

E-PAYMENT SCAM DETECTION SYSTEM

Jay Mahesh Acharya*¹, Shreshth Shetty*², Ajitkumar Khachane*³

*^{1,2}Student, Dept. Of Engineering, Vidyalkar Institute Of Technology, Mumbai, India.

*³Professor, Dept. Of Engineering, Vidyalkar Institute Of Technology, Mumbai, India.

ABSTRACT

It is vital that credit card companies can identify fraudulent credit card transactions so that customers are not charged for items that they did not purchase. Such issues can be overseen with Data Science furthermore, its significance, alongside Machine Learning, could not be more significant. This venture plans to outline the displaying of an informational collection utilizing AI with Credit Card Fraud Detection. The Credit Card Extortion Detection Problem incorporates displaying past charge card exchanges with the information of the ones that ended up being extortion. This model is then used to perceive whether another exchange is false or not. Our target here is to identify 100% of the deceitful exchanges while limiting the erroneous extortion orders. Mastercard Fraud Detection is an average example of order. In this interaction, we have zeroed in on dissecting and pre-handling informational collections just as the arrangement of numerous inconsistency identification calculations, for example, Local Outlier Factor and Isolation Forest calculation on the PCA changed Credit Card Transaction information.

Keywords: Books, Recommendation System, Machine Learning, Product Filters, AI.

I. INTRODUCTION

'Fraud' in Visa exchanges is unapproved and undesirable utilization of a record by somebody other than the proprietor of that record. Fundamental anticipation measures can be taken to stop this maltreatment and the conduct of such deceitful practices can be examined to limit it and secure against comparable events in the future. In different words, Credit Card Fraud can be characterized as a situation where an individual utilizes another person's Visa for individual reasons while the proprietor and the card giving specialists are uninformed of the way that the card is being utilized. Misrepresentation identification includes observing the exercises of populaces of clients to assess, see or stay away from frightful conduct, which comprise of misrepresentation, interruption, and defaulting. This is an exceptionally prominent issue that requests the consideration of networks, for example, AI what is more, information science where the answer for this issue can be mechanized. This issue is especially difficult according to the viewpoint of learning, as it is portrayed by varied factors, for example, class lopsidedness. The quantity of legitimate exchanges far dwarf false ones. Additionally, the exchange designs continually change their measurable properties throughout the process of everything working out.

II. METHODOLOGY

The methodology that this paper proposes, utilizes the most recent AI calculations to recognize peculiar exercises, called exceptions. Inside this dataset, there are thirty-one sections out of which twenty-eight are named as v1-v28 to secure delicate data. The different segments address Time, Amount and Class. Time shows the delay between the main exchange and the accompanying one. Sum is the measure of cash executed. Class 0 addresses a substantial exchange, and one addresses a fake one.

Software Used:

- Python
- Machine learning algorithm
- SCIKIT
- GOOGLE COLLAB

III. MODELING AND ANALYSIS

The waterfall model is a sequential design process, often used in software development processes, in which progress is seen as flowing steadily downwards (like a waterfall) through the phases of Conception, Initiation, Analysis, Design, Construction, Testing and Maintenance.

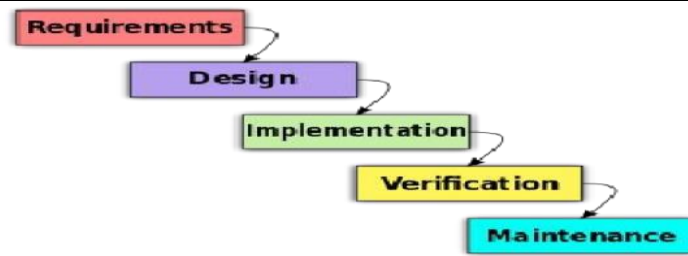


Figure 1: Flowchart

1. It allows for departmentalization and control.
2. It can be set with deadlines for each stage of development.
3. All the requirements are documented beforehand.
4. The waterfall model progresses through easily and explainable phases and thus it is easy to use.
5. It is easy to manage due to the rigidity of the model – each phase has specific deliverables and a review process.
6. In this model, phases are processed and completed one at a time and they do not overlap.

A. Feasibility Study

This is an evaluation and analysis of the potential of the proposed project which is based on extensive investigation and research to support the process of decision making. It assesses the operational technical and economic merits of the proposed project.

1. Technical Feasibility

This assessment is focused on gaining an understanding of the present technical resources of the organization and their applicability to the expected needs of the proposed system. It is an evaluation of the hardware and software and how it meets the needs of the proposed system.

2. Economic Feasibility

This assessment aims to determine the positive economic benefits to the organization that the proposed system will provide. It typically involves a cost/ benefits analysis, and it is the most frequently used method for evaluating the effectiveness of a new proposed system.

3. Schedule Feasibility

It is the measure of how reasonable the project timetable is, or the deadline is reasonable or not. During the lack of time or the time become mandatory, we must finish the project within a given time period. It mainly addresses : Can the project really be completed in the given period of time.

IV. LITERATURE SURVEY

In paper [1] Credit Card Fraud Detection system Survey link

Assistant Professor, Department of Computer Science,

V.V. Vanniaperumal college Virudhunagar. 2Assistant Professor, Department of Computer Science, Government Arts College, Melur, Madurai.

In this paper they have used five classification methods were used to detect fraud in credit card approval process in the banking sector. This work demonstrates the advantages of applying the data mining techniques including decision trees (C5.0 & CART), SVM using SMO, Logistic Regression and BayesNet to reveal fraudulence in credit card approval process. It reduces the financial institution’s risk. Any financial institution needs to retain its customers. A legitimate customer must not be classified as Fraud, or a Fraud must not be classified as Legitimate.

In paper [2] Credit Card fraud detection system. survey link

In this paper they have discussed Fraud function as the unlawful or criminal deception intended to result in financial or personal benefit. It is a deliberate act that is against the law, rule, or policy with an aim to attain unauthorized financial benefit.

In paper [3] Credit Card fraud detection system URL <https://towardsdatascience.com/credit-card-fraud-detection-9bc8db79b956>

In this paper they, let us walk you through a Kaggle competition regarding credit card fraud detection. A deep neural network and two machine learning models will be built to tackle the challenge and compare different model performance. Additionally, data sampling techniques will be implemented to improve the model.

V. RESULTS AND DISCUSSION

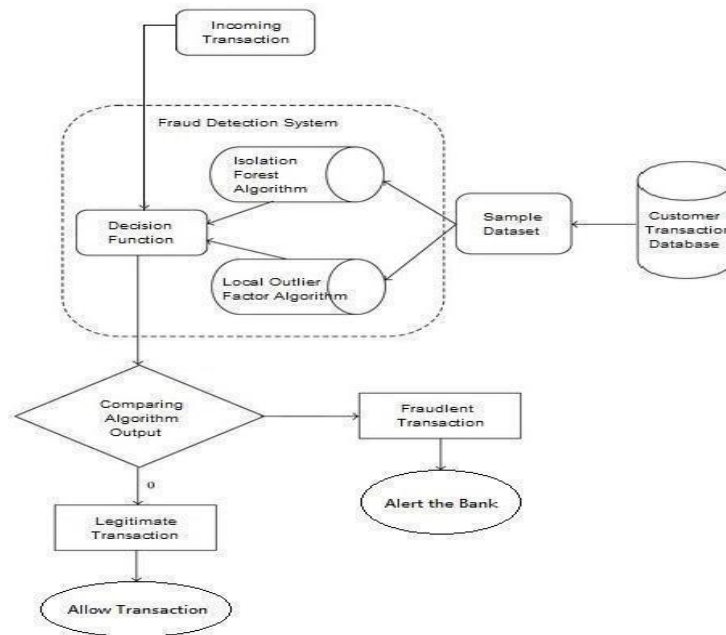


Figure 2: Flowchart

Algorithm to be used:

Logistic Regression and k nearest neighbor model and isolation forest algorithm. The comparison between the algorithm is a crucial factor so that we get a brief idea on result accuracy.

Data Visualization is a multi-component field that deals with graphic representation of data. It is a highly effective means of communication where data is plentiful for example a timeline.



Figure 3: Data Visualization

Transaction class distribution

We plot various diagrams to check for irregularities in the dataset and to outwardly grasp it. In the wake of checking this dataset, we plot a histogram for each section. This is done to get a graphical portrayal of the dataset which can be utilized to check that there are no missing any qualities in the dataset. This is done to guarantee that we do not need any missing worth ascription and the AI calculations can measure the dataset easily.

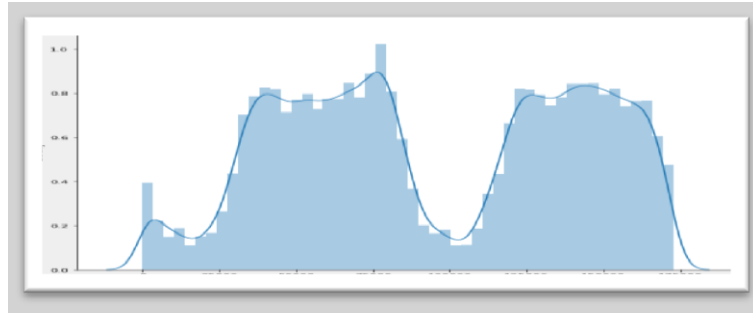


Figure 4: Transaction class distribution

The result with the complete dataset processed is also determined and is given in the results section of this paper. Around 284807 of transaction 284315 was found normal while 492 found fraud. Which is shown in this fig.4



Figure 5: Time of Transaction vs amount of class

VI. CONCLUSION

While we were unable to connect objective of 100% precision in misrepresentation identification, we wound up making a framework that can, with sufficient opportunity and information, get exceptionally near that objective. Likewise with any such venture, there is some opportunity to get better here. The actual idea of this undertaking takes into consideration numerous calculations to be incorporated together as modules and their outcomes can be consolidated to build the exactness of the outcome. This model can additionally be improved with the option of more calculations into it. Notwithstanding, the yield of these calculations should be in a similar arrangement as the others. When that condition is fulfilled, the modules are not difficult to add as done in the code. This gives an incredible level of measured quality and flexibility to the task. More space for development can be found in the dataset. As exhibited previously, the accuracy of the calculations increments when the size of dataset is expanded. Subsequently, more information will make the model more exact in identifying fakes and diminish the quantity of bogus up-sides. In any case, this needs official help from the actual banks.

VII. REFERENCES

- [1] S P Mani raj Assistant Professor (O.G.) Department of Computer Science and Engineering SRM Institute of Science and Technology. International Journal of Engineering Research & Technology (IJERT) ISSN: 2278- 0181Vol. 8 Issue 09, September- 2019
- [2] Credit Card Fraud Detection Based on Transaction Behavior -by John Richard D. Kho, Larry A. Veal" published by Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, 2017.
- [3] "Credit Card Fraud Detection through Parentic Network Analysis- By Massimiliano Zanin, Miguel Romance, Regino Criado, and SantiagoMoral" published by Hindawi Complexity Volume 2018, Article ID 5764370, 9 pages. <http://ieeexplore.ieee.org/Xplore>
- [4] "Credit Card Fraud Detection Using Bayesian and Neural Networks," Proceedings of Neuro Fuzzy, Havana, Cuba.