

Detect Frauds in Credit Card using Data Mining Techniques



Mehak Mahajan, Sandeep Sharma

ABSTRACT: In today era credit card are extensively used for day to day business as well as other transactions. Ascent within the variety of transactions through master card has junction rectifier to rise in the dishonest activities. In trendy day's fraud is one in every of the most important concern within the monetary loses not solely to the merchants however additionally to the individual purchasers. Data processing had competed a commanding role within the detection of credit card in on-line group action. Our aim is to first of all establish the categories of the fraud secondly, the techniques like K-nearest neighbor, Hidden Markov model, SVM, logistic regression, decision tree and neural network. So fraud detection systems became essential for the banks to attenuate their loses. In this paper we have research about the various detecting techniques to identify and detect the fraud through varied techniques of data mining.

KEYWORDS: Credit Card, Types of frauds, Data Mining techniques

I. INTRODUCTION

Fraud in credit card may be the provision for thievery and fraud committed victimization or involving payment card like MasterCard. The paper study the analysis and detection of MasterCard frauds and is employed to attenuate the credit card frauds. 2 mechanism are planned to avoid fraud and losses. They're Fraud Detection and Fraud hindrance. Fraud hindrance is an intense technique that stops frauds from being continuing whereas Fraud Detection is employed once a deceptive group action is tackle by the fraudster. Group action with MasterCard is polished in 2 that are physical transaction (in which credit card is directly concerned and in digital transaction credit card isn't used).

1.1 TYPES OF FRAUDS:

(a) **APPLICATION FRAUD:** Application fraud happens once people apply for credit or a replacement MasterCard in your name. They'll first of all steal the supporting document that are then went to justify their fallacious activity. Banks have various measures to prevent this kind of fraud from being happening. The foremost necessary one is requiring the initial documents solely.

(b) **CNP (Card Not Found) FRAUD:** If someone is aware of the termination date and account range of your card, they will commit CNP fraud against you. This could be done through phone, email or net. This suggests that if someone uses your card while not being in physical possession of it. A lot of usually merchants would force the verification code, creating CNP fraud slightly harder,

however if fraudster will get your account range, they'll get your range too. As there are solely 999 potential combination for the verification code and fraudster will work out the correct range.

(c) **LOST AND PURLOINED CARD FRAUD:** one among the sort of frauds is lost and stolen card fraud. Here the cardboard are going to be taken from your procession, either through thievery otherwise you could lost it. The criminals UN agency can get your card will then use to create the payments through it. It is troublesome to try and do this through machines as they'll enkindle the PIN number. But it's simple enough to use a found or purloined card to create on-line purchases. So, it's necessary to cancel your card as shortly as you understand that the cardboard is missing.

(d) **CARD ID THEFT:** It happens once the small print of your card has been notable to the criminal and this data is then went to take over your card account or open a replacement one. This can be one among the foremost troublesome styles of frauds to spot and to endure, as a result of it take a protracted time before you may recognize that it's happened.

(e) **FAKE CARD:** It takes lots of time to form a faux credit cards, however there are several techniques used to create false cards in order to stop the criminal.

(f) **ACCOUNT TAKEOVER:** The foremost common kind of MasterCard fraud is Account Takeover. The criminal can get the account data and connected documents that are sometimes done on-line. The account can be takeover by the criminal either providing the cardboard range or the account range of the client. During this case a criminal can contact cardboard establishment, disguise because the real cardholder to raise him to send the small print to a replacement address.

II. LITERATURE SURVEY:

Sathyapriya, Dr. V. Thiagarasu [1] used the unsupervised learning algorithm i.e. K-means clustering algorithm to facilitate the pattern recognition work and is based on the past activity of a credit card users in which the conventional usage pattern is applied.

Janaki K, B.V. Harshitha, Keerthana.S, Harshitha .Y.V, Ramyashree.K, used the hybrid method that are the combination of three methods that are decision tree, random forest and Naive Bayes algorithm to detect the frauds where x-axis represents the different methods and y-axis represents the year.

Aswathy MS, LIJI Samuel [2] has used most commonly technique like rule induction, SVM, Decision Tree, ANN, Logistic Regression and Genetic algorithm. The foremost algorithm that is used to detect the fraud is neural network. Moreover, these algorithm may be used as a single model or can be used in combination.

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Mehak Mahajan*, Dept. of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, Punjab, India

Dr. Sandeep Sharma, Dept. of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, Punjab, India

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

Detect Frauds in Credit Card using Data Mining Techniques

S.Vimala, K.C. Sharmili [3] has used the assorted techniques that are used to avoid that risk from various types of attacks like decision tree, neural network, K-mean clustering, Genetic Algorithm, Hidden Markov Model. Decision tree and Hidden Markov Model are the most effective techniques to detect frauds.

Goncalo Figueira, Nuno Carneiro and Miguel Costa [4], for classification problem they have chosen the method of supervised learning. SVM, Logistic Regression and Random Forests are the three different models they opted.

Okokpujie Kennedy O, F.Olajide, S.N. John, Chinyere Grace Kennedy and C.Anele [5].herewith addresses the fraud detected in bank via the employment or the techniques used in knowledge mining like association, classification, clustering, prediction and sequential patterns to investigate the client data so as to identify the patterns that may results in frauds. As per the authors association rule mining is that technique which is created by interpreting data for frequent if/then pattern and to spot the foremost relationship the confidence and criteria support is used.

Suchita Anand Padvekar, Pragati Madan Kangane, Komal Vikas Jadhav [6] has worked on the hidden markov model that are used throughout the transactions through which high fraud coverage as well as low false alarm rate are controlled. Saravanan Sagadevan, Nurul Hashimah Ahamed Hassain Malim and Ong Shu Yee [7] has mentioned about the supervised based classification using the Tree Augmented Naïve Bayes, Bayesian Network Classifiers namely K2, Logistic and J48 classifiers.

Masoumeh Zareapoor, Pourya Shamsolmoali [8] have investigated on five states techniques performance to predict credit fraud like SVM, Naïve Bayes, Bagging ensemble classifier, KNN. They found that bagging ensemble is the best technique as attribute values of bagging ensemble is independent to the data.

A.D. Thakare, Pooja Chougule, Prajakta Kale, Priyanka Nanekar and Madhura Gole, [9] has used various techniques to detect fraud detection like K-Mean Algorithm, Dempster-Shafer Theory, Genetic Algorithm, Hidden Markov Model, Biometrics. The simple genetic and K-means algorithms shows the significant results.

Suman Mitali Bansal [10] have used unsupervised and supervised learning techniques which help us to distinguish two types of frauds that are detected in credit card that are legitimate and fraudulent transactions. In order to detect the fraud transaction they worked on unsupervised learning technique which have a tendency to train network.

R Gayathri, A Malathi [12] has used classification techniques to detect fraud are Naive Bayes, Neural Network, SVM, K-NN and decision tree.

Suman, Nutan has come back to the conclusion that there are multiple techniques to detect the frauds. The techniques are Gas Algorithm, Hidden Markov Model, Bayesian Network, Fusion approach using Bayesian learning and Dempster-shafer theory, Neural Network, Genetic Algorithm, Decision Tree and Logistic Regression. The chances of fraud may be predicted if combination of frauds is applied to detect the frauds in bank.

Tejpal Singh, Amit Sinhal [11] have implemented or enforced the method known as Hidden Markov Model which will simply detect and remove the complexity of system. Hidden Markov Model is used to detect whether the incoming transaction is fraudulent or not.

Masoumeh Zareapoor, Seeja KR, M .Afshar Alam[13] have proposed the fraud detection methods based on credit card like Artificial Immune System(AIS), Genetic Algorithm, Fuzzy System ,Neural Network, Rule-Induction Techniques, Decision Tree, SVM, K-NN. To create the classifiers, we can use these techniques using enable or meta- learning.

Francisca Nonyellum Ogwueleka [14], the technique used to detect the fraud in card is neural network which is based on unsupervised method. The carrying out problem of optimal classification for each transaction can be solved using the technique known as self-organizing map neural network in which prior output is unknown.

V.Dheepa, Dr.R. Dhanapal [15] have used the three methods to detect the frauds. Firstly, classify the fraudulent and legal transactions the clustering model is used. Secondly, from the probability density of credit card user's past behavior we can calculate the current probability by using the Gaussian Model. Lastly, to describe the statistics of different fraud situation the Bayesian networks are used. All three models can be useful to enhance the system of fraud detection.

AUTHOR	YEAR	CLASSIFICATION APPROACH USED	OTHER MODEL/TECHNIQUES
M Sathyapriya, Dr V Thiagarasu	2019	Not Used	Hidden Markov Model, Clustering, k-mean clustering algorithm
Janaki K, B.V. Harshitha, Keerthana.S, Harshitha .Y.V, Ramyashree.K	2019	Decision Tree, Naïve Bayes, Random Forest	Not Used
Aswathy MS, LIJI Samuel	2018	SVM, Logistic Regression, Neural Network, Decision Tree.	Rule Induction, Genetic Algorithm
S Vimala, KC Sharmili	2017	Decision Tree, Neural Network	K-mean clustering, Hidden Markov Model, Genetic Algorithm
Nuno Carneiro Miguel Costa et.al.	2017	Random Forest, Logistic Regression, Support Vector Machine	Not Used
SN John, Okokpuje Kennedy O, Chinyere Grace Kennedy , C.Anele, F.Olajide	2016	Classification	Association Rule Clustering, Prediction
Suchita Anand Padvekar, Pragati Madan et.al.	2016	Not used	Hidden Markov Model

Ong Shu Yee, Saravanan Sagadevan et.al	2015	Supervised Based Classification using Bayesian Classifier like Naïve Bayes , K2, TAN, Logistic and J48	Not used
Masoumeh Zareapoor, Pourya Shamsolmoali	2015	Bagging Ensemble Classifier based on Decision Tree ,Naive Bayes Classifier, SVM, K-Nearest Neighbor Algorithm, K-NN	Not Used
Pooja Chougule, A.D.Thakare , Madhura Gole et.al.	2015	Not used	K-Mean Algorithm, Hidden Markov Model, Biometrics, Dempster-Shafer Theory, Genetic algorithm
Suman Mitali Bansal	2014	Supervised	Unsupervised
Ashphak Khan, Tejpal Singh et.al	2012	Not Used	Hidden Markov Model
R Gayathri, A Malathi	2013	Support Vector Machine, Neural Network, Naive Bayes, Decision Tree, K-NN.	Not used
Suman , Nutan	2013	Bayesian Networks, Neural Network , Logistic Regression	Gas Algorithm, Genetic Algorithm, ,Fusion Approach Using Dempster- Shafer Theory , Bayesian Learning, Hidden Markov Model
Masoumeh Zareapoor, Seeja KR ,M .Afshar Alam	2012	K-NN, Neural Network, SVM, Decision Tree.	Rule-Induction Techniques, Fuzzy System, Artificial Immune System, Genetic Algorithm
Francisca Nonyellum Ogwueleka	2011	Neural Network	Not Used
V.Dheepa , Dr R Dhanapal	2009	Bayesian Networks	Clustering Model, Gaussian Mixture Model

1. PARAMETERIC ANALYSIS:

AUTHOR	YEAR	PARAMETERS				
		Accuracy	Time	Count	Speed	Cost
M.Sathyapriya, Dr.V. Thiagarasu	2019	✓	✗	✗	✗	✗
Janaki K, B.V. Harshitha, Keerthana.S, Harshitha .Y.V, Ramyashree.K	2019	✓	✗	✗	✗	✗
Aswathy MS , LIJI Samuel	2018	✗	✗	✗	✗	✗
S.Vimala, KC Sharmili	2017	✗	✗	✗	✗	✗
Nuno Carneiro .Goncalo Figueira, Miguel Costa	2017	✗	✓	✗	✗	✗
SN John, Okokpuje Kennedy O, Chinyere Grace Kennedy , C.Anele, F.Olajide	2016	✗	✗	✓	✗	✗
Suchita Anand Padvekar, Pragati Madan Kangane, Komal Vikas Jadhav	2016	✓	✗	✗	✗	✗
Ong Shu Yee, Saravanan Sagadevan et.al.	2015	✓	✗	✗	✗	✗
Masoumeh Zareapoor, Pourya Shamsolmoali	2015	✗	✓	✗	✗	✗
Pooja Chougule, A.D. Thakare et.al.	2015	✗	✗	✗	✗	✗
Suman, Mitali Bansal	2014	✗	✗	✗	✗	✗
R.Gayathri, A.Malathi	2013	✗	✗	✗	✗	✓
Suman , Nutan	2013	✗	✗	✗	✗	✗
Ashphak Khan, Tejpal Singh, Amit Sinhal	2012	✓	✗	✗	✗	✗
Masoumeh Zareapoor, Seeja KR ,M .Afshar Alam	2012	✓	✗	✗	✓	✓
Francisca Nonyellum Ogwueleka	2011	✓	✗	✗	✗	✗
V.Dheepa , Dr R Dhanapal	2009	✗	✗	✗	✗	✗

Detect Frauds in Credit Card using Data Mining Techniques

2. Gaps In Existing Literature:

Author	Year	Proposed Work	Gap/ Future Work
M.Sathyapriya, Dr.V. Thiagarasu	2019	The proposed system (credit card expenditure model) is better in reducing the rate of false alarm rates that can be achieved by examining the relationship between the transactions marked by the actual fraud and the transactions guessed as fraud.	Not Mentioned
Janaki K, B.V. Harshitha, Keerthana.S, Harshitha .Y.V, Ramyashree.K	2019	The proposed system is help to detect and prevent the fraudulent transaction and activities so to decrease the unit of dropping in economic industry.	Web based study models are used to reached out by considering the strategies.However the other models can be examined through internet study. The location of the extortion cases can be quickly trapped by the use of the internet
S.Vimala, KC Sharmili	2017	We conclude that the best way to find fraud detection is possible through decision tree and Hidden Markov model.	In order to avoid fraud in future there is need of further improvement in hidden Markov Model and decision tree.
Nuno Carneiro .Goncalo Figueira et.al.	2017	The development of automatic and manual system was used to extract the data from database which further help to prepare it, trains the algorithm and run a web application.	There is no such difference in the performance of using balanced set of observation rather than using the bigger unbalanced full sets of records.
Ong Shu Yee, Saravanan Sagadevan et.al.	2015	Classification metrics are tested by using five Bayesian classifiers. The interpretation conducted using two datasets was classified as: the first set of data was dummy which tells the characteristics of credit card data and other was a newly transformed dataset using data normalization and Principal component Analysis techniques.	In future, we can explore fraud detection in credit card using real time data.
Pooja Chougule, A.D. Thakare et.al.	2015	The clusters generated by K-Means clustering algorithm is applied through genetic algorithm and the outcome generated by K-mean clustering will optimize by genetic algorithm. The scalability of system in terms of rules is measured by rule engine.	The accuracy of system can be improved by adding various rules in rule engine which are extended by models (genetic algorithm and K-mean clustering).
Suman, Mitali Bansal	2014	Frauds can be detected by various techniques. The approach is based on unsupervised learning in which networks are trained to detect the fraud.	Not Mentioned
R.Gayathri, A.Malathi	2013	The various methods used to detect frauds using data mining based classification techniques are divide into four different classification such as SVM, K-NN, Decision tree and neural network.	To overcome the problem of missing value, handling voluminous data precisely and handling incomplete data set, various techniques for complete set of pattern recognition can be develop.
Ashphak Khan, Tejpal Singh et.al.	2012	Hidden Markov model is used to detect and remove the complexity of the system.	Not Mentioned
Masoumeh Zareapoor, Seeja KR et.al.	2012	Compared the different methods based on different fraud detection method and had compared methods with different parameters like speed, cost and accuracy	The problem of classification with variable misclassification cost and higher accuracy can be develop by using hybrid approach.
Francisca Nonyellum Ogwueleka	2011	The existing banking software use the design of neural network credit card fraud detection and this prove to be effective and efficient method of fraudulent transactions.	Operational fraud detection system should be used
V.Dheepa, Dr .Dhanapal	2009	Three methods are used to detect the fraud detection that are clustering , Gaussian mixture and Bayesian networks	The health insurance fraud detection can be extended by using these models.

III. CONCLUSION:

In financial services, detecting fraud in credit card is very serious problem. Loss due to the credit card fraud is increasing day by day in the e-commerce. In this paper we have discussed various techniques of data mining through which we can detect the fraud. Various techniques like Hidden Markov Model, K-mean clustering algorithm, K-nearest neighbor, Decision Tree, Fusion approach due using dumpster Shafer, Bayesian Network, Neural Network, SVM and Logistic Regression are used. In future work for research, we will use hybrid method for detecting fraud using these above techniques.

REFERENCES:

1. M Sathyapriya, Dr.V. Thiagarasu, " A Cluster Based Approach for Credit Card Fraud Detection System using HMM with the Implementation of Big Data Technology" ,International Journal of Applied Engineering Research ISSN Volume-14,Number 2(2019)pp.393-396.
2. Aswathy MS, LIJI Sameul, "Survey on Credit Card Fraud Detection", International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 11, Nov2018, pp.1291-1294.
3. S.Vimala, K.C.Sharmili, "Survey Paper for Credit Card Fraud Detection Using Data Mining Techniques", International Journal of Innovative Research in Science, Engineering and Technology, Vol.6, Special Issue 11, September 2017,pp.357-364.
4. Nuno Carneiro, Goncalo Figueira, Miguel Costa, " A data mining based system for credit card fraud detection in e-tail", Elsevier,pp.91-101.

5. Suchita Anant Padvekar, Pragati Madan Kangane, Komal Vikas Jadhav, "Credit Card Fraud Detection System", International journal of engineering and Computer Science ,Volume 5 Issue 4 April 21016,pp.16183-16186.
6. S.N.John, Okokpujie Kennedy O, C.Anele, F.Olajide, Chinyere Grace Kennedy , "Real Time Fraud Detection in the Banking Sector Using Data Mining Techniques/Algorithm", Research Gate Publication, December 2016,pp.1186-1191.
7. N.Sivakumar, Dr.R.Balasubramanian, "Fraud Detection in Credit Card Transaction: Classification, Risks and Prevention Techniques", International Journal of Computer Science and Information Technologies, Volume (2), 2015, pp.1379-1386.
8. Ong Shu Yee, Saravanan Sagadevan and Nurul Hashimah Ahamed Hassain Malim , "Credit Card Fraud Detection Using Machine Learning As Data Mining Technique" ,Journal of telecommunication, Electronic and Computer Engineering, Vol.10,No.1-4.
9. Masoumeh Zareapoor, Pourya Shamsolmoali, "Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier", Elsevier, 2015, pp.679-685.
10. Pooja Chougule, A.D.Thakare , Prajakta Kale, Madhura Gole, Priyanka Nanekar "Genetic k-mean algorithm for credit card fraud detection", International Journal of Computer Science and Information Technologies, Volume 6(2), 2015, pp.1724-1727.
11. Suman, Mitali Bansal, "Survey Paper on Credit Card Fraud Detection", International Journal of advanced research in Computer Engineering And Technology, Volume 3 Issue 3, March 2014.
12. R.Gayathri, A.Malathi, "Investigation of Data Mining Techniques in Fraud Detection: Credit Card", International Journal of Computer Applications, Volume 82-No.9, November 2013.
13. Suman, Nutan, "Review Paper on Credit Card Fraud Detection", International Journal of Computer Trends and Technology, Volume 4, Issue 7-July 2013.
14. Ashphak Khan, Tejpal Singh, Amit Sinhal, "Implement Credit Card Fraudulent Detection System Using Observation Probabilistic In Hidden Markov Model", Research Gate Publication, 2012 .pp.1-6.
15. Masoumeh Zareapoor, Seeja.K.R, M.Afshar Alam,
16. "Analysis of Credit Card Fraud Detection Techniques: based on certain design criteria", Research Gate Publication, Volume 52-No.3, August 2012.
17. Francisca Nonyelum Ogwueleka, "Data Mining Application in Credit Card Fraud Detection System", Journal of Engineering Science and Technology, Vol. 6, No.3 (2011) 311-322.
18. V.Dheepa, Dr .Dhanapal, "Analysis of Credit Card Fraud Detection Method", International Journal of Recent Trends in Engineering, , Volume 2, No.3 November 2009.
19. Janaki K, B.V. Harshitha, Keerthana.S, Harshitha .Y.V, Ramyashree.K "A Hybrid Method for Credit Card Fraud Detection Using Machine Learning Algorithm", International Journal of Recent Technology and Engineering , Volume-7, April 2019