Unveiling Fraud: Harnessing Predictive Data Mining for Credit Card Fraud Detection

Felipe Trinidad Pilier & Kieran Scanlon
Fordham University
CISC 5660 Data Sci. for Cybersecurity L01
Professor Weiss
May 8, 2024,

***Abstract*–**Credit card fraud poses a significant challenge to financial institutions and consumers worldwide, necessitating advanced detection mechanisms to safeguard against financial losses and maintain trust in the banking system. Credit card fraud not only leads to substantial financial losses but also undermines consumer confidence in digital transactions, impacting the overall integrity of the financial ecosystem. The sophistication of fraudulent schemes continually evolves, with perpetrators employing tactics such as identity theft, card skimming, and phishing scams to circumvent detection measures. Furthermore, the rise of digital transactions and the increasing interconnectedness of financial systems have expanded the avenues for fraudulent activities, necessitating constant innovation in fraud detection technologies to stay ahead of increasingly sophisticated cyber threats. As financial transactions become more complex and occur in real time, traditional rule-based fraud detection systems struggle to keep pace with the dynamic nature of fraudulent activities. Therefore, there is a pressing need to leverage advanced data mining techniques and machine learning algorithms to analyze vast amounts of transactional data and identify subtle patterns indicative of fraudulent behavior. This paper addresses this need by examining the efficacy of predictive data mining methods in enhancing the accuracy and efficiency of credit card fraud detection, offering insights into the potential of these technologies to bolster financial security in an increasingly digital world. However, before we begin, we must acquire a large enough dataset to perform our experiments adequately. We acquired our dataset from Kaggle.com, which contains over 280,000 different transactions with nearly 500 known fraudulent transactions. We attempted to tackle two significant issues within our dataset during the experiment. The first is the steep imbalance between entries of fraudulent and legitimate transactions, with fraudulent transactions being the minority class by a landslide. This simply means we must balance the dataset before we begin our tests. The second major issue is that most of the features are illegible due to a Principal component analysis (PCA) transformation. The PCA problem is the more significant of the two as twenty-eight out of thirty-one features have been changed due to the protection of sensitive information. Due to the protected nature of the information we are experimenting with, we can only go so far with the data available to the general public. Now, using the data, we can understand and use undersampling by reducing the instances of legitimate transactions and oversampling through the Synthetic Minority Over-sampling Technique (SMOTE) to address the class imbalance. Through experimentation, we evaluate the performance of various machine learning algorithms, including logistic regression, decision trees, k-nearest neighbor, and k-clustering. Our results demonstrate that balancing the dataset through SMOTE proved to yield much higher accuracy, precision, recall, and F1-score metrics. Thanks to our results, this study shows the efficacy of using advanced data mining techniques and machine learning algorithms, along with carefully balancing the dataset, in significantly increasing our credit card fraud detection capabilities.

I. BACKGROUND

Behind debit cards, credit cards are the most common method of financial transactions worldwide. According to Forbes Advisor, an organization centered around providing financial advice and insight, thirty-three percent of respondents with a bank account say they primarily pay for purchases with a credit card. Credit cards are payment cards issued by financial institutions that allow users to borrow funds from the provider in order to make purchases. When a credit card is used, the card owner essentially agrees with the bank to pay them back for the transaction. Thanks to its ability to be used in online transactions, credit and debit cards have become the two most used payment methods worldwide. Banks typically offer rewards for users who make purchases with their credit cards and uphold a good credit score according to their standards. These rewards include cashback, discounts, and gift cards, providing an incentive for their use. Additionally, credit cards offer the convenience of making payments from anywhere the cardholder is while delivering better security than cash, thanks to features like fraud protection and the ability to dispute unauthorized charges. Moreover, responsible use of credit cards can contribute to building a positive credit history, which is crucial for obtaining future loans, mortgages, and other financial products.

II. INTRODUCTION

As with any form of payment, there is a danger of the credit card holder's personal information being compromised by a threat. Credit card fraud is a danger that has become a worldwide concern for cardholders and banking institutions. It is a form of identity theft that comes from the unauthorized use of someone else's credit card. In the United States, using a stolen debit or credit card is illegal and punishable by up to ten years in prison, a ten thousand dollar fine, or a combination of the two.

Credit card fraud can happen in many different ways. The most common method of fraud is a card-not-present (CNP) scam. This type of compromise is generally self-explanatory. It is any unauthorized transaction that takes place without the physical credit card being used. CNP scams are performed by utilizing the three major components of a credit/debit card. The first component is the card number, commonly seen as a sixteen-digit number, but may vary depending on the card provider. This number uniquely identifies the card, as no other card will have the same sequence of numbers. The card's expiration date is the next component needed in a CPN scam. Credit cards expire every two to five years, depending on the bank that is issuing the card. Cards expire as a protective measure against fraud. When a card expires, it is no longer valid for use and will be declined if a transaction is attempted. A renewed card may also give the user an opportunity to upgrade it with new technology, like the contactless payment option currently being rolled out by various banks. The last component needed is the card verification value (CVV) number. It is a three-digit number commonly found on the backside of a credit or debit card. This is an extra layer of protection that the user must provide when purchasing online.

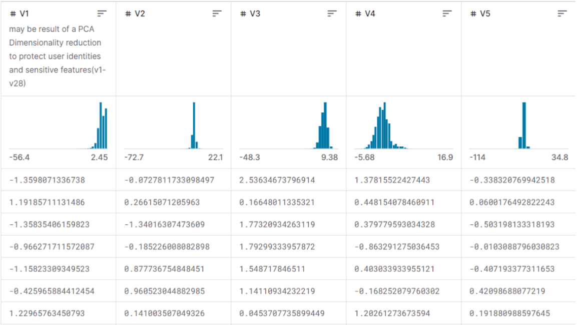
A card-not-present scam is focused on obtaining these three details that somebody can get through various means. Phishing is the most common method of obtaining the necessary information for a credit card scam. The phishing emails typically contain links to fake websites that closely resemble the legitimate ones. Recipients are asked to click on these links and provide sensitive information, such as credit card numbers, expiration dates, security codes, and personal identification details. Cybs capture the data inputted by the victim in real time. Vishing is another method of CNP. It is similar to phishing, requiring social engineering to trick victims into providing personal information. A threat actor uses a voice modification and confronts the target through a phone call. Because of the recent rise in artificial intelligence, vishing has become a more formidable threat. Generative AI is able to produce identical sounds based on a given sample. Take, for example, the story of Jennifer DeStefano's experience of being frightened by the similarity of generative AI and her daughter. On June 13, 2024, she received a call from her daughter and was convinced that she was kidnapped and held hostage. Pleading for her life, another person demanded DeStefano to wire fifty thousand dollars or else he would harm her daughter. According to DeStefano herself, "AI is revolutionizing and unraveling the very foundation of our social fabric by creating doubt and fear in what was once never questioned – the sound of a loved one's voice." Artificial intelligence has constantly pushed the limits of technology, especially recently with the release of ChatGPT 3. Since then, it has become a regular topic of conversation as we explore the unknown. In this exploration, AI's various uses in cybercrime have been highlighted as a cause for concern. Card-not-present scams may also occur through application fraud, where the cybercriminal opens a bank and false alarm rateset's name and makes unauthorized transactions. Account takeover occurs when a threat actor gains access to the cardholder's bank account login. Lastly, CNP scams can happen through skimming machines that look identical to official card scanners.

Despite the several layers of security banks implement to prevent fraud, cybercrime can never be foolproof. That introduces the topic of this paper: fraud detection using predictive data mining. The primary goal of this project is to detect fraudulent transactions accurately and efficiently. Financial institutions can take proactive measures to prevent fraudulent transactions and mitigate economic losses by identifying suspicious activities in real-time or near-real-time. Most cases of credit card fraud are uncovered after the threat actor has made their transaction. Fraud is inherently a data mining issue as it involves leveraging large volumes of transaction data to uncover patterns, anomalies, and trends indicative of fraudulent activity. By leveraging data mining techniques, banks can detect and prevent fraud more effectively, protecting consumers and businesses from financial losses.

III. EXPERIMENT METHODOLOGY

The dataset used for our experiments was sourced from Kaggle.com. It serves as a hub for data scientists, machine learning practitioners, researchers, and enthusiasts to collaborate and share knowledge. Kaggle offers various datasets across different domains, allowing users to explore, analyze, and build predictive models using real-world data. Additionally, it provides tools, resources, and educational materials to support learning and development in the field of data science and machine learning. The dataset we used is titled “Credit Card Fraud Detection.” The dataset comprises transactions by European cardholders using credit cards in September 2013. All of the entries were recorded within a two-day period. There are thirty-one features present. The first of note is time; this refers to the number of seconds elapsed between a given transaction and the next one in the sequence. The transaction amount is in euros. Class refers to whether the transaction is fraudulent or legitimate, labeled as zero or one. Lastly, V1 to V28 are features shown as ordinary numbers. These twenty-eight features pose one of two major issues when working with this dataset.

The first roadblock we faced with this dataset was regarding the features V1 to V28 previously mentioned. They contain only numerical input variables resulting from a principal component analysis (PCA) transformation. PCA is a dimensionality reduction method used to lower the number of irrelevant and trivial features. As more features become introduced in a given data set, the volume of space between data points increases, making them more sparse. Sparse data points are generally unwanted because they produce unwanted challenges and complications when analyzing them.



# V1	# V2	# V3	# V4	# V5
may be result of a PCA Dimensionality reduction to protect user identities and sensitive features(v1-v28)				
-56.4 2.45	-72.7 22.1	-48.3 9.38	-5.68 16.9	-114 34.8
-1.3598071336738	-0.0727811733098497	2.53634673796914	1.37815522427443	-0.338320769942518
1.19185711131486	0.26615071205963	0.16648011335321	0.448154078460911	0.0600176492822243
-1.35835406159823	-1.34016307473609	1.77320934263119	0.379779593034328	-0.503198133318193
-0.966271711572087	-0.185226008082898	1.79299333957872	-0.863291275036453	-0.0103088796030823
-1.15823309349523	0.877736754848451	1.548717846511	0.403033933955121	-0.407193377311653
-0.425965884412454	0.960523044882985	1.14110934232219	-0.168252079760302	0.42098688077219
1.22965763450793	0.141003507049326	0.0453707735899449	1.20261273673594	0.191880988597645

Fig 1. PCA Transformation

Fig. 1 visualizes the PCA transformation performed on the original dataset. Due to confidentiality issues, the original content and description of features V1, V2, …, V28 cannot be provided. However, we can reasonably assume that some of the features are based on the different factors recorded in a credit card transaction. These factors may include the following: the cardholder’s name, the merchant of the purchase, the frequency of which the same credit card was used, the location of the purchase, the time of the transaction, the device that was used, as well as the card number, the expiration date, and CVV. Due to the nature of the PCA transformation, the features in question are illegible. The only features which have not been transformed with PCA are “Time” and “Amount.”

The second major roadblock we faced with this dataset is that the classes are highly imbalanced. In total, two hundred-four thousand eight hundred and seven (284,807) points are labeled. Out of this number, only four hundred ninety-two (492) cases are fraudulent.

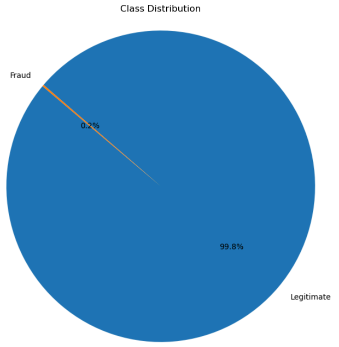


Fig. 2. Class Imbalance 492 Fraudulent to 284,315 Legitimate

Fig. 2 shows the class imbalance being incredibly steep. The data is highly unbalanced, with the positive class being fraud, making up 0.172% of all transactions. Implementing prediction methods at this point of the experiment would be unwise. When classes are imbalanced, the rare class may not be predicted very often and, in some cases, may not even be predicted at all. Class imbalance is commonly found in the detection of intrusion, malware, diseases, and credit card fraud. Even if a high accuracy metric is calculated, it may not be helpful since the minority class, fraud in this case, is never predicted. Because it is the most important class, it is imperative that the prediction method be able to pick up fraudulent cases every time they appear. Standard evaluation metrics like accuracy, recall, precision, and f-measure are not well suited for use within unbalanced datasets, especially in cases where cost is a factor. The cost of misclassifying a rare class is higher than the majority class being misidentified. Therefore, the best way to visualize the efficiency of the prediction method is to produce the area under the receiver operating curve (AUC-ROC). This is a graphical approach for displaying the trade-off between detection and false alarm rates. Ideally, our ROC curves will be as close to the left and upper border of the graph as possible. Ordinarily, a confusion matrix would be sufficient. However, in a situation like this, there is a high probability of false negatives. We aim to keep these false negatives as close to zero as possible. To solve this issue, we need to rebalance the class distribution.

There are two approaches we considered to circumvent the issue of class imbalance. The first is undersampling, which involves discarding data points from the majority class, being legitimate transactions, until they match the number of fraudulent points. We tested this out by undersampling the majority class and implementing the k-nearest neighbor algorithm. KNN is a simple yet powerful algorithm that memorizes the entire training dataset and makes predictions based on the similarity between new instances and existing data points. After undersampling the data, the current class distribution is four hundred ninety-two points for each class. The result is shown below in Fig. 3:

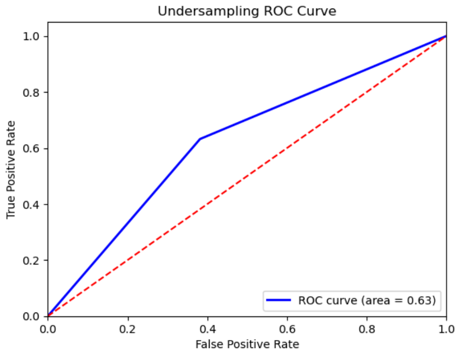


Fig. 3 Undersampling ROC using KNN

The result shown in Fig. 3 is very undesirable as the AUC-ROC is 0.63. This means that it may not be surprising for fraudulent examples to be detected as legitimate if the k-nearest neighbor is used. An AUC-ROC at 0.5 is essentially flipping a coin. We believe the metrics turned out like this because over two hundred thousand data points were discarded. That is a lot of precious data that is being thrown away. This prompted us to move on to the next experiment: oversampling.

Oversampling works in an opposite methodology to undersampling. The goal is to balance the classes by creating new instances of the minority class until the class distribution is even. Synthetic Minority Oversampling Technique (SMOTE) is a technique that is designed for such an issue. It works by synthesizing new examples of the minority class in the dataset, thereby balancing the class distribution. The SMOTE algorithm is an improved oversampling technique based on the random oversampling algorithm and is currently the most widely used sampling method. Utilizing the same method of testing as undersampling, we will save the oversampled dataset in a separate file and test k-nearest neighbor. Oversampling with SMOTE caused the creation of two hundred eighty-three thousand eight hundred twenty-three (283,823) new data points belonging to the fraud class. The result is shown below:

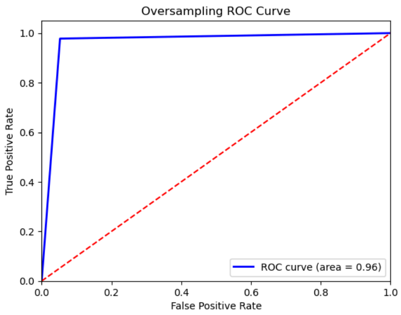


Fig. 4 Oversampling ROC using KNN

As demonstrated in Fig. 4, the ROC curve produces a much more desirable outcome. With this in mind, we will go forward with the SMOTE dataset, which we have titled “creditcard_oversampled_dataset.csv.”

Because of the previously mentioned PCA transformation that caused many features to become numerical inputs, it is impossible to explain the model for any of the different prediction algorithms. This is the limit of how much we are able to explain about the experimentation. In place of this, we will present the various algorithms used, explain what they are, and present the results printed. The first prediction method used was a decision tree. A decision tree is a supervised learning algorithm that is used for classification and regression. It is structured similarly to a flow chart where each internal node represents a decision based on a particular feature. After each condition, the tree splits and will continue to do so until a class is predicted. In this project, a tree was generated. We cannot explain the different criteria for splitting due to features V1 to V28 being illegible.

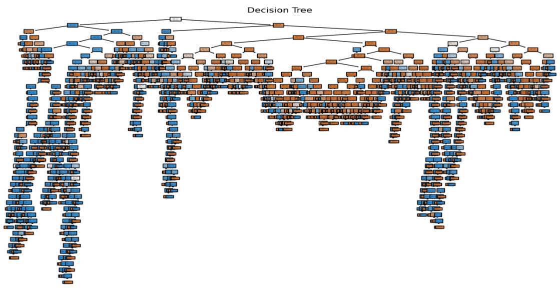


Fig. 5 Decision Tree

Despite the numerical values of features V1 to V28, the decision tree was able to run and produce predictions at the end of its respective branches. As shown in Fig. 5, the output is very complicated due to its use of every feature available. The tree may contain a lot of noise and be more accessible to interpret if fewer features were incorporated. Due to the nature of credit card transactions, every feature contains precious data; as such, it would be unwise to throw away points simply because we do not know how important they may be.

The next noteworthy experiment is with logistic regression. Logistic regression is a statistical method used for binary classification tasks. In the case of this project, it is perfect, considering the classes are labeled as the binary values 0 and 1. Logistic regression is preferred due to its simplicity, ease of interpretation, and efficiency, mainly when the association between the features and the target variable is roughly linear. The experiment is shown below in Fig. 6.

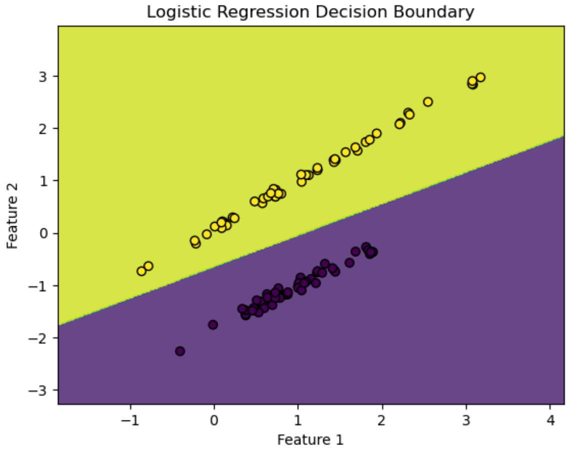


Fig. 6 Logistic Regression

IV. CONCLUSION

As technology continues to advance, new methods of cybercrime emerge. Thanks to the rise of artificial intelligence, vishing has become one of the most dangerous yet shockingly convincing social engineering tools today. As new risks of credit card fraud develop, we must also learn how to efficiently and accurately identify fraudulent activity in real time. The results of our project highlight how we can tackle one of the most significant issues with datasets concerning this topic: class imbalance. Because of SMOTE, newly generated samples of fraudulent transactions can be used and fed into predictive learning algorithms. As time progresses, we will acquire more data to learn the red flags seen in fraud more accurately. As data scientists, there is a limited amount of open-source datasets in which to practice machine learning techniques. The only way To have access to a more realistic dataset, one must work in collaboration with a bank. Under a non-disclosure agreement, they will likely provide the necessary data and more robust hardware to process the vast amount of points. The dataset acquired is the upper limit of what we can work with as researchers. Overall, confidentiality and safety remain the primary aspects that cybersecurity professionals aim to uphold. Credit card fraud detection is one major step toward clearing the blurry line between fraudulent and legitimate transactions.

REFERENCES

1. Woman falls prey to AI voice scam, loses Rs 1.4 lakh by mistaking caller for nephew; The woman transferred the money immediately into the man’s account only to realise that she had been scammed using AI voice scam. (2023, November 20). The Financial Express (New Delhi, India).
2. https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud
3. Ke Zhou, et al. “Improving SMOTE Technology for Credit Card Fraud Detection Category Imbalance Issues.” Engineering Letters, vol. 31, no. 4, Dec. 2023, pp. 1780–85. EBSCOhost,search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=173982005&site=eds-live.
4. W.-C. Lin, C.-F. Tsai, Y.-H. Hu, and J.-S. Jhang, “Clustering-based undersampling in class-imbalanced data,” Information Sciences, vol. 409, pp. 17–26, 2017.
5. Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.