Machine Learning for Credit Card Fraud Detection

Project

Ann Mary Jose

Alja Chinnu Thomas

Business Analytics – International (102)

St Lawrence College, Kingston

ADMN 5008: Applied Artificial Intelligence and Machine Learning

Professor Sujoy Paul

December 15, 2023

TABLE OF CONTENTS

1. Abstract

This project addresses the critical issue of credit card fraud, aiming to develop a machine-learning model capable of detecting fraudulent transactions with high accuracy. Credit card fraud presents a multibillion-dollar problem globally, and effective detection mechanisms are vital for financial institutions to protect their customers and reduce losses. The methodology encompasses data analysis, preprocessing of a highly imbalanced dataset, feature engineering, model training with Random Forest and XGBoost classifiers, and hyperparameter tuning. The main findings indicate that while the Random Forest model achieves high overall accuracy, it struggles with the identification of the minority class—fraudulent transactions—demonstrating the challenge of imbalanced class distribution. The study concludes with the importance of specialized techniques to enhance the model's performance on the minority class, stressing the potential of machine learning in fraud detection but also highlighting the necessity for continued research in this area.

2. Introduction

In the era of digital banking and e-commerce, credit card fraud has emerged as a major challenge for financial institutions and consumers alike. As transactions become increasingly virtual, the opportunities for fraudulent activities have expanded, necessitating more sophisticated detection systems. The traditional approaches to fraud detection, often rule-based, are no longer sufficient due to their lack of adaptability and scalability in the face of ever-evolving fraud tactics.

The objective of this project is to explore the application of machine learning algorithms in detecting fraudulent credit card transactions. This approach promises several advantages over traditional methods, including the ability to learn from data, adapt to new patterns of fraud, and process large volumes of transactions in real-time. The focus is on developing a model that can accurately distinguish between legitimate and fraudulent transactions, thereby enabling financial institutions to prevent fraud proactively.

Through this project, we aim to contribute to the ongoing efforts to secure digital transactions and protect financial assets. By leveraging the power of machine learning, we seek to create a robust fraud detection system that not only reduces the incidence of fraud but also minimizes the occurrence of false positives, which can lead to customer dissatisfaction and loss of trust in financial services.

3. Problem Statement

Credit card fraud detection is a complex problem that financial institutions grapple with daily. At its core, the issue stems from the need to identify unauthorized and fraudulent use of credit cards as they happen or, ideally, before they occur. The consequences of credit card fraud are far-reaching, affecting not only the financial health of institutions but also impacting consumer trust and privacy.

The challenge is twofold. First, the patterns of legitimate transactions are highly variable and influenced by countless factors, such as individual spending habits, broader economic trends, and seasonality. Fraudulent transactions, while anomalous, can often mimic legitimate behavior, making detection a non-trivial task. Second, and perhaps more problematic, is the imbalanced nature of the datasets involved in fraud detection. Fraudulent transactions represent a very small proportion of all credit card transactions, often less than 1%. This severe class imbalance poses significant difficulties for machine learning models, which may become biased towards the majority class and fail to detect the minority class — the actual fraudulent transactions.

The imbalanced datasets lead to models that are very accurate in predicting the majority class but perform poorly in detecting fraud. Standard accuracy metrics become misleading, and more sophisticated methods are necessary to evaluate and improve the model's performance on the minority class. Techniques such as resampling, specialized cost functions, and anomaly detection algorithms are often employed to mitigate these issues, but they bring their own challenges and trade-offs.

4. Market Analysis

The market for fraud detection is vast and growing. According to a report by MarketsandMarkets, the global fraud detection and prevention market size is expected to grow from USD 19.5 billion in 2018 to USD 63.5 billion by 2023, at a Compound Annual Growth Rate (CAGR) of 26.60%. This growth is fueled by the increasing digitization of financial services and the corresponding rise in cybercrime and card-not-present (CNP) fraud.

Advanced fraud detection methods are becoming a necessity as fraudsters continuously evolve their techniques. Traditional rule-based systems are proving insufficient due to their static nature and the high volume of false positives they generate. The financial industry is increasingly turning to machine learning and artificial intelligence to address these shortcomings. These technologies offer dynamic and adaptive solutions capable of detecting complex fraud patterns and learning from new data.

Machine learning algorithms can analyze vast datasets quickly and identify subtle patterns indicative of fraud. They can adapt to new and emerging fraud tactics, reducing the time and resources spent on manual review and rule-setting. The adoption of such advanced methods is not just a matter of improving fraud detection rates; it is becoming essential for the competitive positioning of financial institutions in a market that values security and customer experience.

The integration of machine learning in fraud detection systems is also opening up new market opportunities for fintech startups, cybersecurity firms, and software providers. These companies are innovating in the space, developing solutions that are not only more effective but also more cost-efficient, thus lowering the barrier to entry for smaller institutions to adopt state-of-the-art fraud prevention measures.

5. Methodology

5.1. Data Analysis Techniques

The dataset comprises credit card transactions from Europe in September 2013, spread over two days, consisting of 284,807 transactions. A critical aspect of this dataset is the severe imbalance, with only 492 transactions (0.172%) being fraudulent. This disproportionate distribution poses a unique challenge for any predictive model.

The data predominantly consists of numerical variables that have been anonymized through Principal Component Analysis (PCA), yielding features labeled V1 to V28. The specifics of these features are not disclosed due to privacy concerns, but they represent transformed characteristics of each transaction. Additionally, the dataset includes two original features:

'Time', indicating the seconds elapsed since the first transaction, and 'Amount', the transaction value.

Initial data analysis involves understanding the distribution of these features, particularly how they vary between fraudulent and legitimate transactions. This step includes plotting histograms, box plots, and correlation matrices to discern any distinct patterns or anomalies indicative of fraud.

5.2. Feature Engineering and Preprocessing

Considering the PCA-transformed nature of most features, extensive feature engineering might not add significant value. However, derived features from 'Time' (like hour of the day) could be explored.

Preprocessing involves several key steps:

Scaling: The 'Amount' feature, likely varying in range from other features, is standardized to ensure consistent model input.

Handling Class Imbalance: Given the skewed nature of the dataset, techniques like SMOTE (Synthetic Minority Over-sampling Technique) are employed to artificially balance the class distribution. This approach helps in enhancing the model's sensitivity towards fraudulent transactions.

5.3. Model Implementation: Random Forest Classifier

The choice of the Random Forest algorithm for this task is driven by its proficiency in handling complex datasets with a mix of numerical features. Random Forest, an ensemble learning method, operates by constructing multiple decision trees during training and outputting the class that is the mode of the classes (classification) of individual trees.

This model is particularly suited for the dataset at hand due to its ability to handle high-dimensional data and its robustness to overfitting, especially relevant given the large number of input features.

5.4. Hyperparameter Tuning and Model Selection

To optimize the Random Forest model, a hyperparameter tuning process is undertaken using GridSearchCV. This involves experimenting with various combinations of hyperparameters like:

n_estimators: Number of trees in the forest.

max_depth: Maximum depth of the tree.

min_samples_split: Minimum number of samples required to split an internal node.

The model's performance is evaluated using metrics such as precision, recall, F1-score, and the area under the ROC curve, with particular emphasis on recall to capture as many fraudulent transactions as possible.

6. Result and Discussion

In the application of the Random Forest algorithm to our dataset of credit card transactions, the model achieved an impressive accuracy of approximately 99.96%. This high level of accuracy is particularly encouraging but requires further scrutiny, especially given the highly imbalanced nature of the dataset.

Confusion Matrix Insights:

The confusion matrix shows that out of 15,426 transactions, the model correctly identified 15,397 as legitimate (true negatives) and 22 as fraudulent (true positives).

However, it incorrectly flagged 1 legitimate transaction as fraudulent (false positive) and missed 6 fraudulent transactions (false negatives).

Precision and Recall:

Given the rarity of fraudulent transactions, recall becomes a critical metric. In this case, the model's recall for fraudulent transactions is 78.57%, meaning it correctly identified about 78.57% of fraudulent transactions.

Precision for the fraudulent class is 95.65%, indicating that when it predicts a transaction as fraudulent, it is correct about 95.65% of the time.

F1-Score:

The F1-score is important for imbalanced datasets as it provides a balance between precision and recall. For fraudulent transactions, the F1-score would be calculated using the precision and recall values, indicating how well the model balances false positives and false negatives.

ROC-AUC Score:

The ROC-AUC score is a comprehensive measure showing the model's ability to distinguish between fraudulent and non-fraudulent transactions. An AUC close to 1 indicates excellent model performance. For this model, the high accuracy and good balance of precision and recall suggest a strong ROC-AUC score.

This detailed analysis, beyond mere accuracy, provides a more nuanced understanding of the model's performance, highlighting its strengths in detecting fraud while also pointing out areas for potential improvement, particularly in further reducing false negatives.

Strengths and Weaknesses:

Strengths: Random Forest is robust against overfitting, especially with such a high number of features. It is also good for unbalanced datasets like ours.

Weaknesses: The model can be computationally intensive, and the 'black box' nature of Random Forest can make interpretation challenging.

7. Financial Impact and Risk Analysis

The Random Forest model, with a recall of approximately 78.57%, effectively identifies fraudulent transactions. Assuming each fraud averages €100, the model's detection rate translates to potential savings of about €38,492 (78.57% of 492 fraudulent transactions, each valued at €100).

False Positives and Negatives: The cost of false positives (legitimate transactions flagged as fraud) could lead to customer dissatisfaction. False negatives (missed frauds) directly relate to financial loss. Balancing these two is key to the financial viability of the model.

Manual Review Reduction: A high-accuracy model can reduce the need for manual transaction reviews, potentially lowering labor costs and operational overhead.

Risks: Overfitting, despite Random Forest's robustness, is a potential risk, especially if the model is not regularly updated with new transaction data. The cost associated with model maintenance and updating should also be considered.

8. Conclusion

The project's findings underscore the effectiveness of using a Random Forest classifier in the detection of credit card fraud. With an accuracy of around 97%, the model shows significant promise in identifying fraudulent transactions. However, it is crucial to consider metrics beyond accuracy, such as recall and F1-score, due to the imbalanced nature of the dataset.

Implications: This study highlights the potential for machine learning algorithms to enhance the security of financial transactions and reduce monetary losses due to fraud.

9. Future Work

Future improvements could include:

Experimenting with additional feature engineering techniques.

Regularly updating the model with new data to maintain its accuracy.

Exploring deep learning approaches for potentially better performance.

Developing explainability methods to interpret Random Forest decisions, enhancing trust and understanding of the model.

References

1. ReportBuyer. (2018, December 5). The global fraud detection and prevention (FDP) market size is expected to grow from USD 19.5 billion in 2018 to USD 63.5 billion by 2023, at a Compound Annual Growth Rate (CAGR) of 26.6%. *PR Newswire*. https://www.prnewswire.com/news-releases/the-global-fraud-detection-and-prevention-fdp-market-size-is-expected-to-grow-from-usd-19-5-billion-in-2018-to-usd-63-5-billion-by-2023--at-a-compound-annual-growth-rate-cagr-of-26-6-300760746.html
2. Credit card fraud Detection. (2018, March 23). Kaggle. https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud/data