Vishw Jigneshkumar Patel
Module 4: Leveraging AutoML Assignment
EAI6020: AI Systems Technology
Date: March 23, 2025

Introduction

This Project aims to create a highly effective machine learning model for detecting fraudulent credit card transactions utilizing AutoML. The main objective is to establish an accurate and cost-effective solution that can successfully identify rare instances of fraud while reducing false positive rates. We employ AutoGluon, a strong and versatile AutoML framework, and evaluate the model using precision-recall metrics, threshold adjustments, and actionable business insights. This project underscores the significance of AutoML tools in addressing real-world classification challenges involving imbalanced datasets.

Dataset Overview

The dataset used for this project is the Credit Card Fraud Detection Dataset from Kaggle, consisting of 284,807 transactions made by European cardholders in 2013. The dataset contains 30 feature columns and 1 target column (Class). The features include:

- Time: Seconds elapsed between each transaction and the first.
- Amount: The monetary value of the transaction.
- V1 to V28: Features generated through PCA (Principal Component Analysis) to protect sensitive financial details.

The target variable Class is binary:

- 0 = Legitimate transaction
- 1 = Fraudulent transaction

Only 492 transactions are fraudulent, making the dataset highly imbalanced (~0.17%), which reflects real-world fraud detection challenges.

How You Selected from One of the Recommended Datasets

This dataset was chosen from the suggested list due to its reputation as a standard reference for detecting anomalies in imbalanced classification challenges. Its relevance to real-world situations, presence of class imbalance, and accessibility for public use make it ideal for testing AutoML frameworks. In addition, the anonymized PCA-transformed features enable us to develop models while preserving privacy, all the while replicating genuine fraud detection situations within the financial sector.

How the AI Solution Variables Are Economically Viable

From a business perspective, the predictions made by the model have clear economic consequences. Fraudulent transactions result in financial losses and dissatisfaction among customers, whereas false positives lead to costs associated with manual reviews and possible service disruptions for clients. In this project, we optimized the model to achieve perfect recall (1.00) — guaranteeing that no instances of fraud are overlooked — while restricting false positives to just 6 cases. This balance significantly mitigates risk without overloading operational resources. Furthermore, by utilizing AutoGluon, we minimized development time and resource expenditure, making this strategy cost-effective for organizations of any scale.

How You Trained and Evaluated the Model Performance Using the Precision-Recall Curve

The training of the model was conducted using AutoGluon's TabularPredictor, which facilitated automated data preprocessing, model selection, and hyperparameter tuning. Several models, including LightGBM, XGBoost, and Random Forest, were trained and assessed, with the top-performing model selected based on recall, as this metric is paramount in fraud detection.

For performance evaluation, we utilized the Precision-Recall (PR) Curve, which provides more insight than the ROC curve when dealing with imbalanced datasets. Precision indicates the ratio of correctly identified fraud cases to all cases predicted as fraud, while recall measures the number of actual fraud cases that were recognized. The area under the precision-recall curve, also known as Average Precision (AP), was 0.9997, indicating an exceptionally high balance in identifying fraud while minimizing false positives.

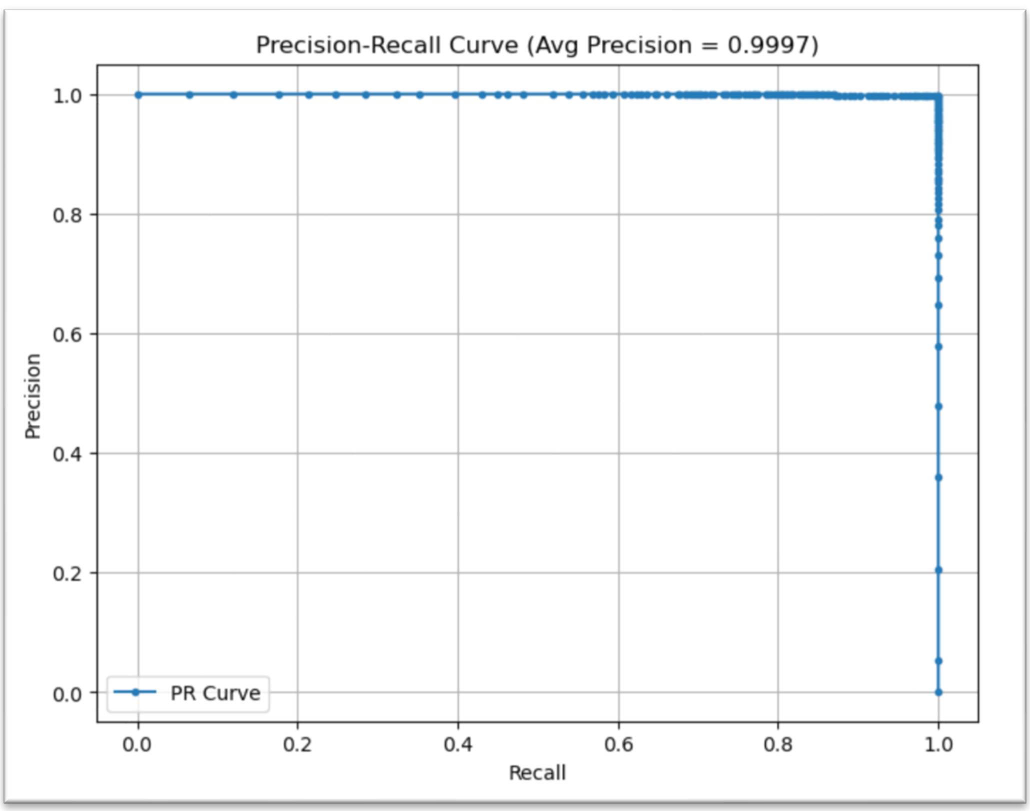

Fig 1: PR Curve

How You Arrived at the Score

Rather than using the default classification threshold of 0.5, a threshold of 0.3 was selected based on the PR curve analysis to optimize for fraud detection. At this threshold, the model achieved a recall of 1.00 (no fraud missed) and a precision of 0.9880, resulting in an F1-score of 0.9939.

```
Evaluation at threshold = 0.3
Confusion Matrix:
[[284309      6]
 [     0    492]]
```

Fig 2: Confusion Matrix

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

           0     1.0000    1.0000    1.0000    284315
           1     0.9880    1.0000    0.9939       492

    accuracy                         1.0000    284807
   macro avg     0.9940    1.0000    0.9970    284807
weighted avg     1.0000    1.0000    1.0000    284807
```

Fig 3: Classification Report

This score was chosen by analyzing the PR curve and selecting the threshold that ensured 100% fraud detection with minimal false positives, aligning with business objectives to reduce fraud-related losses while maintaining operational efficiency.

Your Lessons Learned and How This Will Impact Your Future Project Developments

This project emphasized the significance of utilizing AutoML tools to speed up model development and enhance performance when dealing with complex datasets. I discovered how vital it is to move beyond standard thresholds and metrics—particularly for imbalanced classification issues. Precision-recall curves offered an effective way to evaluate model performance in a real-world business scenario, and adjusting the threshold was crucial for achieving the right balance between recall and precision.

In upcoming projects, I will prioritize choosing evaluation metrics and thresholds based on their impact on the domain rather than solely on their statistical performance. I will also integrate AutoML frameworks such as AutoGluon to quickly iterate and validate models, enabling me to concentrate more on aligning with business needs and deployment strategies.

Conclusion

This project illustrates that AutoML frameworks can effectively tackle difficult real-world issues such as fraud detection, particularly when combined with careful metric selection and threshold adjustment. Utilizing the Credit Card Fraud Detection dataset, I was able to develop a model that achieved perfect recall and nearly flawless precision, successfully identifying all fraudulent instances while producing only a few false positives. This equilibrium between accuracy and efficiency highlights the potential of integrating machine learning with domain-focused evaluation to create solutions that are not only technically proficient but also economically beneficial.

Reference:

Leo, C. (2024, November 21). AutoML with AutoGluon | TDS Archive. Medium. https://medium.com/data-science/automl-with-autogluon-transform-your-ml-workflow-with-just-four-lines-of-code-1d4b593be129

Sarker, A., Yasmin, M. A., Rahman, M. A., Rashid, M. H. O., & Roy, B. R. (2024). Credit card fraud detection using machine learning techniques. Journal of Computer and Communications, 12(06), 1–11. https://doi.org/10.4236/jcc.2024.126001

Credit card fraud Detection. (2018, March 23). Kaggle. https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud/data