Review on Credit Card Fraud Detection Using Machine Learning

Bhumika Maheriya,
M.Tech. in Cyber Security
Sankalchand Patel University

Mehul S. Patel,
Assistant Professor, IT Department,
Sankalchand Patel University

Abstract: MasterCard extortion happens when an individual uses another person's charge card for individual reasons or gain while the proprietor doesn't know about this reality. In this our mean to identify and discover extortion in charge card. For discovery of Visa misrepresentation have numerous methods to look for same objective for staying away from charge card extortion. In this paper I use to AI procedures for identification the extortion. A directed learning calculation utilized for preparing information and produces a surmised work and an unaided learning calculation utilized for bunch investigation. In this paper different discovery strategies subtleties of MasterCard misrepresentation.

Keywords: Decision Tree, Neural Networks, Logistic regression, bayesalgorithm, support vector machine, k-nearest neighbor algorithm.

1. INTRODUCTION

Monetary extortion is a developing worry with broad outcomes in the government, corporate associations, account industry, In this day and age high reliance on web innovation has appreciated expanded MasterCard exchanges yet MasterCard misrepresentation had likewise quickened as on the web and disconnected exchange. As Visa exchanges become an inescapable method of installment, center has been given to late computational approaches to deal with the credit card extortion issue. There are numerous extortion discovery arrangements and programming which forestall fakes in organizations, for example, charge card, retail, online business, protection, and industries.[1] by and large, there are three classes of MasterCard extortion specifically, customary fakes (for example taken, phony and fake), online fakes (for example bogus/counterfeit trader destinations), and shipper related fakes (for example trader intrigue and triangulation) .

HIGH RISK COUNTRY

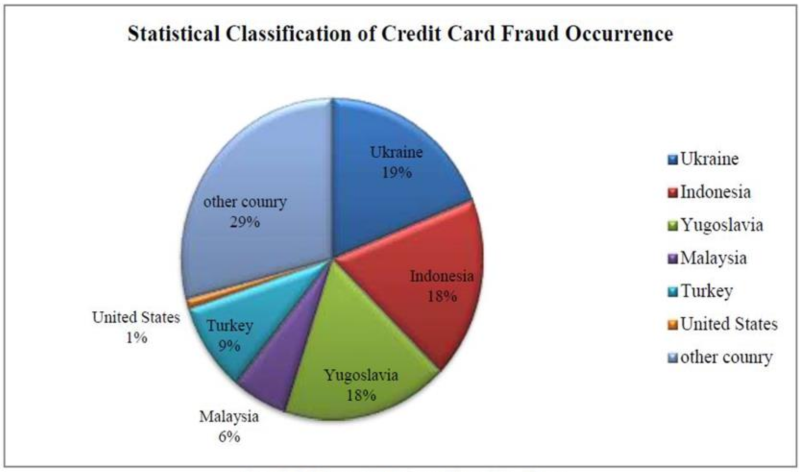

Fig1. High risk countries facing credit card fraud threat

[1]

2. Methodology

Extortion location is a double order task in which any transaction will be anticipated and named as a misrepresentation or genuine. In this paper cutting edge grouping methods were gone after for this undertaking and their exhibitions were looked at. The accompanying subsections momentarily clarify these characterization procedures, informational index and measurements utilized for execution measure.

What is credit card fraud?

Fraud detection is a set of activities undertaken to prevent money or property from being obtained through false pretenses. Fraud detection is applied to many industries such as banking or insurance. In banking, fraud may include forging checks or using stolen credit cards.

Credit Card Fraud Detection Techniques

- Decision Tree.
- Genetic Algorithms and a Range of Additional Algorithms.
- Clustering Techniques.
- Neural Networks.
- Logistic regression
- Naive Bayes Classifiers.
- K-Nearest Neighbor Algorithms.
- Support Vector Machines (SVMs)
- Bagging Ensemble Classifier.

3. Classification of algorithms with detail explanation:

Logistic Regression

Logistic Regression is a supervised classification method that returns the probability of binary dependent variable that is predicted from the in depend variable of dataset that is logistic regression predict the probability of an outcome which has two values either zero or one, yes or no and false or true. Logistic regression has similarities to linear regression but as in linear regression a straight line is obtained, logistic regression shows a curve. The use of one or several predictors or independent variable is on what prediction is based, logistic regression produces logistic curves which plots the values between zero and one. Regression is a regression model where the dependent variable is categorical and analyzes the relationship between multiple independent variables. There are many types of logistic regression model such as binary logistic model, multiple logistic model, and binomial logistic models. Binary Logistic Regression model is used to estimate the probability of a binary response based on one or more predictors.

ANN

An artificial neural network (ANN) is the piece of a computing system designed to simulate the way the human brain analyzes and processes information. It is the foundation of artificial intelligence (AI) and solves problems that would prove impossible or difficult by human or statistical standards.

- How can neural networks be used to detect fraud?

By employing neural networks, effectively, banks can detect fraudulent use of a card, faster and more efficiently. ... In more practical terms neural networks are non-linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data.

- How can neural networks be used to detect fraud?

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 with the pattern of the original card holder on which the neural network has been trained, if the pattern matches the neural network declare the transaction ok.

Bayesian technique

Bayesian inference is a method of statistical inference in which Bayes' theorem is used to update the probability for a hypothesis as more evidence or information becomes available. Bayesian inference is an important technique in statistics, and especially in mathematical statistics.

$$\boxed{P(A|B) = \frac{P(B|A)P(A)}{P(B)}}$$

Figure - Bayes Rule (James V Stone, 2013)

4. Model building:

Dataset:

Our objective is to credit card fraud detection is to find pure data.The datasets contains transactions made by credit cards in September 2013 by european cardholders. This dataset presents transactions that occurred in two days, where we have 492 frauds out of 284,807 transactions. The dataset is highly unbalanced, the positive class (frauds) account for 0.172% of all transactions.[3]

Parameters:

Four basic metrics are used in evaluating the experiments, namely True Positive (TPR), True Negative (TNR), False Positive (FPR) and False Negative (FNR) rates metric respectively.

- TPR= TP/P
- TNR=TN/R
- FNR=FP/R
- FPR=FP/R

For evaluating algorithm, following parameter required to measured and analyzed.

- ACCURACY= TP+TN/TP+FP+TN+FN
- SENSITIVITY=TP/TP+PN
- SPECIFICITY=TN/FP+TN
- PRECISION= TP/TP+FP

- When FN , FP ,TP,TN, and are the quantity of false negative false positive ,genuine positive and genuine negative experiments ordered while complete number of positive and negative class cases under test are addressed by P and N. Cases characterized appropriately as discredit are named with genuine negative and cases delegated positive which are really sure are named with Genuine positive .Cases named positive however are negative cases are named as bogus positive and cases named negative yet are genuinely sure are named as bogus negative. The presentation of Classifiers is assessed dependent on exactness, accuracy, explicitness and affectability. Sensitivity (Review) gives the exactness on certain (extortion) cases grouping . Particularity gives the exactness on negative (authentic) cases characterization. Exactness gives the precision in cases named extortion (positive).

5. Literature review

In this paper mainly focus on overview of detection types and techniques.in machine learning two mainly techniques use for a detection supervised and unsupervised. In this paper logistic regression,svm, knn etc. Based machine learning approach is utilized to detect credit card fraud.Extortion go about as the unlawful or criminal trickery proposed to result in monetary or individual advantage. It is a conscious demonstration that is illegal, rule or strategy with an intend to accomplish unapproved monetary benefit. Numerous written works relating to inconsistency or misrepresentation identification in this area have been distributed as of now and are accessible for public usage.[4]

In this paper strategic relapse, based AI approach is used to identify MasterCard extortion. The outcomes show strategic relapse based methodologies beats with the most noteworthy exactness and it very well may be viably utilized for misrepresentation specialists. An extensive overview directed by Clifton Phua and his partners have uncovered that techniques employed in this space incorporate information mining applications, robotized misrepresentation discovery, ill-disposed location. In another paper, Suman, Exploration Researcher, GJUS&T at Hisar HCE introduced procedures like Regulated and Unaided Learning for Visa misrepresentation discovery. Despite the fact that these strategies and calculations brought an unforeseen achievement in certain territories, they neglected to give a perpetual and reliable answer for extortion detection. A comparable examination space was introduced by Wen-Tooth YU and Na Wang where they utilized Anomaly mining, Exception discovery mining and Distance entirety calculations to precisely anticipate deceitful exchange in an imitating analysis of credit card transaction informational collection of one certain business bank. Exception mining is a field of information mining which is essentially utilized in money related and web fields. It manages distinguishing objects that are isolates from the fundamental framework for example the exchanges that aren't real. They have taken ascribes of client's conduct and dependent on the estimation of those credits they've determined that distance between the noticed estimation of that trait and its foreordained worth. Unusual strategies, for example, mixture information mining/complex organization arrangement calculation can see illicit occurrences in a genuine card exchange informational index, in view of organization recreation calculation that permits making portrayals of the deviation of one case from a reference bunch have demonstrated productive ordinarily on medium estimated online exchange. There have likewise been endeavors to advance from a totally new perspective. Endeavors have been made to improve the ready input collaboration if there should be an occurrence of deceitful exchange. In the event of deceitful exchange, the approved framework would be cautioned and an input would be shipped off deny the progressing exchange. Fake Hereditary Calculation, one of the methodologies that shed new light in this space, countered extortion from an alternate heading. It demonstrated exact in discovering the deceitful exchanges and limiting the quantity of bogus alarms. Despite the fact that, it was joined by characterization issue with variable misclassification costs.[4]

In This Writing Audit characterize the all thing whatever face by the writer in doing explore .here likewise characterize as the information about MasterCard extortion discovery.

6. Future work:

Nowadays, in the global computing environment, online payments are important, because online payments use only the credential information from the credit card to fulfill an application and then deduct money. Due to this reason, it is important to find the best solution to detect the maximum number of frauds in online systems. Here in this using different machine learning classification we can try to pureify the credit card data,use it to find out the fraud.

7. Summary / Conclusion:

The evaluations conducted using two datasets, where, the first dataset was a dummy dataset that represented the characteristics of credit card data and a newly transformed dataset using data normalization and Principal Component Analysis techniques.

Decision trees, Logistic Regression and Neural Network algorithms were used in developing four fraud detection models to classify a transaction as fraudulent or legitimate. Three metrics were used in evaluating their performances. The results showed that there is no data mining technique that is universally better than others. Performance improvement could be achieved through developing a fraud detection model using a combination of different data mining techniques (ensemble).

8. References:

1. Jain R., Gour B., Dubey S., A hybrid approach for credit card fraud detection using rough set and decision tree technique,International Journal of Computer Applications 139(10) (2016)
2. Friedman, N., Geiger, D. and Goldsmith, M. Bayesian network classifiers. Machinelearning,Vol. 29, pp 131-163. Kluwer Academic Publishers, Boston.
3. Jon TS Quah and M Sriganesh. Real-time credit card fraud detection using computational
4. G. Liu ,W.luan,Z ,li, and Y Zang,a new FDS for credit card fraud detection based on behaviors certificates ,2018.
5. Hamzah Ali Shukur et al, International Journal of Computer Science and Mobile Computing, Vol.8 Issue.3, March- 2019, pg. 257-260 s
6. “Statistics for General and On-Line Card Fraud,” http://www.epaynews.com/statistics/fraud.html, Mar. 2007.
7. Bolton, R, and D Hand. "Unsupervised Profiling Methods for Fraud Detection." Credit Scoring and Credit Control VII, 2001.
8. Maes, S., Tuyls, K., Vanschoenwinkel, B., &Manderick, B. (2002). Credit Card Fraud Detection Using Bayesian and Neural Networks. Brussels, Belgium.

9. Zaki, M., &Meira, W. (2014). Data Mining and Analysis: Fundamental Concepts and Algorithms. New York City, New York: Cam-bridge University Press. https://doi.org/10.1017/CBO9780511810114.
10. Sahin, Y., &Duman, E. (2011). Detecting Credit Card Fraud by Decision Trees and Support Vector Machines. Hong Kong, China: The International MultiConference of Engineers and Computer Scientist
11. Huang, S. (2013). Fraud Detection Model by Using Support Vector Machine Techniques. Chiayi, Taiwan: International Journal of Digital Content Technology & its Applications.
12. Balaji, G. N., T. S. Sabatini, and N. Chidambaram. "Detection of heart muscle damage from automated analysis of echocardiogram video." *IETE Journal of Research 61.3 (2015): pp: 236-243. https://doi.org/10.1080/03772063.2015.1009403*.