Credit Card Fraud Detection

Machine Learning Project- Phase 1

Malak Gaballa
900201683
Data Science
The American University in Cairo
malakhgaballa@aucegypt.edu

Masa Tantawy
900201312
Data Science
The American University in Cairo
masatantawy@aucegypt.edu

Introduction

Living in the 21st century, technology is an essential part of our daily lives and plays a vital role in almost all financial procedures. With the ongoing concerns regarding sustainability and the environment along with the rise in fintech institutions, several communities are starting to adopt a cashless and contactless lifestyle allowing individuals to undergo financial transactions without carrying money, but instead using digital options such as cards or mobile wallets. However, with the rising number of transactions that occur due to the very large mass of users and their spread, misuse appears. Similar to traditional financial transactions, frauds are a critical concern when using these technologies; criminals are afforded more opportunities to take money that does not belong to them from users hence causing multiple drawbacks to the parties involved such as the customer and the bank. These losses, which may be financial, reputational, or otherwise, may sometimes be very severe if the transaction was not identified as a fraud and stopped at an appropriate time.

One of the most common and prevalent forms of financial technologies is credit cards as almost every individual above the age of 18 nowadays holds at least 1 credit or debit card. In fact, many countries are creating channels to allow those below this age to use one since banks are the main place to store money. With this in mind, this technology must be made very safe to use, meaning that the chance of fraudulent transactions must be minimized. There are two possible ways for that: first, trying to block a transaction that appears to be fraudulent before the money is transferred, which is known as *fraud prevention*; second, *fraud detection* is identifying successful fraud transactions that have occurred [1]. Conspicuously, technological advances have the capacity to operate on these massive numbers of transactions unlike humans and can be of higher accuracy. In other words, machine learning models can be utilized to implement an effective credit card fraud detection solution making them safer to use. Thus, a model will be designed to identify whether or not a future transaction, given its details, is fraudulent or legitimate.

1 Literature Review

In this section, machine learning techniques adopted by previous works to develop a credit card fraud detection system will be reviewed. The most popular are neural networks, including multi-layer neural networks and bayesian networks, decision trees, logistic regression, and K-nearest neighbour classifier. These approaches are compared to evaluate their performance. Other models found are: a fusion approach using Dempster–Shafer theory and Bayesian learning, BLAST-SSAHA hybridization, hidden Markov model, and Fuzzy Darwinian detection system.

The first technique is neural networks, which are a set of weighted connected nodes that function as a human brain. One type of neural networks is *multi-layer neural network* which consists of multiple layers of weighted nodes in between the input and output. During data training, the weights are varied to achieve minimization of the mean square error between actual and predicted values of a network. In one of the studies examined, a sigmoid function was implemented for the available nodes in the layers [2]. Another type is *Bayesian networks*, also known as belief networks, which create layers using artificial intelligence and data mining techniques through supervised learning [3]. These are able to deal with skewed data distributions such as in the case of fraudulent transactions which are always fewer than legitimate transactions. Despite the prevalence of neural classifiers in fraud

detection due to their ability to learn on unexplained data, they require naturally clustered data [2]. Also, they require excessive and efficient training on the data thus presenting a time and cost limitation.

One example of a multi-layer neural network is a pilot study implemented at Mellon Bank that uses P-RCE , a member of radial basis function networks that have been developed for application to fraud recognition [4]. In the study, the output was a single cell that displays a numerical response known as *fraud score*, which is similar to credit scoring systems. Bearing in mind that perfect separability of fraudulent from non-fraudulent transactions is not possible due to the natural distribution of the procedure in real life, this model was found very useful since it was able to reduce total fraud losses from 40% to 20% by detecting fraudulent patterns on credit card accounts with less human attention. Yet, this is mainly, as it highly depends, on payment-related information as input features to the neural network. In other words, this model's performance depends on payment-related information.

Other frequently adopted machine learning algorithms for credit card fraud detection are decision trees, logistic regression and K-nearest neighbour (kNN). A decision tree is a tree shaped table with branches and leaves. At the end of a node, it can be connected to other nodes, known as a *branch node*, or unconnected, known as a *leaf node* [7]. Similarity trees were developed based on decision trees. These have high flexibility since they do not rely on a parameter from the data distribution and are easily implemented, displayed and understood. Nevertheless, decision trees are inefficient as they inspect each node one by one [5]. Similar to linear regression, logistic regression's binary output depends on a set of known features where it can predict whether a transaction is fraudulent or not. The regression coefficients of the model estimate odds ratios for each of the predictors [5]. As for kNN, this instance based learning model carries out classification based on similarity measures, such as Euclidean, Manhattan and Minkowski distances [6]. The majority class among the closest K points, hence must be an odd number, according to the similarity measure used, is returned.

The aforementioned techniques were assessed in several studies to design an efficient credit card fraud detection system. In multiple comparative studies mentioned in one of the papers [6], it was found that logistic regression had lower asymptotic error than Bayesian networks, Bayesian networks surpassed multi-layer neural networks in credit card fraud detection, and neural networks and logistic regression outperformed decision trees. On the other hand, a group of researchers tested logistic regression, neural networks and decision tree models for their applicability in fraud detections and concluded that neural networks had the best performance, followed by logistic regression [5]. However, consistent with the findings above, they found that these techniques outperformed decision trees. The datasets they used for training and testing were real transactions from 2005 and 2006 respectively. The usefulness of each technique was determined based on *lift*, a metric that assesses classification models. Another experiment was conducted using 3 different classifiers, Bayesian networks, logistic regression and kNN, on different samples of a dataset [6]. The metrics used were *accuracy, sensitivity, specificity, precision, Matthews correlation coefficient (MCC) and balanced classification rate (BCR)* which were computed using the true positive, false positive, true negative, and false negative rates for each model. For the kNN classifier, the Euclidean distance measure was used and the value of k=3 showed optimal performance after parameter tuning. After evaluating their performance, the kNN classifier was the best in all the metrics used since its false positive rate was 0 while logistic regression was the worst.

Additionally, other classifiers for credit card fraud detection were found, yet they were not popular and were more complicated [3]. A fusion approach using Dempster–Shafer theory and Bayesian learning combines information from current and past behaviour with a preset initial belief. Depending on it, the transaction is classified as fraudulent or not; then, the initial belief is either strengthened or weakened using Bayesian learning. BLAST-SSAHA hybridization classifies a transaction through two steps each consisting of an analyzer. First, a *profile analyzer* (PA) compares the incoming transaction with the card holder's spending history. If a transaction appears unusual, a *deviation analyzer* (DA) matches it to previously known fraudulent transactions to determine a decision. A hidden Markov model is an advanced stochastic model that is trained on the card holder's normal behaviour. Incoming transactions submitted for verification are classified based on a detailed analysis of each user's log. Lastly, the Fuzzy Darwinian detection system divides the data into levels of suspicion: low, medium and high using a set of guides from previous transactions. To confirm the decision, non-

overdue payments are considered genuine while those overdue for over three months are categorized as suspicious thus detecting stolen credit card frauds.

Based on studying previous researchers' works, there have been discrepancies on which machine learning algorithm is the best to develop a credit card fraud detection system. Logistic regression and kNN classifiers appear to be among the most efficient techniques due to their effectiveness, with uncertainty about the performance of neural networks. Contrarily, decision trees are inferior to the other methods of classification. Accordingly, we aim to develop a credit card fraud detection system using the best-performing algorithms previously stated. This system can be applied at financial institutions that issue cards, such as banks, to monitor credit card activity and hinder fraudulent transactions.

2 Datasets

In this project, the dataset chosen plays a critical role in the performance of the machine learning model. While searching for a dataset to use, one was found on Kaggle that contains actual transactions of European card holders over the course of 2 days during the month of September 2013, of which some are fraudulent and the rest are legitimate [7]. This dataset contains 284,807 transactions, of which 492 are fraudulent accounting for 0.172% of all transactions which despite being immensely unbalanced, mimics that real-world context. It contains 31 features including the time of the transaction and its amount as well as a binary feature labeling each instance as fraudulent (1) or legitimate (0). Although this dataset appears to be highly relevant to the topic and of an appropriate size, the remaining features are a result of principal component analysis transformation, which means that they are not the original features and that they are all numerical values. This presents a limitation, so this dataset can not be used.

Another dataset found on Kaggle contained the details of 94,682 credit card transactions that happened online, each labeled as fraudulent (FRAUD) or legitimate (LEGIT) [8]. Other details of this dataset such as when or where it was collected cannot be found. It contains 20 features of which 3 are categorical such as the masked domain name of the customer's email address and ZIP code, and other numeric ones like the hour features of the transactions and the total transaction amount. A major setback to this dataset is that 13 of these features are anonymized to maintain the privacy of the customers and their financial transactions, so the real details of the transactions are unknown. Hence, despite its availability, relevance and large size, this dataset will not be used.

One dataset found of mobile money transactions was generated using a simulator, *PaySim*, to be used by the public to overcome privacy issues that arise with real transactional data and the scarcity of data in this domain [9]. It was synthesized using financial logs extracted from actual transactions of a multinational company's mobile money service in an African country. It consists of 1,048,576 instances, representing a quarter of the original dataset, and 11 features. These include the amount of the transaction, the type of the transaction, customer and recipient IDs, account balances before and after the transaction, and a binary label to identify whether or not the transaction is fraudulent (1 or 0 respectively). Even though this data is very suitable for use and is available, the constraint to using this dataset is its small number of features, as they are less than 20, which will restrict the performance of the machine learning model making it difficult to predict whether a transaction is fraudulent or legit.

Finally, the dataset chosen, similar to the previous one, consists of simulated credit card transactions over the span of 2 years, from the 1st of January 2019 to the 31st of December 2020 [10]. It was found on Kaggle and was generated using a simulation tool in python, created by Brandon Harris, where transactions for 1000 customers with 800 merchants were created based on factual information. This dataset is divided into a training set of 1,296,677 instances and a testing set of 555,720 instances. For each transaction, there are 23 features such as the category and name of the merchant, credit card number, location details of the transaction, and its amount. The label is a binary feature that categorizes the transaction as fraudulent (1) or legitimate (0). Considering the size of this dataset and its features, it is excellent for use. A machine learning model to label an unknown transaction as fraud or not can be created on it.

References

[1]
Pozzolo, A.D. and Bontempi, G. (2015) *Adaptive Machine Learning for Credit Card Fraud Detection, Université Libre de Bruxelles*. Available at:

https://di.ulb.ac.be/map/adalpozz/pdf/Dalpozzolo2015PhD.pdf (Accessed: February 13, 2023).
[2]
Khyati Chaudhary, Jyoti Yadav, and Bhawna Mallick. 2012. A review of Fraud Detection Techniques: Credit Card . *International Journal of Computer Applications* 45, 1 (May 2012), 39-44.
[3]
S. Benson Edwin Raj and A. Annie Portia. 2011. Analysis on credit card fraud detection methods. *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)* (March 2011), 152-156. DOI: https://doi.org/10.1109/icccet.2011.5762457
[4]
Ghosh and Reilly. 1994. Credit card fraud detection with a neural-network. *Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences HICSS-94* (January 1994), 621-630. DOI: https://doi.org/10.1109/hicss.1994.323314
[5]
Aihua Shen, Rencheng Tong, and Yaochen Deng. 2007. Application of Classification Models on Credit Card Fraud Detection. *2007 International Conference on Service Systems and Service Management* (June 2007), 1-4. DOI: https://doi.org/10.1109/icsssm.2007.4280163
[6]
John O. Awoyemi, Adebayo O. Adetunmbi, and Samuel A. Oluwadare. 2017. Credit card fraud detection using machine learning techniques: A comparative analysis. *2017 International Conference on Computing Networking and Informatics (ICCNI)* (October 2017), 1-9. DOI:https://doi.org/10.1109/iccni.2017.8123782
[7]
Credit Card Fraud Detection. *Kaggle*. Retrieved February 13, 2023 from https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud
[8]
Online Credit Card Transactions. *Kaggle*. Retrieved February 13, 2023 from https://www.kaggle.com/adityakadiwal/credit-card-fraudulent-transactions
[9]
Synthetic Financial Datasets For Fraud Detection. *Kaggle*. Retrieved February 13, 2023 from https://www.kaggle.com/datasets/ealaxi/paysim1
[10]
Credit Card Transactions Fraud Detection Dataset. *Kaggle*. Retrieved February 13, 2022 from https://www.kaggle.com/datasets/kartik2112/fraud-detection?resource=download