Credit Card Fraud Detection Model Report

By

Muhammad Ahmed Zaheer

Brainwave Matrix Solutions

AI/ML Intern

August, 2024.

Table of Contents

List of Figures

Abstract

This project focused on developing a robust credit card fraud detection system using both supervised learning methods and anomaly detection techniques to identify fraudulent transactions. Given the significant class imbalance in the dataset, data balancing techniques such as SMOTE and AdaSyn were applied to enhance model performance. Supervised methods, including Logistic Regression and XGBoost, were implemented, with hyperparameter tuning used to optimize their performance. The models were evaluated on metrics such as accuracy, precision, recall, and the F1-score. Anomaly detection techniques, specifically Autoencoders and One-Class SVM, were also employed. The One-Class SVM model underwent further improvement through feature selection and dimensionality reduction. Despite the time inefficiency in hyperparameter tuning for One-Class SVM, the final models achieved high accuracy and robust detection capabilities, highlighting the effectiveness of the applied methods.

1 Introduction

Credit card fraud is a significant and growing problem in the financial sector, affecting businesses and consumers alike. With the increasing volume of online transactions, the need for effective fraud detection systems has never been more critical. Fraudulent activities lead to substantial financial losses and can severely damage the reputation of financial institutions. Consequently, developing reliable and accurate fraud detection models has become a crucial focus of research and development.

In this project, we aim to address the challenge of credit card fraud detection by developing a robust model capable of identifying fraudulent transactions with high precision. The complexity of this problem arises from the highly imbalanced nature of the dataset, where legitimate transactions vastly outnumber fraudulent ones. To tackle this, we employed both supervised learning methods and anomaly detection techniques.

Our approach involved the following key methodologies:

1. **Supervised Learning**: We implemented supervised learning techniques to classify transactions as fraudulent or legitimate. Given the imbalanced nature of the dataset, techniques such as SMOTE (Synthetic Minority Over-sampling Technique) and ADASYN (Adaptive Synthetic Sampling) were utilized to address data imbalance, thereby improving model performance.
2. **Anomaly Detection**: We explored anomaly detection methods, specifically Autoencoders and One-Class SVMs, which are well-suited for detecting rare and unusual events like fraud.

While our initial supervised models achieved high accuracy, we recognized the importance of exploring additional techniques such as anomaly detection. This exploration led to the development of an autoencoder-based model and a One-Class SVM model, both of which provided valuable insights into the detection of fraudulent transactions. Furthermore, efforts were made to enhance the One-Class SVM model through hyperparameter tuning,

though this approach was ultimately deemed too time-consuming to be practical within our project constraints.

In this report, we present the details of our approach, including the methodologies employed, the results obtained, and the challenges faced during the development process. Our findings demonstrate the effectiveness of combining both supervised and anomaly detection methods in building a comprehensive fraud detection system.

2 Methodology

The methodology outlines the systematic approach adopted to develop and evaluate the credit card fraud detection model using supervised and anomaly detection methods. This section details the dataset used, data preprocessing steps, and information about the techniques and evaluation metrics used.

2.1 Dataset

The dataset used for this project was sourced from Kaggle and contains 284,807 credit card transactions made by European cardholders over two days in September 2013. The data is highly imbalanced, with only 492 fraudulent transactions (approximately 0.172%) out of the total. The dataset includes 31 features: 28 anonymized numerical features obtained via PCA, along with Time, Amount, and the target variable Class, where Class is 1 for fraudulent transactions and 0 for legitimate ones.

2.2 Data Preprocessing

Data preprocessing is crucial for preparing the raw data into a format suitable for modeling. Initially, the Time and Amount features were normalized using MinMaxScaler to scale their values between 0 and 1. Since the anonymized features were already standardized, no further scaling was required. The dataset had no missing values, so imputation was

unnecessary. The data was split into features (X) and the target variable (y), setting the stage for training and evaluation across different models.

2.3 Supervised Methods

Supervised learning involves training a model on labeled data, meaning the model learns from input-output pairs to make predictions on new data. Two supervised methods, Logistic Regression and XGBoost, were employed for fraud detection.

2.3.1 Logistic Regression

Logistic Regression is a linear model used for binary classification tasks. It calculates the probability that a given input belongs to a particular class by applying the logistic function to a linear combination of input features. Despite its simplicity, Logistic Regression is highly effective for classification problems and serves as a solid baseline.

In this project, Logistic Regression was first applied to the imbalanced dataset. Given the dataset's imbalance, the model naturally struggled to identify fraudulent transactions. To improve performance, logistic regression was also used with balanced data generated by SMOTE and AdaSyn.

2.3.2 XGBoost

XGBoost (Extreme Gradient Boosting) is an advanced boosting algorithm known for its high performance and scalability. It builds an ensemble of weak learners, typically decision trees, and combines them to create a strong predictive model. XGBoost handles missing data, captures complex patterns, and offers several regularization techniques to avoid overfitting.

In the project, XGBoost was applied directly to the imbalanced dataset, producing better results than Logistic Regression due to its ability to learn from the intricate patterns in the data. XGBoost was also utilized with data balanced by SMOTE and AdaSyn to compare its performance under different conditions.

2.4 Handling Data Imbalance

Due to the extreme class imbalance in the dataset, specialized techniques were necessary to ensure the models could effectively detect fraudulent transactions. Two popular oversampling methods, SMOTE and AdaSyn, were employed.

2.4.1 SMOTE

SMOTE (Synthetic Minority Over-sampling Technique) is an oversampling technique that generates synthetic samples for the minority class (fraudulent transactions) by interpolating between existing samples. By doing so, it reduces the class imbalance and allows models to learn more effectively from the minority class.

In the project, SMOTE was applied to create a balanced dataset, which was then used to train both the Logistic Regression and XGBoost models. The models showed marked improvements in detecting fraud with SMOTE-applied data.

2.4.2 AdaSyn

AdaSyn (Adaptive Synthetic Sampling) is another oversampling technique, similar to SMOTE, but with a focus on generating more synthetic data for harder-to-learn minority class samples. It dynamically adjusts the number of synthetic samples created based on the data distribution, aiming to further refine the model's ability to learn from imbalanced data.

AdaSyn was also used to balance the dataset. The Logistic Regression and XGBoost models were trained on this balanced data, and their performance was compared with that of the SMOTE-based models.

2.5 Supervised Methods with Balanced Data

After balancing the dataset using SMOTE and AdaSyn, both Logistic Regression and XGBoost models were retrained to assess the impact of oversampling on model performance.

2.5.1 Logistic Regression with SMOTE

Logistic Regression was first trained using data balanced with SMOTE. The model demonstrated a significant improvement in recall for detecting fraudulent transactions, though some trade-offs were observed in precision.

2.5.2 Logistic Regression with AdaSyn

Similarly, Logistic Regression was trained on AdaSyn-balanced data. This model showed comparable improvements to the SMOTE-based version, with subtle differences in precision and recall, further highlighting the importance of data balancing techniques.

2.5.3 XGBoost with SMOTE

XGBoost was also trained on the SMOTE-balanced data, showing improved detection of fraudulent transactions. The model's ability to capture complex patterns made it particularly effective when combined with SMOTE.

2.5.4 XGBoost with AdaSyn

Finally, XGBoost was trained on data balanced with AdaSyn. Like the SMOTE version, this model exhibited strong performance in fraud detection, further proving the robustness of XGBoost when combined with oversampling techniques.

2.6 Anomaly Detection Methods

Given the rarity of fraudulent transactions, anomaly detection methods were explored as an alternative approach. These methods focus on identifying outliers in the data, which are assumed to be fraud.

2.6.1 Autoencoders

Autoencoders are a type of neural network designed for unsupervised learning, where the goal is to compress data into a lower-dimensional representation and then reconstruct it. By training the autoencoder on non-fraudulent transactions, the model learns to accurately reconstruct them. Fraudulent transactions, being anomalies, result in larger reconstruction errors and are thus identified as outliers.

In this project, a basic autoencoder was first implemented, followed by an improved version that employed feature engineering and regularization techniques to enhance performance.

2.6.2 One Class SVM

One-Class SVM is a popular anomaly detection method that attempts to learn the decision boundary that encloses the majority class (legitimate transactions) while identifying outliers (fraudulent transactions). The model was trained exclusively on legitimate transactions and then used to predict anomalies in the dataset.

Initially, a basic One-Class SVM model was implemented, which yielded decent results. To improve its performance, feature selection and dimensionality reduction were applied to the data before training the One-Class SVM. While attempts were made to further improve the model through hyperparameter tuning, it was ultimately deemed unfeasible due to excessive time requirements.

2.7 Evaluation Metrics

Model performance was evaluated using several metrics, including accuracy, precision, recall, F1-score, and ROC AUC score. Precision and recall were especially important due to the class imbalance, as they provide insight into how well the models detect fraudulent transactions without generating excessive false positives.

3 Results and Discussion

The supervised models, Logistic Regression and XGBoost, initially performed well in detecting fraudulent transactions. However, the imbalance in the dataset caused significant challenges, with both models achieving high accuracy but lower recall for the minority class. This was particularly evident with Logistic Regression, which struggled more with recall compared to XGBoost. The results obtained with each model on the test set are shown below:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	56866
1	0.77	0.56	0.65	96
accuracy			1.00	56962
macro avg	0.89	0.78	0.83	56962
weighted avg	1.00	1.00	1.00	56962

Figure 1: Logistic Regression

	precision	recall	f1-score	support
0	1.00	1.00	1.00	56866
1	0.90	0.74	0.81	96
accuracy			1.00	56962
macro avg	0.95	0.87	0.91	56962
weighted avg	1.00	1.00	1.00	56962

Figure 2: XGBoost

	precision	recall	f1-score	support
0	1.00	0.97	0.99	56866
1	0.06	0.90	0.11	96
accuracy			0.97	56962
macro avg	0.53	0.94	0.55	56962
weighted avg	1.00	0.97	0.99	56962

Figure 3: Logistic Regression with SMOTE

	precision	recall	f1-score	support
0	1.00	0.91	0.95	56866
1	0.02	0.96	0.03	96
accuracy			0.91	56962
macro avg	0.51	0.93	0.49	56962
weighted avg	1.00	0.91	0.95	56962

Figure 4: Logistic Regression with AdaSyn

	precision	recall	f1-score	support
0	1.00	1.00	1.00	56866
1	0.70	0.79	0.74	96
accuracy			1.00	56962
macro avg	0.85	0.90	0.87	56962
weighted avg	1.00	1.00	1.00	56962

Figure 5: XGBoost with SMOTE

	precision	recall	f1-score	support
0	1.00	1.00	1.00	56866
1	0.66	0.77	0.71	96
accuracy			1.00	56962
macro avg	0.83	0.89	0.86	56962
weighted avg	1.00	1.00	1.00	56962

Figure 6: XGBoost with AdaSyn

Classification Report:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	85307
1	0.50	0.01	0.01	136
accuracy			1.00	85443
macro avg	0.75	0.50	0.51	85443
weighted avg	1.00	1.00	1.00	85443

Figure 7: Autoencoder

Classification Report:

	precision	recall	f1-score	support
0	1.00	0.97	0.98	284315
1	0.04	0.66	0.07	492
accuracy			0.97	284807
macro avg	0.52	0.82	0.53	284807
weighted avg	1.00	0.97	0.98	284807

Figure 8: One Class SVM

```
Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.97      0.98    284315
           1       0.04      0.67      0.07       492

    accuracy                           0.97    284807
   macro avg       0.52      0.82      0.53    284807
weighted avg       1.00      0.97      0.98    284807
```

Figure 9: One Class SVM with feature selection and dimensionality reduction

To address the data imbalance, we applied SMOTE and AdaSyn. These techniques significantly improved recall for the minority class, particularly for Logistic Regression, although precision was affected. XGBoost also showed improvements, with a more balanced performance between precision and recall. However, it's clear that these resampling techniques improved detection but introduced trade-offs, particularly in the form of false positives.

In contrast, anomaly detection methods, including Autoencoders and One-Class SVM, focused on identifying fraud as an anomaly. The One-Class SVM initially showed promising results, but after feature selection and dimensionality reduction, the model's performance improved, especially in recall. However, these models still struggled with precision, which is a common challenge in anomaly detection for fraud detection.

In summary, each approach presented its strengths and weaknesses, highlighting the complexity of fraud detection in highly imbalanced datasets. Supervised methods offered more consistent results, particularly with data resampling, while anomaly detection provided an alternative perspective but faced difficulties in precision.

4 Conclusion

In this project, we explored various approaches to detect fraudulent credit card transactions, addressing the inherent challenge of working with highly imbalanced data. Our investigation spanned both supervised learning techniques—Logistic Regression and XGBoost—and anomaly detection methods—Autoencoders and One-Class SVM. Each method demonstrated distinct advantages and limitations.

Supervised learning methods, particularly XGBoost, achieved high accuracy and recall, especially after applying SMOTE and AdaSyn to balance the data. However, these improvements came with the trade-off of increased false positives. The anomaly detection techniques, while innovative, faced challenges with precision, particularly in the One-Class SVM. Despite improvements through feature selection and dimensionality reduction, the One-Class SVM struggled to maintain a balance between recall and precision.

Ultimately, our results underscore the complexity of fraud detection, where no single model or method excels in all metrics. While supervised learning with resampling techniques provided the most practical solution, anomaly detection remains a valuable tool, particularly in scenarios where labeled data is scarce. Future work could focus on refining these models through hyperparameter tuning or exploring hybrid approaches that combine the strengths of both supervised and unsupervised methods.