**

***Abstract*: Transactional data handling in real time is the most security concerning task for companies these days due to the recent advancements in attacks occurring in financial data. One of these attacks modifies or adds fake transactions to the input data thereby reducing their genuineness and credibility of the system in place for the financial corporations. In order to improve the security of the financial transactions and detect fake transactions, researchers have developed many techniques, some of which are reviewed in this paper. This paper also compares these state-of-the-art techniques and concludes which techniques are better for handling financial data. The paper concludes with some interesting observations about these algorithms and suggests improvements in them so that a better and more secure system can be built upon..**

***Keywords: Transactions, fraud, attacks, financial***

## I. INTRODUCTION

Finance deceit can be portrayed as "The unapproved utilization of a person's individual data to produce buys, or to dismiss assets from the client's document." according to the review of statista [1] 41% of world wide web consumers have purchased things online in 2013. In 2011, the amount of computer-aided buyers overall augmented 792.6 million. After a year, the number rocketed to 903.6 million. In 2013, 41.3% of worldwide web clients had obtained items on the web. In 2017, this total is required to amplify to 46.4%. In the investigation by BBC discovery, Accidents from web-based trading trick soared to near 48% in 2014 contrasted and 2013 as users progressively led their commercial operations on the network. So with an enlarged amount of such cashless transaction and internet purchasing, dishonest transactions are additionally extending.

Pretenders can be brought about by taking or bartering off funding nuances by email phishing, telephonic phishing, malware, non-secure protection nuances, simple connection targets, and shoulder surfing. Dishonest trades can be recognized either order approach or by distinguishing remote exchange from conventional replacements. For an organization strategy, the start prototype is developed from preparing knowledge. Highlights are liberated and changed from raw knowledge while offering it to prepare show [2].

During everyday utilization of Visa swaps, the procurement of details and officials supports online trades or card swiping acquirements.

This incites gain in online trades appropriating credit and monitor cards evolving to nature of spontaneous worth. Pretenders correlated with the charge card area have made drastic infliction the consumers and the specialist system and is supposed to be distant extra monstrous in future times. Deceits patrol and regulate to the expeditious developments in the discovery and discover bright methods to incorporate in unauthorized operations [6].

Pretenders induced because of certain calculating coders are hazardous and unreliable. A trained cheat can make a few dispositions and direct charge card replacements without doing touched. Discussing web-based enterprise transfers the grave point examined because of those deceitful operations is so like authentic things. Consequently becoming a prolific and twisted scam classification structure is an undeniable urgency to keep these crooked activities. The examination division of this problem is to recognize scams in an enormous dataset where the established trades are more and the false transfers are complete lowest or next to petty. There are not very many articles on Mastercard cheat locating tacts because of a particular behavior that certain procedures can't be investigated without a dataset. Thus it's laborious to illustrate the heartiness or even the reasonableness of accomplishment symmetry of the procedures. As we recognize that the charge card data is secret, the bank possessors and authority co-ops don't commission in yielding this data for interpretations also.

In the present quick moving world the requirement for information mining turns out to be the increasingly basic immense measure of information is put away in an information distribution center and it is basic to separate the destitute data from the stockroom. Distinctive information mining systems are there to get the data required for the specific issue. A scam can be portrayed as being unfair or felonious misleading aimed to produce in wealth-related or specific bonus [3], or to harm another person without essentially prompting direct lawful outcomes. Fraud aversion and fraud identification are two different ways to maintain a strategic distance from misfortunes and frauds. Fraud anticipation is done before the event of fraud. Fraud recognition is done when fraudsters went through fraud aversion systems and begin a fraudulent exchange. None of them is certain whether the exchange went through the counteractive action framework. As needs are, the fundamental point of the location methods is to discover whether an exchange is fraudulent or not at the earliest opportunity [4]. The diverse sorts of frauds that can happen are fraudulent exchanges in Visa

**Revised Manuscript Received on December 22, 2018**.

**Rohan Bhimrao Rangari**, Department of Computer Science, Shri Ramdeobaba college of Engineering, India.

**Dr. Manoj B. Chandak** , Department of Computer Science, Shri Ramdeobaba college of Engineering, India.

frameworks and internet business frameworks, fraud account in monetary frameworks, fraudulent calls or administration uses in media transmission frameworks and fraudulent cases in protection frameworks. Charge card frauds are of two kinds as on the web and disconnected. Serious burglary, purpose scam, bogus cards go under detached and cardholder isn't should have been open in online trick as the transfer is executed remotely and card's innuendoes are needed. A hand-operated check, a password or a credit card imprint is not obligated at a specific term concerning obtainment. Presently a day, online frauds are expanded a great deal since utilization of charge card is unavoidable. In European nations, above half of the fraud misfortunes in 2008 is a result of online frauds which are accounted for by visa.

The next section compares the techniques used for transactional fraud detection, mainly targeted to credit card fraud detection, later we compare these techniques to find the ones which are better than others, and finally, we conclude with some interesting observations about these techniques and ways to improve their security

## II. Literature review

In [5] the experts suggested a Visa scam classification illustrate that detects fraud from especially imbalanced and anonymous charge card exchange datasets. Visit property set digging is employed for exposing licit and outlawed cases of transfers which handles the exemption issue in class. To discover whether the approaching exchanges of the clients have a place with the lawful or unlawful example, a coordinating calculation is proposed and concurring that exchange nearer to the examples are recognized and choices are made. No uncommon consideration on credits is given to deal with the unknown idea of exchange information and each trait is dealt with similarly for example finding. On UCSD Data Ming Contest 2009 Dataset, Evaluation of accomplishment for this prototype is fulfilled and noticed to have scarce false warning rate differentiated with whacking advantage classifiers, a degree of trickery identification is high, order rate is adjusted, Matthews connection coefficient [6]. Characteristic calculations utilized by the creators are KNN, Random Forest, SVM, and Naïve Bayes. The key system proposed is "Fraud Miner". Visit thing set mining is utilized to make designs for legitimate and fraud exchanges for every client from official and quick transfers of them definitely in the development stage. The approaching exchanges are checked by the coordinating calculation to identify which design it has a place. The lawful example of the client is coordinated with approaching exchange then "0" is rendered by the computation. That swindle problem of the customer is organized with resembling trade then "1" is returned by the calculation [7].

Any vote tree [8] comprises a graphical replica of believable results concerning a determination conditional at precise events. A choice tree begins with root hub, isolates into discrete branches, these branches are associated with different hubs, etc. Voice tree completes up toward a core termed leaflet core. All core in Decision tree addresses via an examination, members related with it signifies to its trustworthy outcomes and a leaflet core owns an emblem of class. With this imperative methodology of isolating and choosing, choice tree more often than not seclude the unpredictable issue into basic ones. A straightforward case of choice recognizes plausibility of exchange being genuine or fraudulent. Affiliation Rules are produced to distinguish fraudulent exchanges and ordinary exchanges. In fraud recognition, created guidelines will be utilized to characterize fraudulent and real exchanges. In this manner, rules are produced according to conduct. This strategy is like a choice tree.

In [9] the examiners suggested a task that employs the original dataset for the achievement correlation of choice tree calculations and SVM. Relatively, choice tree models are superior to SVM models on the test dataset. At the point when the preparation datasets are utilized correlation results in the turn around the structure with the end goal that preparation information is over-adapted by SVM models. The achievement factor of this issue is the task of numerous fraudulent exchanges as fraudulent. Notwithstanding whether the exchange is genuine fraud or genuine typical task, the rate of genuine assignments is appeared by precision. As per accuracy, when exhibitions of models are thought about as the expansion in the number of preparing information, overfitting turns out to be less and SVM models execution end up similar to choice tree demonstrate. SVM models got just less number of frauds than choice tree models. An exactness of assignments by models isn't identified with the quantity of genuine fraudulent exchanges relegated as fraudulent. In this issue, the execution metric isn't coordinated with exactness in like manner. C5.0 display is the best contrasted with different models however C&RT show gets more frauds from tests. C&RT and C5.0 are picked by the above key factor [10].

In [11] it is recommended work on charge card fraud location dependent on neural systems. In spite of the fact that distinctive information mining advances are there to identify fraud, every one of them is not available to detect crooked transfer in improvement. Two exceptional attributes of Visa fraud identification are the constrained time to take choice whether to acknowledge or reject and the vast measure of Visa exchanges handled in stipulated time. The neural system based fraud identification is like human cerebrum working. The neural system made a PC to think as human cerebrum that learns through past involvement. The learning background or information is utilized to tackle and settle on choice in issues in a day to day life. A similar strategy is for charge card fraud recognition. The customer utilizes settled example of charge card use. This example is taken for the past a couple of years to prepare a neural system. The diverse different classifications of data can moreover be put away like area for children buy, incidences of tremendous buy, etc in constrained time. The neural system prepares the different essences of Mastercard fraud alongside Mastercard utilization design which is given by the bank. Visa use design is taken by the forecast calculation to separate fraudulent and non-fraudulent. Unapproved [12] client's example is coordinated with the unique card holder's example which is prepared by a neural system, and if a design is the same the choice made as the authentic exchange. Example coordinating isn't really to be careful rather little varieties can be acknowledged and on the off chance that there exists enormous contrast in an example, at that point risks that specific exchange is illicit exchange is more. The yield of a neural system will be in the middle of 0 and 1. On the off chance that the yield is beneath .6 or .7, it suggests exchange legitimate and whenever yield is over .7 then the likelihood of an illicit exchange is high. In a few events, legitimate clients may make the exchange that will be very unique and at times fraudster makes exchanges that coordinate the example prepared by a neural system. Because of impediment issues, lawful clients will utilize card for constrained sum yet fraudster will attempt to do huge buy before the move made by the charge card holder which will be a befuddle with the prepared example by neural system. The procedure of business will be available dependably in neural system design acknowledgment frameworks structure. History descriptors give subtleties utilization subtleties of card and installments made. Different descriptors have data about date if an issue, etc [13]. Neural Network is one of the focal classifiers to grasp achieved experiment in the nucleus of conspicuous highlights. NN achieves the equivalent being a human's cerebrum. NN involves numerous layers inside which this fundamental layer is the input layer moreover the outermost layer is yield layer. It might have an amount of a shrouded layer or no secreted layer. On the off chance that Neural arrangement constitutes further than one shrouded layer, at that point, it is deep-felt learning. Each layer has assorted neurons, and every neuron is associated with weighted edges. A yield of all neuron is a part of its system. This function is called actuation work. Each illustration of various performance capabilities utilized is sigmoid capacity, step work, limit work, direct capacity and so forth. Numerous employed potential is Sigmoid capacity amongst all. Yield layer possesses the corresponding amount concerning neurons as composition result. All neuron of yield layer supplies the probability of obtaining such class. Neurons of the second layer connected with yield layers' neurons. The neural system completes on a selection of weights on edges from learning proffered to it for providing and change loads appropriating back spread computation [14]. In [15] they have proposed a three-layer feedforward neural system to recognize Mastercard tricks. Features from 50 assets were combined into 20 constitutes an augmentation to the neural system. They served the neural system on a wide informational index of named exchanges taken from Mallon bank. These exchanges involve the case of various fraud cases. Complexities toward appropriating neural system are to determine the number of layers and number of neurons in each layer, number of weight and learning flow. [16] Learning rate is a size of the succession taken at every sequence, before the revision of shots. Each big motivation toward knowledge flow can create the model train expeditious, yet it can overshoot nearby minima. Selecting Activation to achieve concurring dataset is moreover experimenting trial. In 1993, VISA Company had added Neural Network innovation to encounter card fraud. Each neural system can moreover comprise associated with the ancestral computation to determine parameters regarding Neural Network[17].

Convolutional Neural Network (CNN) [18] implies one part concerning intelligent knowledge. Mapping concerning augmentation to shrouded layer articulates to an individual component map. Individually element map articulates toward item brand. Each idea about arranging neurons toward highlight map is designated convolution. Subsampling diminishes parameters of highlight map. A completely associated layer is the same as the neural system [19]. To help the issue of the not adjusted information, they utilized cost-based examining a system to make a different number of made fakes to set up the model. They related CNN seem in the light of how it is fair for equipping a comprehensive size of information and CNN has the portion to escape over-fitting.

The Mastercard issuing bank runs a fraud recognition framework (FDS). FDS checks every single internal exchange. The card subtleties and buy subtleties are utilized by FDS to discover veritable or counterfeit exchange [20]. FDS checks for the distinction by contrasting spending subtleties of the Visa holder, conveyance address, etc. On the off chance that there is a distinction, FDS affirms that the exchange is phony and the exchange is declined. Perception images ought to be resolved to deal with Mastercard exchange by HMM. Confine the x estimations of procurement into M value ranges, for example, V1, V2, … , VM, builds up perception images to a bank. An HMM is prepared for each charge card holder. A grouping calculation is executed to get the perception images of individual cardholder‟s exchange separately. Numerous characteristics are put away in the database of the issuing bank. The spending subtleties of cardholder assume a noteworthy job which can be separated into three classifications in particular high-spending, low – spending and medium-spending. The proposed model is prepared with a couple of exchanges so it will be less demanding to recognize frauds and which is additionally created with redresses for future references to productively distinguish the fraud. Starting image grouping is shaped from the images taken from the cardholder‟s preparing information in the wake of learning HMM factors. A hidden Markov shows up (HMM) is a quantifiable Markov resemble inside which the core is registered is admitted to be a Markov chain with secured positions. An HMM is a duplex embedded likelihood distribution process with pecking order levels. Fraud location Approach utilizing HMM is proposed in [21]. People need to be held three utility extends low, medium and high {l,m, h} asset of conceivable perception. For instance, let l = (0,100$], m=($100,$500], h=($500,credit card limit]. On the off chance that a client makes an exchange of $320, at that point, the resultant perception image will be m. Every exchange sum more often than not relies upon the equal sort of procurement. The arrangement of every single imaginable sort of procurement and the arrangement of every conceivable line of business of traders shapes the arrangement of concealed conditions of the HMM. The proposed methodology in [22], Hidden Markov Model (HMM)- based charge card FDS does not require fraud marks and still it can identify frauds by considering a client's spending design.

Joseph Pun, Yuri Lawryshyn pursues the meta-learning procedures presented by Chan and Stolfo [23] in their proposed work. The meta-learning endeavors to join the aftereffects of different students to a precision of forecast and qualities and shortcoming of strategies are complemented with one another. The two different ways of joining calculations are mediator and combiner techniques. Through the trials, Chan and Stolfo came resolution that combiner technique is increasingly successful contrasted with referee strategy [24]. In the combiner technique, the characteristics and right groupings are utilized to prepare base classifiers. The subsequent

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

expectations are then sustained into the meta-level classifier. The blend of Original traits, expectations from base classifier and right arrangement for every single occasion is utilized to make another "consolidated" dataset which is then utilized as preparing information for meta-classifier. In the combiner procedure, the last expectation is the forecast from the meta-level classifier [25].

## III. THEORETICAL ANALYSIS

The following table shows the analysis of various techniques and their advantages and drawbacks w.r.t. the datasets used,

TABLE 1. Comparison of techniques

| Algorithm/ Methods | Advantages | Drawbacks |
|---|---|---|
| Logistic Regression | 1.It produces a simple probability formula for classification.<br>2.It works well with linear data for credit card fraud detection. | 1.It cannot be applied on non-linear data.<br>2. It is not capable of handling fraud detection at the time of transaction. |
| Decision Tree | 1.Used for both linear and non-linear input<br>2.Easy to interpret | 1. Algorithm is complex<br>2. Not real time<br>3. Useful for simple data<br>4. Pre-processed data needed |
| Artificial Neural Network | 1.Real time operation<br>2.Adaptive and used for complex input | 1. Initialization needed<br>2. Needs processing delay<br>3. Energy consumption is high<br>4. Difficult to interpret |
| Hidden Markov Model (HMM) | 1.Real time & highly accurate.<br>2.Scalable for handling large volume of data | 1. Needs time to train<br>2. Highly expensive |
| Support vector machine | 1.Real time operation | 1.Less accuracy |
| K-Nearest Neighbor Algorithm | 1.No need to train | 1. Limited Accuracy<br>2. Non real time |
| Rule based method | 1.Easy to understand and implement | 1.Not good for new classes |

## IV. CONCLUSION

In this paper, From the analysis of the algorithms we can observe that the neural network based algorithms, and more specifically the convolution neural network is the best choice for detection of fraud in transactional data. Other techniques like HMM, kNN and SVM are close to the neural network in terms of detection accuracy but they lack in terms of the datasets flexibility, and might not work for large amount of data.

In future, researchers can work on a blockchain based model combined with artificial intelligence in order to improve the security of financial transactions and reduce frauds in the overall banking system, including credit card based systems, and maintain higher level of transparency for the users.

## REFERENCES

[1] Smt.S.Rajani, Prof.M. Padmavathamma, “A Model for Rule Based Fraud Detection in telecommunications”, International Journal of Engineering Research & Technology (IJERT), Vol. 1 Issue 5, July – 2012.

[2] Minewiskan, Microsoft Neural Network Algorithm Technical Reference (2017, March 14) available at https://docs.microsoft.com/enus/sql/analysis-services/data-mining/microsoft-neural-networkalgorithm-technical-reference

[3] 3.Divya.Iyer,Arti Mohanpurkar,Sneha Janardhan,Dhanashree Rathod,Amruta Sardeshmukh” credit card fraud detection using hidden markov model ” 978-1-4673-0126-8/11/$26.00_c 2011 IEEE

[4] 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.

[5] Alejandro Correa Bahnsen, Djamila Aouada, Aleksandar Stojanovic, Björn Ottersten, “Feature engineering strategies for credit card fraud detection”, 0957-4174/ 2016 Elsevier.

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

[7] Ghosh, S., Reilly, D.L.: Credit card fraud detection with a neuralnetwork. In: Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, 1994, vol. 3, pp. 621–630. IEEE (1994)

[8] Ekrem Duman, M. Hamdi Ozcelik "Detecting credit card fraud by genetic algorithm and scatter search". Elsevier, Expert Systems with Applications, (2011). 38; (13057–13063).

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

[10] Statista the statistic portal (2017, March 14) available https://www.statista.com/topics/871/online-shopping/

[11] Priya Ravindra Shimpi, Prof. Vijayalaxmi Kadroli, "Survey on Credit Card Fraud Detection Techniques", International Journal Of Engineering And Computer Science, Volume 4 Issue 11 Nov 2015, Page No. 15010-15015.

[12] Hidden Markov model (2017, March 15) available https://en.wikipedia.org/wiki/Hidden_Markov_model

[13] A.J. Graaff A.P. Engelbrecht agraaff "The Artificial Immune System for Fraud Detection in the Telecommunications Environment" 20 November 2014.

[14] P.Jayant,Vaishali,D.Sharma," Survey on Credit Card Fraud Detection Techniques", International Journal of Engineering Research & Technology (IJERT), Vol. 3 Issue 3, March – 2014,pg.1545-1551

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

[16] K.RamaKalyani, D.UmaDevi" Fraud Detection of Credit Card Payment System by Genetic Algorithm" Volume 3, Issue 7, July-2012

[17] S. Benson Edwin Raj, A. Annie Portia "Analysis on Credit Card Fraud Detection Methods", IEEE-International Conference on Computer, Communication and Electrical Technology, (2011), pg.152-156.

[18] Venkata Ratnam Ganji," Credit card fraud detection using Anti-k Nearest Neighbor Algorithm",International Journal on Computer Science and Engineering (IJCSE) Vol. 4 ,06 June 2012,(1035-1039)

[19] Emin Aleskerov, Bernd fieisleben and Bharat Rao, "CARDWATCH: A Neural Network based database Mining System for Credit Card Fraud Detection"

[20] 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

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

[22] Renu, Suman" Analysis on Credit Card Fraud Detection Methods" volume 8 number 1– Feb 2014

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

[24] Raghavendra Patidar, Lokesh Sharma, "Credit card fraud detection using Neural Network", IJSCE Volume-1, Issue-NCAI2011, June 2011.

[25] Abhinav Srivastava, Amlan Kundu, Shamik Sural, and Arun K. Majumdar" Credit Card Fraud Detection Using Hidden Markov Model" VOL. 5, NO. 1, JANUARY-MARCH 2008

## AUTHORS PROFILE

**Vinay Yogendra Mishra** is currently Pursuing MTech in CS from Shri. Ramdeobaba College of Engineering and completed Bachelor of Engineering in Computer Engineering from Dr.D.Y.Patil College Of Engineering,Akurdi, (PUNE University) in 2017. His interest of research is Deep Learning, Computer Vision, Feature Detection, Recognition, Internet-of-Things(IoT), Neural Networks, Machine Learning, and Data Analytics.

**Dr. Manoj B. Chandak** holds a Ph.D. in Computer Science and Engineering. He is an acknowledged academician with over 22 years of teaching experience with more than 60 research publications in referred journals and conferences. A recognized supervisor in doctoral research in RTM Nagpur University and SGB Amravati University, his research interests are in NLP, Big Data, Information Retrieval and Mobile & Wireless Technology.

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*