IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Review of Credit Card Fraud Detection Using Data Mining

Reshma Farooqui*, Sifatullah Siddiqi#

*M.Tech, CSE, Integral University, Lucknow, Uttar Pradesh, India

#Assistant Professor, dept. of CSE, Integral University, Lucknow, Uttar Pradesh, India

Abstract— **Data mining (DM) involves a core algorithm that enables data deeper than basic insights and knowledge. In fact, data mining is more part of knowledge discovery process. Credit card (CC) providers provide multiple cards to their customers. All credit card users must be genuine and sincere. Giving a card to any kind of mistake can lead to a financial crisis. Due to the rapid growth in cashless transactions, it is unlikely, Fake transactions can also be increased. A fraudulent transaction can be identified by studying credit cards of various behaviors as a previous transaction history dataset. If there is any deviation from the available cost pattern, it is a bogus transaction. DM & machine learning techniques (MLT) are widely applied in credit card fraud detection (CCFD). In this survey paper we show an indication of various widely available DM & MLT for detecting credit card fraud.**

Keywords Data mining, Credit card, machine learning, financial crisis, transaction history.

1. INTRODUCTION

Information mining is interaction of seeing as genuinely solid, mysterious and noteworthy data. Moreover, DM issue should be well distinct, can't be made sense of with question and revealing apparatuses, and can be coordinated in DM process model. This information should be accessible, important, sufficient and clean [1]. The bank is monetary establishment that gets ventures from local area. Being defenseless against a misrepresentation turns into a significant exclusion for the bank. 'K Chan and J Stolpho et al' note that various types of extortion and monetary misrepresentation are ones generally impacted by bank. Attributable to quickly developing internet banking action, we came to know that 44% of US individuals utilized these web-based exchanges. 'John T The MistyLook Theme' expressed that It is assessed to have misfortune $ 8.2 billion of every 2006 with $ 3 billion in USalone. 'Philip K Keener' says that DM is recently creating apparatus that can distinguish CCF rapidly. Characterized by 'Chan and Wei Fan et al' as they would like to think, information mining can assist us with tracking down connections between stowed away examples and informational indexes. Misrepresentation or criminal extortion because of monetary or work force benefits.

Card giving bank should bear regulatory and foundation costs. Concentrates on say normal time in the midst of fake exchange dates and chargebacks can be as long as 72 days, giving fraudster sufficient opportunity to bargain serious damage [2]. Online Mastercards or disconnected exchanges for actual cards are utilized for day to day existence Visas for good and administrations. In actual exchanges, a Mastercard is embedded into an installment machine at the dealer's store to buy the products.

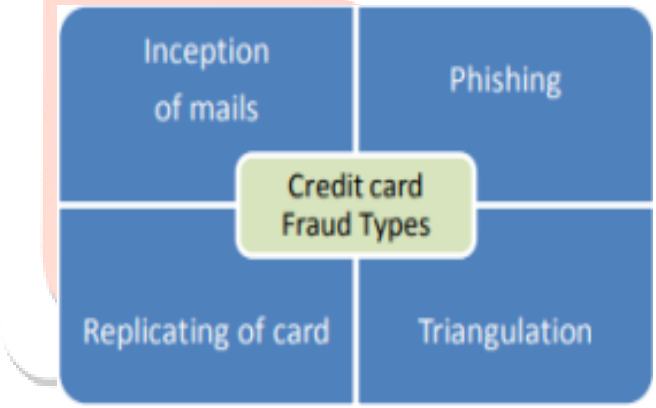

Figure.1 Types of Credit Card Fraud

This mode may not be able to track forged transactions because the attacker already theft a credit card. By online payment mode, attackers have very little data to counterfeit transactions (safe codes, card numbers, end dates, so on [3].

In this paper section I contains the introduction, section II contains the literature review details, section III contains the details about credit card fraud, section IV shows the advantage of credit card, V describe the type of fraud, Section VI explains the credit card fraud detection techniques and section VII provide conclusion of this review paper.

2. LITERATURE REVIEW

Anuruddha Thennakoon et al. [2019] Real-world exchanges in four significant extortion cases. Each trick is settled by ML model, and most ideal way is through valuation. This assessment gives exhaustive manual for choosing the ideal algo for the kinds of tricks and loads we view as most fitting relief measures. One more key part that we explanation in our venture is ongoing CCFD. That's what to do, we use prescient investigation to decide if specific exchange to AI models and API module is genuine or false. We are likewise assessing new methodology

that tends to contorted appropriation of information. Data applied in our examinations are as of private divulgence arrangement [8].

Chunzhi Wang et al.[2018] BP Neural Network, an optimizing framework that streamlines the BP brain organization, depends on tackling slow union rate issues, which can bring about neighborhood ideal, network blackouts, and unfortunate framework strength. Utilizing the Whale Group Optimization algo to improve weight of BP organization, we first methodology WOA algo to get essential worth, and Next BP network algo to exact shortcoming values. The ideal qualities are gotten [9].

Sahil Dhankhad et al. [2018] In various regulated ML algos, distinguish CC fake exchanges and execute genuine world datasets. Furthermore, we utilize these algos to execute super classifier by installed learning approaches [10].

Zahra Kazemi et al. [2017] To eliminate the best highlights from Mastercard exchanges, recommend a top to bottom auto encoder, and afterward add a delicate max organization to the class mark. Such information utilizing super-complete auto encoder can be utilized to guide to an enormous sum ofspace, and the meager model can be valuable for grouping targets [12].

Kosemani Temitayo Hafiz et al. [2016] center around building scorecards as of significant appraisal standards, viewpoints and capacities of prognostic examination seller arrangements as of now used to CCFD. Record gives synchronous examination of five merchant CC forecast examination seller arrangements in Canada. Affirming concentrate on results, rundown of CCFs has illustrated PAT seller's tests, dangers and limitations [13].

V. Mareeswari et al. [2016] Owing to limits of current framework, this paper recommended new algo with present algo. Limits of current adaptability issues, profoundly lopsided classes, and time requirements. Extortion discovery for local area and spike recognition utilizing CC application mixture support vector machine (HSVM). HSVM is regularly used procedure to design acknowledgment and arrangement [14].

Fahimeh Ghobadi et al. [2016] Progress CCFD Model Based on ANN and Meta Cost Process to Improve Risk and Loss. ANN methodology was utilized for charge card extortion anticipation and identification. Because of unequal nature of data (misrepresentation and non-extortion cases), misrepresentation can be challenging to distinguish. Added Meta CostProcess to manage issue of unsound data. Cost Sensitive NN (CSNN) depends on misuse identification strategy. In light of the examination of the Artificial Immune System (AIS), this model tracked down cost reserve funds and development rates. This concentrate on information was gotten from a huge Brazilian charge card backer who gave genuine exchange information [15].

Krishna Modi et al. [2017] past exchange information from clients dissecting cost conduct. On the off chance that there is any deviation from the accessible expense pattern,itis a bogustransaction. Banks and Mastercard organizations utilize various information mining techniques to recognize misrepresentation, for example choice points of view, rule-based mining, NN, fluffy grouping draws near, stowed away Markov models or crossover draws near. Both of that approaches are utilized to investigate normal use designs in view of the past activities of the clients. This paper gives examination of various strategies to recognizing extortion [11].

3. CREDIT CARD FRAUD

Unapproved system of CC or data denied of proprietor's information is called CCF. The unique CCFtrick applications and ways of behaving are connected with two gatherings of cheats. When application extortion happens, fraudsters apply for another card from the bank or give itto organizations that utilization misleading or other data. A client can record numerous applications with a solitary common of portrays (named copy misrepresentation), or an alternate client with comparative depicts (named personality extortion). All things being equal, there are essentially 4 primary sorts of social extortion: taken/lost cards, mail burglary, counterfeit cards, and 'current card holder doesn't exist' misrepresentation. When a taken/lost card extortion happens, fraudsters take a Mastercard or get lost card. Mail robbery misrepresentation when a fraudster gets individual data from a bank via the post office before a Mastercard or unique card holder. Counterfeit and Card Holders Fraud and Mastercard depicts are not introduced. In past, far off correspondences should be possible utilizing card subtleties by means of mail, telephone or web. Second, counterfeit cards are made on card information.

4. Advantages of using a credit card

- **Ease of purchase**

Credit cards make life easier. A payment made over the Internet, by telephone, & by an ATM allows customers to borrow credit at a time, place & amount without paying for an efficient payment method.

- **Keep customer credit history**

Having good credit history is often key to finding loyal customers. This history is valuable not simply to CCs, but also for other financial services, e.g. loans, rental application or certain jobs. Lender & issuer of credit mortgage companies, CC companies, retail stores & utility companies can evaluation credit scores, timely & responsible customers' history of how wellthey operate on their loans.

- **Protection of Purchases**

Credit cards can provide other protection to customers if they are lost, damaged or stolen. Buyer’s CC statement & corporation can ensure that original receipt has been lost or taken. Additionally, specific CC companies offer large purchases for insurance [4].

5. Types of Frauds

This section covers credit cards fraud, telecommunication fraud, computer penetration, bankruptcy fraud, theft / fake fraud, application fraud & conduct fraud. CCF: CCF is classified in 2 categories:

- Offline Fraud: At a call center or other location on a physical card stolen using offline fraud.
- Online Fraud: Online fraud is by a cardholder with shopping, Internet, phone, web or absence. Telecom fraud: Use of telecom services for other types of fraud.
- Its victims are consumers, businesses & communications service providers. Computer Intrusion: Intrusion is distinct a warranty or invasion without entering work; this means “unauthorized attempts to access data, & manipulate data.
- Infiltrators can be since any environment, outsider (or hacker), & person who recognizes layout ofsystem. Bankruptcy Fraud: This column attentions on bankruptcy fraud.
- Bankruptcy fraud resources not by CC. One of most complex scams is bankruptcy fraud. Theft Fraud/ Counterfeit Fraud: In this section, we attention on each

other's related theft & Counterfeit fraud. Theft fraud states card that is not yours.

- Once holder gives some feedback & approaches bank, bank will proceed action to investigate thief as soon as likely. Similarly, credit fraud is used remotely when fraud is committed, Wherever CC details are required only.
- Applications Fraud: Once a person relates to credit card, he or she is given false data, which is called application fraud.
- Toward detect application fraud, two dissimilar scenarios need to be considered. While apps with the same information from similar user, it is termed duplicate, & when applications derived as of different people by same information, it is called identity fraud.
- Phua et al. describes application fraud as "demonstration of identity crime, occurs when application forms contain possible, & synthetic (identity fraud), or real but also stolen identity information (identity theft)" [5].

6. CREDIT CARD FRAUD DETECTION TECHNIQUES

Genetic algorithm - Genetic algorithm is often recommended as fraud prediction methods. An algorithm developed by Bentley is based on genetic software design to create the classification of CC transactions in questionable & non-doubtful classes. Essentially, this method follows scoring procedure. In their study, database consisted of 62 regions with over 4,000 transactions. As similar point of view, training & testing models were utilized. Dissimilar types of rule were verified by different fields. Best rule is to have best prediction. Their technique has proven outcomes of real home insurance data, & is an effective way to combat credit card fraud.

Decision Tree- Decision perspective is a graphical demonstration of probable solution to an option based on positive circumstances. The decision view starts from root node, divided into separate spaces, which are linked to added nodes. Decision tree termination up node is named leaf node. At every node, decision view signifies an experiment, related by branch, representing its outcomes, & leaf node is class of labels. Through this strategic method to differentiation & decision-making, decision perspectives are usually simplified in a complex problem.

Artificial Neural Network(ANN)- ANN is mostinfluential classifiers with different characteristics among hidden patterns. ANN functions similarly to human brain. The first layer is input layer & last layer is output layer. It may have either any number of hidden layers. If neural networks have more hidden layer of stability, it is intensive learning. Each layer has dissimilar neurons & every neuron is associated with heavier edges. Every neuron of output has its private unit of action. This function is named activation function. E.g., various beginning functions are used: linear function, step function, threshold function, sigmoid function, & so on. There is commonly applied function is public sigmoid function.

Convolution Neural Network (CNN)- CNN is measure of intensive education. The feature map represents the hidden layer within the mapping. Each feature map represents a feature. The feature map in the compressing neurons of the process is called convolution. The feature of the sub-sample reduces the map parameters. The fully connected layer is the same neural network.

Outlier Detection- Outlier are basic method of substandard attention that can be applied to detect fraud. An observation that deviates so much from other explanations that it is suspicious another observation is known externally. This model uses unsupervised learning approach. In general, outcome of unread study is new description or demonstration of detected information, followed by better future decisions. Unfeasible approaches do not require prior information of fraudulent & non-fraudulent transactions, but rather sense variations by unfeasible learning behavior & uncommon transactions.

Clustering techniques- Two clustering methods to behavior fraud reported in Bolton & Hand (2002). Peer group study is system that identifies account that act otherwise as of others in a moment. These are some of the accounts that are called suspicious. & then there are cases of fraud. Peer cluster study behind assumption is that if an account is still operating differently for specified period of time, then this account needs for reported. Other method, Breakpoint Analysis, usages another theory that suggests that the card should be investigated if the change in card procedure is on separate beginning.

Logistic Regression- There are more & more statistical models that discriminate data mining functions such as study, regression analysis, & multiple logistic logic. Logistic regression (LR) is a set of predictive variables that are valuable to predicting presence or deficiency of attribute or outcome. This is parallel to linear regression model, but itis suite for model with reliant on variable dichotomies.

Deep learning - Deep Learning is a sophisticated technology that has recently attracted the attention of IT circles. Deep Learning Theory is an ANN with many hidden layers. In contrast, deep learning forward neural networks have only one hidden layer.

Rule based method- Association rules have been created by perceive fraud-based transactions & common transactions. In fraud detection, the rules created can be applied to categorize fraud & legal relations. There are rules for created behavior. This technique is related to the decision perspective.

Hidden Markov Model (HMM)- The HHM is modelingof hybrid, embedded stochastic procedure. This generalized process of complexity exceeds the Markov model. If the learner with a high potential probability does not approve the hidden Markov model bank transaction, this is measured dangerous & fake transaction. Baum Welch algo is applied for model learning, & K-Means algo to data classification. Model categorizes transactions in high, average, & low levels.

7. CONCLUSION

In this review paper, credit card investigations have been conducted on various methods of detecting fraud. First, it stated the importance of the topic & mentioned the current shortcomings in traditional practices. Counterfeit transactions have different levels of risk, & they must find ways to quickly& accurately detect high-risk transactions. Typical data mining methods are not sufficient to identify these transactions. Advanced algorithms should be used to find the best answer.

REFERENCE

[1] Clifton Phua, Vincent Lee, Kate Smith & Ross Gayler, "A Comprehensive Survey of Data Mining-based Fraud Detection Research", https://arxiv.org/ftp/arxiv/papers/1009/1009.6119.pdf 2014.

[2] Lakshmisri Surya, "Machine Learning-Future of Quality Assurance", International Journal of Emerging Technologies and Innovative Research (www.jetir.org | UGC and issn Approved), ISSN:2349-5162, Vol.6, Issue 12, page no. pp1078-1082, December-2019, Available at : http://www.jetir.org/papers/JETIR1912145.pdf

[3] Suman, Mitali Bansal, "Survey Paper on Credit Card Fraud Detection", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), ISSN: 2278 – 1323, Pp 827-832, Volume 3 Issue 3, March 2014. Lakshmisri Surya, "HOW GOVERNMENT CAN USE AI AND ML TO IDENTIFY SPREADING INFECTIOUS DISEASES", International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, Volume.6, Issue 1, pp.899-902, March 2018, Available at :http://www.ijcrt.org/papers/IJCRT1133873.pdf

[4] Bilonikar Priya, Deokar Malvika, Puranik Shweta, Sonwane Nivedita4, Prof.B.G.Dhake "Survey on Credit Card Fraud Detection Using Hidden Markov Model", International Journal of Advanced Research in Computer& Communication Engineering, ISSN (Online) : 2278-1021, ISSN (Print) : 2319-5940, Vol. 3, Issue 5, Pp 6445-6448, May 2014.

[5] Samaneh Sorournejad, Zahra Zojaji, Reza Ebrahimi Atani, Amir Hassan Monadjemi, "A Survey of Credit Card Fraud Detection Techniques: Data and Technique Oriented Perspective", sorournejad@yahoo.com, 1611.06439, Pp 1-26.

[6] Suman, Nutan, "Review Paper on Credit Card Fraud Detection", International Journal of Computer Trends and Technology (IJCTT) – volume 4 Issue 7–July 2013, ISSN: 2231-2803 , Pp 2206-2215.

[7] Vipul Patil, Dr. Umesh Kumar Lilhore "Survey on Different Data Mining & Machine Learning Methods for Credit Card Fraud Detection", International Journal of Scientific Research in Computer Science, Engineering and Information Technology © 2018 IJSRCSEIT, | Volume 3 | Issue 5 ISSN : 2456-3307, [(3)5 : 320-325.

[8] Mohammad, Sikender Mohsienuddin, Cloud Computing in IT and How It's Going to Help United States Specifically (October 4, 2019). International Journal of Computer Trends and Technology (IJCTT) – Volume 67 Issue 10 - October 2019, Available at SSRN: https://ssrn.com/abstract=3629018

[9] Mohammad, Sikender Mohsienuddin, DevOps Automation Advances I.T. Sectors with the Strategy of Release Management (December 12, 2019). International Journal of Computer Trends and Technology (IJCTT) – Volume 67 Issue 12 – Dec 2019, Available at SSRN: https://ssrn.com/abstract=3628988

[10] Manishaben Jaiswal " SOFTWARE ARCHITECTURE AND SOFTWARE DESIGN" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056, p-ISSN: 2395-0072, Volume: 06 Issue: 11, s. no -303 , pp. 2452-2454 , Nov 2019 Available at: https://www.irjet.net/archives/V6/i11/IRJET-V6I11303.pdf

[11] Manishaben Jaiswal "RISK ANALYSIS IN INFORMATION TECHNOLOGY" , International Journal of Scientific Research and Engineering Development (IJSRED) , ISSN:2581-7175, Vol 2-Issue 6, P110, pp. 857-860, November - December 2019 Available at: http://www.ijsred.com/volume2/issue6/IJSRED-V2I6P110.pdf

[12] Manishaben Jaiswal, Mehul Patel "THE LEARNING ON CRM IN ERPWITH SPECIAL REFERENCES TO SELECTED ENGINEERING COMPANIES IN GUJARAT", International Journal of Management and Humanities Scopus (IJMH) , published by Blue Eyes Intelligence Engineering & Sciences Publication (BEIESP), ISSN 2394-0913, Volume4 Issue-8, April 2020, Pg-117-126,Available At, http://www.ijmh.org/wp-content/uploads/papers/v4i8/H0798044820.pdf

[13] Maja puh, Ljiljana Brkic, " Detecting credit card fraud using selected machine learning algorithms", maja.puh@fe.hr, ljiljana.brkic@fe.hr, MIPRO 2019, May 20-24, 2019, Opatija Croatia, Pp 1250-1255.

[14] Anuruddha Thennakoon, Chee Bhagyani, Sasitha Premadasa, Shalitha Mihiranga, Nuwan Kuruwitaarachchi, "Real-time Credit Card Fraud Detection Using Machine Learning", 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence). 2019, IEEE, pp. 488-493.

[15] Chunzhi Wang Yichao Wang Zhiwei Ye Lingyu Yan Wencheng Cai Shang Pan, "Credit card fraud detection based on whale algorithm optimiz HG BP neural network", The 13th International Conference on Computer Science & Education (ICCSE 2018), 978-1-5386-5495-8/18/$31.00 ©2018 IEEE, Pp 614-617.

[16] Sahil Dhankhad, Emad A. Mohammed, "Supervised Machine Learning Algorithms for Credit Card Fraudulent Transaction Detection: A Comparative Study", 2018 IEEE International Conference on Information Reuse and Integration for Data Science, 978-1-5386-2659-7/18/$31.00 ©2018 IEEE, DOI 10.1109/IRI.2018.00025, Pp 122-125.

[17] Krishna Modi, Reshma Dayma, "Review On Fraud Detection Methods in Credit Card Transactions", 2017 International Conference on Intelligent Computing and Control(I2C2017),krishnamodi1994@gmail.com,ceradayma@gmail.com .

[18] Zahra Kazemi, Houman Zarrabi, "Using deep networks for fraud detection in the credit card transactions", IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI) I Dec. 22QG, 2017, 978-1-5386-2640-5/17/$31.00 ©2017 IEEE, Pp 0630-0633

[19] Kosemani Temitayo Hafiz, Dr. Shaun Aghili, Dr. Pavol Zavarsky, "The Use of Predictive Analytics Technology to Detect Credit Card Fraud in Canada", tkoseman@student.concordia.ab.ca,{shaun.aghili,pavol.zavarsky}@conc ordia.ab.ca, 2016, Pp 1-6.

[20] V.Mareeswari, Dr G. Gunasekaran, "Prevention of Credit Card Fraud Detection based on HSVM", International Conference On Information Communication And Embedded System(ICICES 2016), 978-1-5090- 2552-7.

[21] LAKSHMISRI SURYA , "ARTIFICIAL INTELLIGENCE IN PUBLIC SECTOR", IJIERT - International Journal of

Innovations in Engineering Research and Technology, Volume 6, Issue 8, ISSN : 2394-3696, Page No. 7-12

[22] Fahimeh Ghobadi, Mohsen Rohani, "Cost Sensitive Modeling of Credit Card Fraud Using Neural Network Strategy",Ghobadi.Fahimeh@Gmail.com,m.Rohani@niopdc.ir, 2016