Journal of Engineering and Applied Sciences
MEDWELL PUBLICATIONS

Fraud Detection on Credit Cards using Artificial Intelligence Methods

[1]Trishant Baid and [2]K.Priyadarsini
[1]*Department of Computer Science and Engineering*
[1,2]*Vels Institute of Science Technology and Advanced Studies, Chennai, India*

Key words: Credit card fraud, artificial intelligence, data science, algorithm, machine learning

Corresponding Author:
Trishant Baid
Vels Institute of Science Technology and Advanced Studies, Chennai, India

Page No.: 232-236
Volume: 16, Issue 07, 2021
ISSN: 1816-949x
Journal of Engineering and Applied Sciences

Abstract: It's essential that master or credit cards organizations can distinguish false Visa exchanges with the goal that clients don't need to pay for goods that they didn't buy. These issues are to be handled using Data Science, alongside Machine Learning can't be exaggerated. This venture plans to represent the displaying of an informational collection utilizing AI with Credit Card Fraud Detection. The imbalanced dataset issue happens in light of the fact that the quantity of real exchanges is a lot higher than the false ones though applying the correct component designing is significant as the highlights got from the ventures are restricted and applying highlight building strategies and changing the dataset is pivotal. Additionally, adjusting the recognition framework to continuous situations is a test since the quantity of charge card exchanges in a restricted timespan is exceptionally high. Likewise, we will examine how assessment measurements and AI techniques separate among each examination.

INTRODUCTION

The quantity of cashless exchanges is at its pinnacle point since the start of the advanced period and it is well on the way to increment later on. While that is a favourable position and gives convenience to clients, it additionally makes open doors for fraudsters. Just in 2016, 34,260.6 million exchanges have been performed, making an aggregate of 66,089 exchanges for every second. The overall deficit of the worldwide economy out of false exchanges is $2.17 billion. As the misfortune is very major, there is various exploration to diminish the causalities made with charge card misrepresentation. The quantity of examination papers in the fund application zone utilizing AI spans to thousands. While some of them attempt to fathom this utilizing scientific guideline based calculations as of late, AI and man-made reasoning methods are sought after. That is the aftereffect of the huge information gathered from billions of exchanges and this information by one way or another could be helpful in attempting to foresee whether a next, obscure exchange is really a misrepresentation or not[1, 2].

This issue is especially testing from the point of view of learning as it is portrayed using different views, for example, lopsidedness. Quantity of substantial exchanges far dwarf deceitful ones. Likewise, the exchange designs regularly alter their factual substancialities through the period. This aren't by any means the sole difficulties for usage of a genuine extortion location framework, be that as it may. In genuine world models, the huge stream of installment demands is rapidly filtered via. programmed apparatuses that figure out which exchanges to approve. AI calculations are utilized to break down all the approved exchanges and report the dubious ones.

Recognizing a charge card misrepresentation is in reality a paired arrangement issue, where the result is either bogus or valid. That grouping issue could be settled utilizing three AI errands: directed learning, solo learning and semi-administered learning. An administered learning approach uses past, referred to exchanges that are named as false or real. The previous information is prepared and the made model is utilized to foresee whether another exchange is misrepresentation or not. Solo methods are the ones that don't utilize the marked information, yet utilize unlabeled information to describe the information dispersion of exchanges. With that way, the anomaly information could be worthy as the deceitful exchanges. Grouping and pressure calculations are utilized to take care of solo issues. At the point when the two methodologies expressed are consolidated, it brings out semi-directed calculations. These calculations are commonly utilized when there is less named information in the dataset. The properties of any great misrepresentation recognition framework ought to be[3].

- It ought to have the option to recognize the fakes precisely that implies the quantity of wrong characterizations ought to be least
- It ought to have the option to identify the extortion

Through our research paper we have tried to form a scenario of all the credit card fraud detection and hence we've tried to put in the arrangements of all the methods that we can apply upon this. We, as the authors have made certain commitments about the findings of the paper. They are:

- Writing survey of theresearch on extortion identification frameworks
- Synopsis and arrangement of all the current strategies in extortion identification
- Parametric correlation of all the currentstrategies and proposed model

Literature review: Misrepresentation go accordingly the double dealing planned that brings the budgetary or orsolefully the person to person benefits. It's a very illegal and unlawful act and even literary sources claim it to be very unlawful and hence the immediate effects on the articles and many other sources[4].

Different written databases that relate to abnormality and even we can say misrepresentation identification of the space has been a very amount of distributed and hence kept as open source purposes, so that, can be accessed by anyone in and around the world.. A far reaching study in direction of Clifton Phua with his teammates uncovered that strategies utilized in this area incorporate information mining applications, computerized misrepresentation identification, ill-disposed location. In spite of the fact that the new generation techniques and the calculations brought an unforeseen accomplishment in a few zones, they neglected to give a perpetual and steady answer for misrepresentation identification. Unpredictable procedures, for example, half breed information mining/complex system grouping calculation can see unlawful occasions in a real card exchange informational collection in view of system remaking calculation that permits making portrayals where one case differs from the other case and hence have demonstrated the commonly used medium exchange[5].

MATERIALS AND METHODS

Types of fraud

Theft fraud: This kind of fraud is where the thief has the same copy of card as yours or the card is stolen. The thief makes many a times multiple attempts to make the card transaction successful. We have many a cases where our card is saved in many of the shopping or payment websites. The transaction then just requires the CVV or which is better known as Card Verification Value. If someone unauthorized gets access to any of the sites, he or she can easily attempt the CVV value multiple times and hence once or the other time the card access is stolen. In these cases the thing to do is that the owner of the card needs tp contact the bank immediately and get his or her card blocked in certain urgent measures.The card number should even be replaced and even the bank should try finding the IP Address of the transaction site and hence severe action should be taken against it. This kind of Visa Extortion is developing the danger towards the vendors who are seling their products online[6].

Fraud while application process: This happens when someone gives the application for the card with a bogus data. We need to carry out or recognise these kind of frauds and hence there should be a framework that can identify the mistreated applications. To distinguish application extortion, there must be different criterias that should be recognized. When the applicable forms start from the originating point from an equivalent individual and it contains much of similar kind of subtleties and in the other case when these forms for visa originate from many a same people with very same informations, the socalled personality fraudsters. Many banks across the world have a process of a full fledged application system. The data required incorporates recognizable proof data, area data, contact data, private data and extra data. Intermittent data accessible will be for many recognisable purposes that can include the personal identity of a person and hence the name, complete address, date of birth of the person. The candidate will have to illuminate the bank

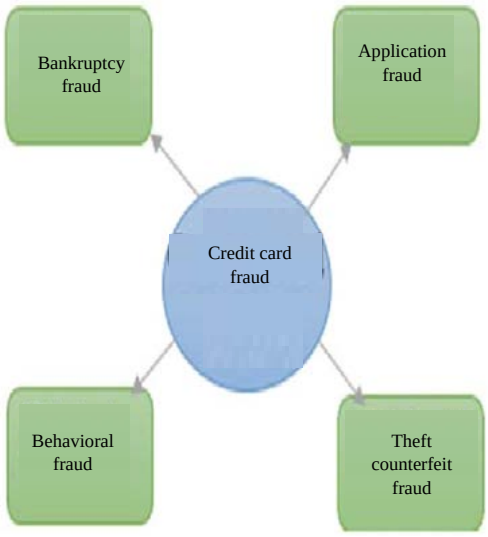

Fig. 1: Types of credit card fraud

Fig. 2: Fraud detection methods

regarding the area subtleties: the exact mapping of the address given, PIN Code and other details like the city and the country mentionary. The provider bank will also likewise request proximity subtleties, for example, email address, land-line and versatile telephone numbers. Private data will be the secret key. Furthermore, the sexual orientation will be given.All these information that is collected from the application of the applicant will be used for identification purposes and if any fraud that can also be considered accordingly by the bank for many checking criterias. To distinguish the alleged copies, cross-coordinating procedures are in like manner use[8].

Bankruptcy fraud: In this process, An exploration request is passed by the bank towards the hand of credit agency. The exploration incorporates many databases like the personal data which is also needed by the credit agency. As a result of which, the credit agency sends a data of all the information that is explored and the credit score of that particular individual is set to the credit agency as a credit report of individual specifics, subtleties of rebelliousness and all the binding promises that are done legally are done, data from open registries and extra positive data, for example, reimbursement of advances that will be done on the basis of contract or the legal paper which is signed on the day of development or before it.Few of these credit card agencies wven move for an in house verification for the address that is given by the applicant and hence they verify it for the sanity of the company in obscure cases. Data witt department information hence needs to be accumulated and is taken a wide range of sources. Banks, purchaser fund organizations, credit associations and assortment organizations are a portion of the elements that intermittently provide all the report to the credit card agency that is in control of the application. There are many other methods that is used for the development and hence the data is taken from the local and the state governments for further verification. Ordinarily, person money related organizations and all the various agencies in contact with the applicant report to the credit card agency about the updates of the data (Fig. 1 and 2)[9].

Existing techniques used for detection of frauds: There are many methods which were used towards this scenario of credit card fraud detection and hence, we have tried to list some of the existing methods and our research paper would resultify all the methods that are used with a accuracy and precision along with the false rate.

Decision tree method of detection: This method is a computational instrument of fraud detection using the criterias of arrangement and expectation. The involvement of a tree is a hub which involves testing on a particular trait in which the Branch tells us the particular result or the outcome of that trait and the leaf hub which is also known as terminal hub also signifies the class name. This has a tree kind of structure in which the node which is on the top is called the root node or the main node. The method goes as follows that it organises the criterias and passes it on to the subsequent leaf node for the application.

Fuzzy logic

- Step 1: You need to input 5 criterias that is time, amount, location, interval and frequency of credit card transaction
- Step 2: It will show an output of credit card classifys that are posted on the terms of linguistic
- Step 3: The inputs will have fuzzy variables
- Step 4: Membership functions are associated with each of the fuzzy variables. It is calculated for each of the fuzzy variables

- Step 4: Now finally the credit card classification is done bt the maximum amount of the selected output and hence the fraud transactions are determined

KNN algorithm: This is a technique that is without the parameters. Refering to which no assumptions were made and data is a still pure data. This algorithm is very simple and termed to be the easiest algorithm for classifying. The prime point around which the algorithm revolves around is whenever there is a new data, the neighbouring data is taken from the training data. We have taken particular variables as an example to explain this algorithm properly:

- Age- Age of the applicant
- Working Experience- The number of years he has been working
- Total Income- Monthly Income of the individual
- We calculate specificity, sensitivity and accuracy on the following data
- Sensitivity= TP/ (TP+FN)
- Accuracy= (TP+TN)/(TP+FP+TN+FN)
- Specificity= TN/(TN+FP)

TN stands for True Negative, TP stands for True Positive, FN stands for False Negative, FP stands for False Negative

Neural networks: These are Artificial Neural Networks (ANN) in which a toolbox is used so that, the results can be tested:

- Step 1: There was an application fitting that was selected for our work
- Step 2: Next step was the selection of the datasets. 100 samples were collected of which the elements chosen in input were 5 in number. And another 100 samples were collected of which 1 element was selected as the target
- Step 3: Now we randomly divide the datasets:
- Training-70%
- Testing-15%
- Validation-15%
- Step 4: There were neurons in the hidden layer and the number of neurons were 10
- Step 5: It goes to the step of testing in which the process was carried out for repeated times with different initial conditions[10]

RESULTS AND DISCUSSION

Support Vector Machines gives about 94% of accuracy with the precision of 85% and fallacy rate of 5%. Logistic Regression gives about 94.7% of accuracy with the precision of 78% and fallacy rate of 3%. Artificial Neural Networks gives about 99% of accuracy with the precision of 99% and fallacy rate of 0.1%. Decision Trees gives about 98% of accuracy with the precision of 98% and fallacy rate of 2%. Bayesian Network gives about 97% of accuracy with the precision of 97% and fallacy rate of 2.5%. KNN Method gives about 97% of accuracy with the precision of 96% and fallacy rate of 3%. Fuzzy Logic System gives about 95% of accuracy with the precision of 87% and fallacy rate of 1%.

The results we found from the above table are that the Artificial Neural Networks have the highest level of accuracy and Support VECTOR Machine having the least accuracy. Detection rate becomes high in case of Decision Trees, Artificial Neural Networks, K-Nearest neighbour and Bayesian Network. While it lowers in the case of Logistic Regression, Support Vector Machine and Fuzzy Based Logic System. Lower False Rate is only provided by Artificial Neural Networks and SVM having the highest False Rate. We on the basis of research have found these major gaps in the existing models that needs to be improved:

- We don't have proper datasets of the credit cards with us because it is the private property of the bank and it remains as a bond between the customer and bank
- Inaccessibility of a solitary amazing calculation which we can perform and can hence outrage every existing calculation which is possible
- There is an absence of the boundaries and hence the assessments cant take place properlyand hence cant depict the precision of the framework and lead to the superior results
- The framework fails in adjusting the viably evolving conditions, new fake strategies also, veritable changes made in buy propensities for a client[11]

CONCLUSION

The extortion is a proof of the criminal nature of people. The research paper has rattled off the most well-known strategies for extortion alongside their discovery techniques and investigated ongoing discoveries that are going on in this domain. This research paper has also shown how the methods of Artificial Intelligence and Machine Learning show signs of improvement brings about misrepresentation discovery alongside the calculation, pseudocode, clarification its execution and experimentation outcomes.

Despite of the fact that many of the very few misrepresentation location strategies accessible today however none can identify all cheats totally when they are really occurring, they for the most part recognize it after the misrepresentation has been submitted. This occurs since an extremely infinitesimal number of exchanges from the all out exchanges are really fake in nature. So we

need an innovation that can identify the fake exchange at the point when it is occurring with the goal that it tends to be halted at that point and there and that too in a base expense.

So, the significant undertaking of today is to manufacture an exact, exact and quick distinguishing misrepresentation identification framework for Mastercard cheats that can distinguish not just cheats occurring over the web like phishing and website cloning yet in addition messing with the Visa itself for example it signals an alert when the altered Visa is being utilized.. There are gaps in everything. To solve those we apply the method of the divide and rule. We combine both the strategies and hybrid way of defining the fraud. We combine different algorithms to enhance the performances and lead to a better outcome of fraud detection methods.

RECOMMENDATIONS

Further enhancement scopes: We are not able to reach the 100% objective with which we started the paper but there are many areas that needs to be explored and carried out many oppurtunities for further researchers to have a perfect interpretation of the credit card frauds that happen in the current fast moving world. There are many outcomes that have been brought out through this research paper and hence we have tried to find the best ideologies. The further ideas include:

- Finding a hybrid means of enhancements that can lead us to a 100% accurate methods
- If the size of dataset is expanded, the calculation data and accuracy should remain the same as earlier
- There can be a new algorithm that can be found and can replace the existing ones and hence lead to a proper methods

REFERENCES

01. Zanin, M., M. Romance, S. Moral and R. Criado, 2018. Credit card fraud detection through parenclitic network analysis. Complexity, Vol. 2018, 10.1155/2018/5764370.
02. Dal Pozzolo, A., G. Boracchi, O. Caelen, C. Alippi and G. Bontempi, 2017. Credit card fraud detection: A realistic modeling and a novel learning strategy. IEEE. Trans. Neural Networks Learn. Syst., 29: 3784-3797.
03. Delamaire, L., Abdou, H.A.H. and J. Pointon, 2009. Credit card fraud and detection techniques: A review. Banks Bank Syst., 4: 57-68.
04. Yu, W.F. and N. Wang, 2009. Research on credit card fraud detection model based on distance sum. Proceedings of the International Joint Conference on Artificial Intelligence, April 25-26, 2009, Hainan Island, pp: 353-356.
05. Jain, Y., N. Tiwari, S. Dubey and S. Jain, 2019. A comparative analysis of various credit card fraud detection techniques. Int. J. Recent Technol. Eng., 7: 402-407.
06. Kho, J.R.D. and L.A. Vea, 2017. Credit card fraud detection based on transaction behavior. Proceedings of the Tencon 2017-2017 IEEE Region 10 Conference, November 5-8, 2017, IEEE, Penang, Malaysia, pp: 1880-1884.
07. Weston, D.J., D.J. Hand, N.M. Adams, C. Whitrow and P. Juszczak, 2008. Plastic card fraud detection using peer group analysis. Adv. Data Anal. Classif., 2: 45-62.
08. Trivedi, I., Monika and M. Mridushi, 2016. Credit card fraud detection. Int. J. Adv. Res. Comput. Commun. Eng., 5: 39-42.
09. Maniraj, S., A. Saini, S. Ahmed and S. Sarkar, 2019. Credit card fraud detection using machine learning and data science. Int. J. Eng. Res. Technol., 8: 110-115.
10. Phua, C., V. Lee, K. Smith and R. Gayler, 2014. A comprehensive survey of data mining-based fraud detection research. Comput. Eng. Finance Sci., 1: 1-14.
11. Sonepat, H.C.E. and M. Bansal, 2014. Survey paper on credit card fraud detection. Int. J. Adv. Res. Comput. Eng. Technol., 3: 287-232.