AN INNOVATIVE APPROACH FOR CREDIT CARD FRAUD DETECTION

Bushara Hamza[1], Santhi P[2], Dr. G Kiruthiga[3]

[1] *Student, Department of Computer Science & Engineering, IES College of Engineering, Thrissur, Kerala, India*
[2] *Assistant Professor, Department of Computer Science & Engineering, IES College of Engineering, Thrissur, Kerala, India*
[3] *Head of Department, Department of Computer Science & Engineering, IES College of Engineering, Thrissur, Kerala, India*

ABSTRACT

Due to the growing number of daily electronic transactions, which make credit cards more susceptible to fraud, they are the most popular form of electronic payment. Card fraud has cost credit card firms a lot of money. The most fre quent problem right now is the detection of credit card fraud. Companies that issue credit cards are searching for th e best systems & technology to reduce & detect credit card fraud. There are numerous ways to detect credit card fra ud. Financial fraud cases, such credit card fraud, have increased as a result of recent developments in e-commerce & e-payment systems. Implementing systems that can identify credit card theft is therefore essential. When using machine learning to detect credit card fraud, features of these frauds must be carefully chosen because t hey play a significant part in the process. For clients to avoid being charged for products they did not buy, credit ca rd issuers must be able to recognise fraudulent credit card transactions. With the use of credit card fraud detection, this project aims to demonstrate how to model a data set using machine learning. Modeling previous credit card tra nsactions using information from those that turned out to be fraudulent is part of the Credit Card Fraud Detection P roblem. The validity of a new transaction is then determined using this approach. Here, finding fraudulent transacti ons while reducing erroneous fraud categories is our goal.

.

Keywords: *- Machine Learning, Credit Card, Fraud Detection.*

1. INTRODUCTION

Credit card fraud is the use of another person's credit card for personal gain without the knowledge of the cardholder or the organization responsible for issuing the card. Credit cards have been widely used for online buying as e-commerce has grown & expanded, which has resulted in a significant increase in credit card fraud. Identification of credit card scams is crucial in the digital age. Monitoring & analyzing user activity is necessary for fraud detection in order to predict, identify, or prevent bad behavior. In recent times, a number of technologies & algorithms have been developed to detect credit card fraud. Recently, machine learning algorithms have been frequently employed to categorize transactions as fraudulent or not. These algorithms make use of datasets that include both labelled & unlabeled transactions. They categorize each transaction after analyzing the dataset. Monitoring user populations' activity in order to estimate, detect, or avoid unwanted behavior, such as fraud, intrusion, & defaulting, is called fraud detection. Algorithms for machine learning are used to examine all permitted transactions & flag any that seem suspect. Professionals look into these reports & get in touch with the cardholders to confirm whether the transaction was legitimate or fraudulent. The automated system uses the feedback from the investigators to train & update the algorithm, thereby enhancing the performance of fraud detection over time.

1.1 Motivation & background

Online purchasing is the most convenient way for clients to pay their bills these days, & credit cards are the most popular form of payment. The first issue that needs to be avoided is the potential of a credit card transaction including fraud. Companies save money by not owning physical storefronts & by avoiding high renting costs, while consumers save time by not having to travel to the store to make their purchases. It appears that the digital age introduced several incredibly valuable features that altered how businesses & customers connect with one another, but at a price. In order to effectively prevent these dangers, there are numerous data mining approaches accessible. In most cases, credit card theft occurs when an unauthorised individual steals the card & the fraudster ex ploits the card's information for his own gain. Making a system that can identify the false user would be a solution to this issue. Companies must use professional software engineers & penetration testers to ensure that all transactions are lawful & honest in order to prevent computational complexity & to provide improved accuracy in fraud detection. They are creating the servers for the company in such a way that the client has no control over crucial transactional elements like the payment amount. Most issues can be solved with proper design, however even the framework that was used to build the server is not flawless.

Using a credit card to make a transaction required manually processing it via a slide machine in the early 1970s, which left an imprint of the credit card number on a multi-part receipt. The customer received a carbon copy of the original copy that was intended for the merchant. The majority of credit card sales are now performed electronically through the phone, computer, or internet, with the information being processed in a matter of seconds thanks to technological advancements. Credit cards have been misused since the era of manual machines up until the advent of current electronic processors.

1.2 Problem Statement

Modelling previous credit card transactions while taking into account the ones that turned out to be fraudulent is part of the Credit Card Fraud Detection Problem. Then, a new transaction is evaluated using this model to determine whether it is fraudulent or not. This project seeks to minimise inaccurate fraud classifications while detecting 100% of fraudulent transactions. The credit card fraud detection system was developed to identify fraudulent activity & can tell if a fraudulent person is attempting to use the card.

2. RELATED WORK

Fraud act is 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. Numerous literatures pertaining to anomaly or fraud detection in this domain have been published already & are available for public usage. A comprehensive survey conducted by Clifton Phua & his associates have revealed that techniques employed in this domain include data mining applications, automated fraud detection, adversarial detection. In another paper, Suman, Research Scholar, GJUS&T at Hisar HCE presented techniques like Supervised & Unsupervised Learning for credit card fraud detection. Even though some of these techniques & algorithms achieved unexpected success, they were unable to offer a reliable, long-lasting answer to fraud detection. Wen-Fang YU & Na Wang presented a similar area of research in which they employed outlier mining, outlier detection mining, & distance sum algorithms to precisely forecast fraudulent transactions in an experiment simulating credit card transaction data from a particular commercial bank. Data mining's field of outlier mining is primarily utilised in the financial & internet sectors. It focuses on identifying detached objects from the main system, or transactions that aren't real. The distance between the observed value of an attribute & its predetermined value was calculated using customer behaviour attributes & their respective values. Unusual methods like hybrid data mining/complex network classification algorithms, which are based on network reconstruction algorithm & allow creating representations of the deviation of one instance from a reference group, have typically proven effective on medium-sized online transactions. These methods are able to detect illegal instances in a real card transaction data set. There have also been initiatives to advance from an entirely new perspective. The alert feedback interaction in the event of a fraudulent transaction has been improved. If a transaction is fraudulent, the authorised system will be notified & feedback will be issued to block the current transaction. One method that provided new insight into this area dealt with fraud in a different way: Artificial Genetic Algorithm. It worked well at identifying fraudulent transactions and reducing the incidence of false alarms. Even so, there was a categorization issue with fluctuating misclassification costs.

3. EXISTING SYSTEM

A credit card fraud detection system named Dempster-Shafer theory & Bayesian Learning based on the integration of 3 approaches, namely rule-based filtering, Dempster-Shafer theory, & Bayesian learning. Dempster's rule is used to combine & associate multiple pieces of evidence from the rule-based component in order to compute the initial belief about each incoming transaction.

BLAST & SSAHA hybridization are both highly efficient sequence alignment algorithms for examining customer's pending habits. This is a 2-stage sequence alignment algorithm in which a profile analyser (PA) compares & determines the similarity of an incoming sequence of credit card transactions with the genuine cardholder's previous spending sequences.

The Fuzzy Darwinian system is designed with the specific goal of detecting insurance fraud, which includes the difficult task of categorising data into 2 categories: safe & suspicious.

3.1 Drawbacks of Existing Systems

Existing systems produced less precise results & didn't return more relevant results. False declines or false positives occur when a system incorrectly flags a legitimate transaction as suspicious & cancels it. There is a reduced ability to recognise new patterns & adapt to changes in ancient systems. Manual work is also required for additional verification.

4. PROPOSED SYSTEM

When contrasted to the customer's prior purchases, card transactions are always foreign. In the real world, this unfamiliarity is a really challenging issue. In line with our suggested methodology, we are developing a system that can determine whether the person using the credit card is actually the real user or a fraudster posing as the real user. To determine whether a user is genuine or not, we have a machine learning model.

4.1 Advantage of Proposed System

The suggested solution enables users to use their credit cards securely & recognises when a false user is attempting to use that user's credit card.

5. METHODOLOGY

Methodology mainly includes data preparation, data analysis, data modelling & testing.

The transaction log file & the fraud cases make up the dataset's 2 different sources. The former includes all instances of credit fraud that have been reported, whereas the latter includes all transactions that a bank has accumulated. By comparing the transaction logs & the documented fraud cases, data annotation will start. The record in the transaction log will have a class value of "real" when a match is found, & "fraud" otherwise.

Transaction log data is subject to several pre-processing operations such as data-sanitation, normalization, binning, & handling null values. A feature selection technique is carried out to assess the significance of each property of the transaction log file prior to the modelling & testing phase. Additionally, this will eliminate the problem of dimensionality, which is a prevalent problem when processing large dimensions of data.

6. CONCLUSIONS

Unquestionably, using a credit card fraudulently is a criminal act of dishonesty. How machine learning can be used to improve fraud detection results has been thoroughly addressed in this study. The program will only get more effective over time as more data is fed into it because it is built on machine learning methods. Additionally, it's crucial to remember that you shouldn't give anyone, including friends, your credit card information since if you give them the right information, they can use it to access your account.

Although we fell short of our objective of 100% accuracy in fraud detection, we did manage to develop a system that, given enough time & data, can come very near to that objective. There is some potential for improvement here, as with any effort of this nature. Due to the nature of the project, it is possible to integrate many algorithms as modules & combine their outputs to improve the final result's accuracy. By including new algorithms, this model can be made even better. The output of these algorithms must, nevertheless, follow the same format as that of the others. The modules are simple to add as done in the code once that criterion is met. This gives the project a high degree of adaptability & versatility. The dataset contains more opportunities for development. When the dataset size is larger, the algorithms' precision rises. Consequently, more data will undoubtedly improve the model's ability to identify frauds & decrease the number of false positives. However, the banks themselves must formally support this.

7. ACKNOWLEDGEMENT

The Department of Computer Science & Engineering at IES Engineering College in Thrissur, Kerala, India, has supported this research work as part of the Master Research Project for the MTech Program, & the author would like to thank all of the faculty there.

8. REFERENCES

[1]. "Credit Card Fraud Detection using Machine Learning and Data Science – by S P Maniraj, Aditya Saini, Swarna Deep Sarkar, Shadab Ahmed" Published by International Journal of Engineering Research & Technology (IJERT) http://www.ijert.org ISSN: 2278-0181 IJERTV8IS090031 Vol. 8 Issue 09, September-2019

[2]. "Credit Card Fraud Detection using Machine Learning and Deep Learning Techniques – by Mohammed Azhan, Shazli Meraj" Published by Proceedings of the Third International Conference on Intelligent Sustainable Systems [ICISS 2020] IEEE Xplore Part Number: CFP20M19-ART; ISBN: 978-1-7281-7089-3

[3]. "A Survey Paper on Credit Card Fraud Detection Techniques – by Aisha Mohammad Fayyomi, Derar Eleyan, Amina Eleyan" Published by International Journal of Scientific & Technology Research Volume 10, Issue 09, September 2021 ISSN 2277-8616

[4]. "A Machine Learning Based Credit Card Fraud Detection Using the GA Algorithm for Feature Selection – by Emmanuel Ileberi, Yanxia Sun, Zenghui Wang" Published by Ileberi et al. Journal of Big Data (2022) https://doi.org/10.1186/s40537-022-00573-8

[5]. "Credit Card Fraud Detection Based on Transaction Behaviour – by John Richard D. Kho, Larry A. Vea" Published by Proc. of the IEEE Region 10 Conference (TENCON), Malaysia.

[6]. "A Comprehensive Survey of Data Mining-based Fraud Detection Research – by Clifton Phua, Vincent Lee, Kate Smith, Ross Gayler"