

Fraud Detection of Credit Card using Data Mining Techniques



Avinash Sharma, Aaditi Verma, Dhananjay Gupta

Abstract: *The handling of credit card for online and systematic purchase is booming and scam associated with it. An industry of fraud detection where cumulative rise can have huge perk for banks and client. Numerous stylish techniques like data mining, genetic programming, neural network etc. are used in identify fraudulent transaction. In online transaction, Data mining acquire indispensable aspect in discovery of credit card counterfeit. This paper uses gradient boosted trees, neural network, clustering technique and genetic algorithm and hidden markov model for achieving upshot of the fraudulent transaction. These all model are emerging in identifying various credit card fraudulent detection. The indispensable aims to expose the fraudulent transaction and to corroborate test data for further use. This paper presents the look over techniques and pinpoint the top fraud cases.*

Keywords: *Machine Learning, Credit Card, Data mining, Algorithms.*

I. INTRODUCTION

The credit card is a lightweight handy nimble card that consist of identification material such as moniker or portrait and permit the person named on it to charge purchase or strong point to its account -charge for which they'll be payable episodically. Automated teller machine (ATM), store reader, bank and online internet banking system are places where the information of card is read. They have exclusive card number which is absolute sense. Its security commit on the corporal safety of the pliable card and the confidentiality of the card number. The credit card transaction become the utmost prevalent means of imbursement for both virtual and down transaction. Fraud is outlined as felonious dishonesty with earnest of achieving monetarist boost. Excessive addiction on internet has adored in credit card transaction. Nowadays, mercurial surge in the figure of credit card transaction which has coerce to hefty upsurge in fraudulent activities. Credit card fraud is a widespread word for theft and fraud stanch spending credit card as a counterfeit cradle of assets in a procedure.

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

Dr. Avinash Sharma, Professor, Maharishi Markandeshwar Engineering college, Mullana, Ambala (Haryana) Constituent institution of Maharishi Markandeshwar University, Mullana is NAAC accredited 'A' grade deemed university

Aaditi Verma, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Delhi NCR, India.

Dhananjay Gupta*, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Delhi NCR, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Credit card fraud fall under two categories inner card fraud and external card fraud. Inner card fraud befalls as an outcome of harmony between cardholder and bank by means of fabricated individuality to commit scam while the external card fraud embroils the use of purloined to acquire money throughout suspicious funds. A portion of researches have been ardent to finding of external card fraud. Detecting fraudulent transactions using Quaint method of blue-collar exposure is time unbearable plus uneconomical. Normally, statistical methods and sundry data mining algorithm are used to elucidate this fraud detection hitch. Extreme of the credit card scam detection scheme are built on artificial intelligence, Meta learning and pattern matching. In this paper, we'll emphasize on credit card deceit and revealing channels. Credit card fraud arise when one distinctive uses other individuals' card for the peculiar use without the acquaintance of the proprietor. Accordingly, we need an elucidation which curtails the unmitigated accessible limit on the credit card which is more protuberant to frauds. Data mining method is one noteworthy modus used to deciphering credit card fraud detection delinquent. Credit card fraud detection is the progression of pinpointing particular transaction that exist falsified into two classes of lawful and falsified transaction. A number of confronts are associated with the credit card fraud detection, namely fraudulent compartment silhouette is vigorous, that is counterfeit transaction manage to gaze lawful ones.

II. NUMEROUS PRACTICES USED IN CREDIT CARD FRUAD

The initiation of credit card has not just postulated us with the contentment and expediency but has also engrossed mischievous charismata as it the leisureliest way to net a mammoth expanse of money over an appropriate petite span of time.

A rare conjoint procedure that fraudster uses are:

- Copying the credit card and someway attainment grasp of the surreptitious dash of the punter.
- Purveyors storming extra money from the punter's credit card rivalled to the agreed and without the latter being sentient of the arraigned dough.
- Fraud detection is then carried out after observing a number of transaction and then identifying and classifying them into the genuine transaction and the fraudulent transaction.

III. NUMEROUS PRACTICES USED IN CREDIT CARD FRAUD

The biggest impediments associated with fraud detection is the absence of both the literature providing investigational upshots and real-world data for academic researches to perform experiments on. The main reason behind the hiding the data is customer's privacy.

To generate proper and accurate upshots we must follow some of the properties:

- Skewed distribution should be handled by the system, since a very small percentage of all credit card transaction is fraudulent.
- Noise should be handled properly with the proper means. Noise is the error present in the data, for example incorrect dates.
- Have common characteristics data is one of the major problems. Many trades may look like fraudulent transaction when actually they are unpretentious transaction and vice versa.
- System must be capable to familiarize themselves to different kind of fraud.

IV. CREDIT CARD FRAUD DETECTION METHODS

There exist numerous approaches that can be used to sense credit card fraud detection.

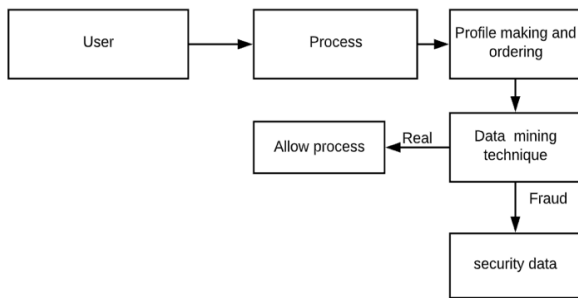


Fig 1: Data mining process

A. NAÏVE BAYES CLASSIFIER

A Supervised machine learning mode that practices instruction set with well-known goal modules to prophesy the class of conceivable cases. This algorithm was first introduced by John and Langley [2]. In Easiest terms, a Naïve Bayes method simulates the “occurrence or lack “of a scrupulous feature of a set is not based on the occurrence and lack of any other characters in similar set. According, to some trials it is seen that Naïve Bayes algorithm work well. However, this technique is named by the name “Naïve” because it naively assumes independence of the attributes given the class [2]. The grouping is completed by “Bayes” rule to analyze the likelihood of the class which is attribute of the operation.

B. GRADIENT BOOSTED TREES

Gradient boosted trees is an iterative procedure that syndicates simple parameterized function with” poor functioning to generate extremely precise prognostication instruction. Gradient boosted trees utilize an ensemble of weak learners to boost performance. This makes it good candidate model for predicting credit card fraud. It requires little data preprocessing and tuning of parameters while yielding interpretable upshots. Gradient boosted trees can

model complex interaction in a simple fashion and be used in both classification and regression with a variety of response distribution including Bernoulli, Gaussian, Poisson, Laplace. The tenets which are missing from the data are managed accordingly. Gradient boosted trees are already in use to manage the management fraud and giving accurate upshots. Gradient Boosted trees advance analytical clout through apt a decision tree to the model's residuals quite than the retort fickle. To update the residuals, new tree is added to it. In gradient boosted trees, trees are grown sequentially. To prevent model to overfit a knowledge rate multiplier is functional.

C. NEURAL NETWORK

Neural network is extolled for credit card fraud detection because it gives efficient upshots in myriad complications. For fraud detection, neural network work like a human brain working principal. Neural network stores the experience from the given datasets.

In credit card fraud detection, neural network splits the information into various categories; first one is based on cardholder income, occupation. Other one category will store the payment details like, number of large purchasing, location etc. These all details are going to help in investigating the future transaction whether the transaction is genuine or fraud.

Neural network contains different layers:

Inner Layer: This layer contains input nodes, it will identify the cardholder use the information it will check the characteristics of the transaction.

Hidden Layer: Hidden Layer are used to identify the transaction fraud or not.

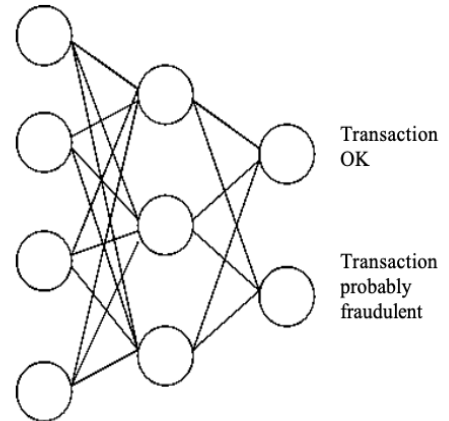
Output Layer: It gives output between 0 and 1.

Age, Income, Occupation, of card Holder

Number of large purchases on the card

Frequency of large purchases

Location where large purchase took



Input Layer Hidden Layer Output Layer

Fig 2: Layer of neural network in credit card [6]

D. HIDDEN MARKOV METHOD

Hidden Markov model is a conformist of conditions with the possibility circulation. All ceremonial gives an upshot bestowing to the probability distribution which comes under a meticulous ceremonial. The state is visible but the only outcome or observation are can be visible that is called Hidden Markov Model [12].

In credit card fraud detection, Hidden Markov Method helps in decree the fraud transaction by the demeanor of cardholder.

With the help of bygone behavior and operation, spending performance is measured. Aspects like operation amount, IP address, shipping address & location of past transaction etc. Three types of categories are there for cardholder:

- 1.Low spending behavior
- 2.Medium spending behavior
- 3.High spending behavior

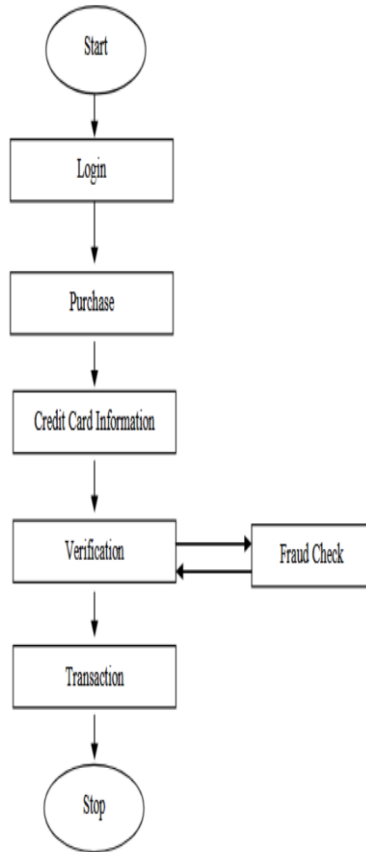


Fig 3: Hidden Markov Method

If cardholder spends less amount then they come under low spending behavior, if the cardholder spending medium level then they come under in medium spending otherwise the cardholder spending high amount they come under the category of high spending behavior.

E. Genetic Algorithm

Genetic Algorithm are evolutionary algorithm which can obtain the better upshot in jettisoning the fraud. A high significance is given to flourish resourceful and assured electronic payment method to detect whether a transaction is fraudulent or not. An Optimization technique that portray natural evolution process. The genetic algorithm amends the population of artificial structure through the solicitation of initialization, selection, crossover and modification operators. Is gives better optimum solution and deceit detection are spawned with the support of given dataset. This algorithm works in the manner of well-organized and protected electronic payment detect whether a transaction is fraud or not [11].

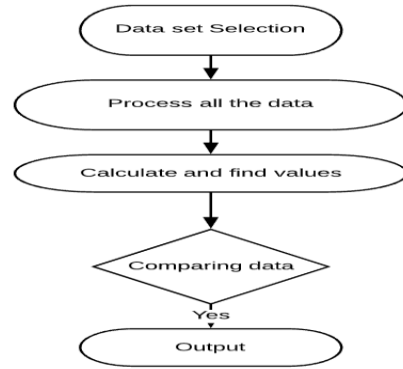


Fig 4: Fraud detection using genetic algorithm

In credit card transaction there are different type of attributes. First step is select the dataset which one going to the process and including all the details of cardholder. Calculating the values with the help of frequency use of card, bank balance, location where they use credit card on that particular transaction, daily spending. After comparing each transaction, we will finally know the transaction is fraud or legal.

V. CONCLUSION

In the past year, credit card usage has been increased very much. Increasing in credit card usage there is increase in counterfeit. There exist many techniques to identify fraud. Main aim is toward understanding the best technique that identifies the fraud cases. Best technique used to detect the fraud transaction is data mining or combination of data mining techniques. The best way to detect the fraud is to finding from the history of transaction. This paper contains the finding of an extensive literature survey kindred of the usability of data mining technique in credit card fraud detection. A list of data mining models apt for use in credit card fraud detection were discussed and difficulty faced in fraud detection. Major problem faced was the absence of dataset convenient for training the models. Comparing the upshots, we can see that best data mining technique are Hidden Markov method and Gradient boosted trees. But there is a possibility of enlargement in both the method to avoid the fraud more in prospect.

REFERENCES

1. Gabriel Rushin ; Cody Stancil ; Muyang Sun ; Stephen Adams ; Peter Beling , “Horse race analysis in credit card fraud-deep learning, logistic regression and Gradient Boosted Trees.” IEEE System and information Engineering Design Symposium(SIEDS), 2017.
2. G.H John,P.Langley(1995). Estimating continuous distribution in Bayesian classifiers. In: Proceeding of the 11th conference on Uncertainty in Artificial Intelligence (1995) ; 338-345.
3. Guelman, Leo. (2012). “Gradient boosting trees for auto insurance loss cost modeling and prediction. Expert Syst. Appl.” 39.36593667.10.1016/j.eswa.2011.09.058.
4. Zare, Masoume&Shamsolmoali, Pourya. (2015). Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier. Procedia Computer Science.48.679-686. 10.1016/j.procs.2015.04.201.
5. M. Zareappor,Seeja , K.R. & M. Alam , Afshar(2012). Analyzing Credit Card: Fraud Detection Technique Based on Certain Design Criteria. International Journal Of Computer Application,52(3):35-42.
6. RaghavendraPatidar, Lokesh Sharma, “Credit Card Fraud Detection Using Neural Network”, International Journal of Soft Computing and Engineering (IJSCE)ISSN: 2231-2307, Volume-1, Issue-NCAI2011, June 2011.

Fraud Detection of Credit Card using Data Mining Techniques

7. Zare, Masoume & K.R., Seeja & Alam, Afshar. (2012). Analysis on Credit Card Fraud Detection Techniques: Based on Certain Design Criteria. International Journal of Computer Applications. 52. 35-42. 10.5120/8184-1538.
8. Trivedi, Ishu. "Credit Card Fraud Detection." (2016).
9. S.Vimala, K. C. Sharmili; "Survey paper for Credit Card Fraud Detection Using Data Mining Techniques," IJRSET vol.6, special Issue 11 September 2017.
10. K.RamaKalyani, D.UmaDevi, "Fraud Detection of Credit Card Payment System by Genetic Algorithm", International Journal of Scientific & Engineering Research Volume 3, Issue 7, July-2012.
11. Pooja Chougule#1, A.D. Thakare, Prajakta Kale, Madhura Gole, Priyanka Nanekar, "Genetic K-means Algorithm for Credit Card Fraud Detection", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (2), 2015.
12. Mr.P.Matheswaran, Mrs.E.SivaSankari ME, Mr.R.Rajesh, "Fraud Detection in Credit Card Using Data Mining Techniques, Volume II, Issue I", IJRSET-February-2015.

AUTHORS PROFILE



Presently Professor, Maharishi Markandeshwar Engineering college, Mullana, Ambala (Haryana) Constituent institution of Maharishi Markandeshwar University, Mullana is NAAC accredited 'A' grade deemed university & the first private engineering college of Haryana established in 1995. Also Dean Faculty of Engineering and Technology & Member of Board of Studies & DRC Committee for Research. Ex-Principal & Professor, Rajasthan College of Engineering for Women (Leading Women's Engineering College in the state of Rajasthan). Publications: International Journals: Published: 45 (Accepted: 61); International Conferences: Published: 90 (Accepted: 102) National Conferences & Workshops: 75 (Accepted: 110 & more) Text Books/EDITED: 06. Total Experience: 20 years (10 years PG) + 03 years Research International Conferences Organized: 10. He is an active member of advisory/technical program committee of reputed International/National conferences & reviewer of number of reputed Journals e.g. Springer, Elsevier Journal Computers & Electrical Engineering. Approximate 20 years of rich experience in Teaching, research and industry managing technical institution, serving in all capacity including Head of Department, Professor, Controller of Examination, Dean Academics Affairs, Principal etc. Played leading role in accreditation of the institution and ISO 9001:2000 certification. (including 05 years of industrial/research experience)



Aaditi Verma is a Bachelor of Technology student in SRM Institute of Science and Technology Department of Computer Science. Her areas of interest are Machine Learning and Cloud Computing with Big Data.



Dhananjay Gupta is a Bachelor of Technology student at SRM Institute of Science and Technology Department of Computer Science and Engineering. His areas of interest are Machine Learning and Artificial Intelligence.