2025

Detecting Deception

MACHINE LEARNING APPROACHES TO CREDIT CARD FRAUD IDENTIFICATION

JENAYA MODESTE, ITCS 3155-091

Introduction

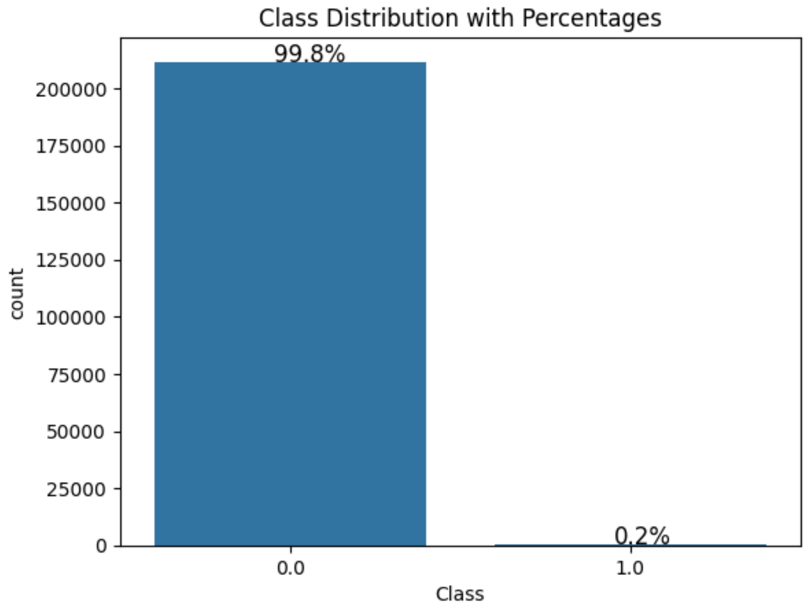
Credit card fraud is a major global challenge that affects individuals and financial institutions worldwide, costing billions of dollars annually and posing significant security risks to the public. Credit card fraud occurs when an attacker uses a victim's existing credit card information and personal data to create fraudulent accounts or perform unauthorized transactions. Using stolen information, perpetrators can open multiple new accounts, further contributing to large-scale data breaches. Although fraudulent transactions represent only a small fraction of overall credit card activity, they are difficult to detect due to their evolving nature. As fraud techniques continue to advance, detection models must be regularly updated to remain effective.

In this project, we applied machine learning techniques to the widely used Credit Card Fraud Detection dataset. The primary objective is to build predictive models capable of identifying fraudulent transactions with high precision and recall while minimizing false positives. Effective fraud detection systems are critical for protecting customers, reducing financial losses, and maintaining trust in digital payment systems.

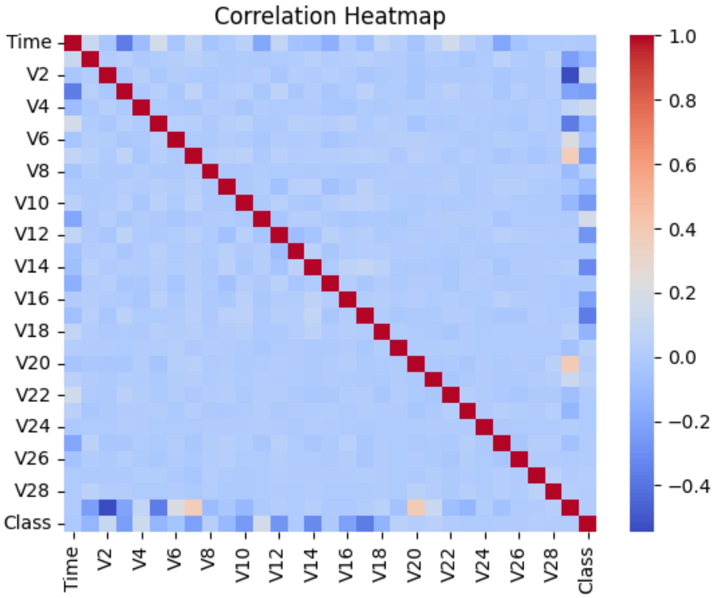
Data

The dataset used in this project is the **Credit Card Fraud Detection** dataset obtained from Kaggle. It contains **284,807 transactions** with **31 total features**, including Time, Amount, Class, and **28 anonymized numerical features** derived using Principal Component Analysis (PCA). A key characteristic of this dataset is its severe class imbalance: most transactions are legitimate, while only a small percentage are fraudulent. Fraudulent transactions make up approximately **2%** of the dataset, as shown in the class distribution visualization.

This imbalance indicates that standard accuracy metrics can be misleading, as a model can achieve high accuracy by simply predicting all transactions as legitimate. Consequently, evaluation metrics such as **precision, recall, F1-score, and AUC** are more appropriate for this analysis. Additionally, the imbalance results in a right-skewed distribution, increasing the likelihood of overlap between legitimate and fraudulent transactions. With these considerations in mind, this project aims to evaluate models using metrics that extend beyond simple accuracy.



The correlation heatmap reveals minimal linear correlation among the features. Since fraudulent behavior often varies by individual case, a lack of strong linear relationships is expected. Furthermore, because the dataset consists largely of PCA-transformed features, each component captures distinct variance rather than direct relationships. This characteristic highlights the potential need for non-linear models, such as random forests, to better capture complex patterns.



Pre-processing

During preprocessing, we applied standard scaling to the Amount and Time features, as their distributions differ significantly from the PCA components. Scaling these features ensures that they contribute appropriately during model training. Prior to scaling, we checked for missing values and duplicate records, both of which could negatively impact model performance. The dataset was found to contain no null values and minimal duplication, indicating that it was relatively clean.

The dataset was then split into **training (80%)** and **testing (20%)** sets, ensuring that the class ratio was preserved in each split. This step is particularly important given the small number of fraudulent transactions, as it helps prevent further imbalance and ensures reliable evaluation on unseen data.

Methods

Logistic regression

Logistic regression is a supervised machine learning algorithm commonly used for binary classification problems. It predicts the probability that a given input belongs to a particular class using the sigmoid function, which maps real-valued inputs to a range between 0 and 1. In this project, logistic regression is used to classify transactions as either fraudulent (1) or legitimate (0).

Sigmoid Function

$$\sigma\,(t) = \frac{1}{1 + e^{-t}}$$

Random Forest Regression

The random forest regressor is an ensemble learning algorithm that combines predictions from multiple decision trees to produce a more robust and accurate result. In the context of fraud detection, random forests are well-suited for capturing non-linear relationships and identifying complex risk patterns associated with fraudulent behavior.

Key advantages of this model include its ability to handle large datasets, reduce overfitting, and remain robust to outliers. However, random forests can be computationally expensive, memory-intensive, and susceptible to overfitting in noisy or severely imbalanced datasets. Given the imbalance present in this dataset, careful preprocessing and evaluation are necessary to ensure optimal model performance.

Results

As described earlier, the dataset was split into **80% training** and **20% testing** sets. Model performance was evaluated using **precision, recall, F1-score, ROC-AUC, and precision–recall AUC**. Precision measures how accurately the model identifies fraudulent transactions, while recall measures how many actual fraudulent transactions are successfully detected. The F1-score provides a balance between precision and recall, and the ROC-AUC evaluates the model's ability to rank fraudulent transactions higher than legitimate ones.

Logistic Regression Results

The logistic regression model achieved an overall accuracy of **99.91%**. However, this high accuracy is largely influenced by the dominance of legitimate transactions in the dataset.

Class 0 (Legitimate Transactions):

- Precision: 1.00
- Recall: 1.00
- F1-score: 1.00
- Support: 42,324

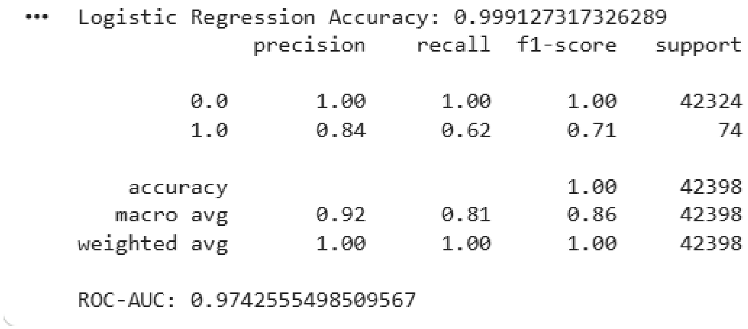
These results indicate that the model correctly classified nearly all legitimate transactions, with very few false positives.

Class 1 (Fraudulent Transactions):

- Precision: 0.84
- Recall: 0.62
- F1-score: 0.71
- Support: 74

When the model predicts fraud, it is correct **84%** of the time. However, it identifies only **62%** of all fraudulent transactions, meaning that approximately **38%** of fraud cases are missed. This recall rate may be insufficient for real-world fraud detection systems.

The model achieved a **ROC-AUC score of 97.42%**, indicating a strong ability to distinguish between fraudulent and legitimate transactions.

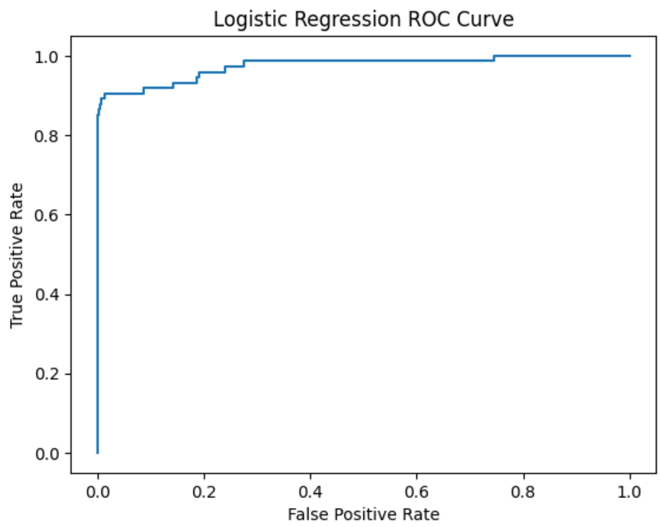


```
Logistic Regression Accuracy: 0.999127317326289
              precision    recall  f1-score   support

         0.0       1.00      1.00      1.00     42324
         1.0       0.84      0.62      0.71        74

    accuracy                           1.00     42398
   macro avg       0.92      0.81      0.86     42398
weighted avg       1.00      1.00      1.00     42398

ROC-AUC: 0.9742555498509567
```



Random Forest Regression Results

The random forest regressor achieved an overall accuracy of **99.95%**, again influenced by the class imbalance.

Class 0 (Legitimate Transactions):

- Precision: 1.00
- Recall: 1.00
- F1-score: 1.00
- Support: 42,324

Class 1 (Fraudulent Transactions):

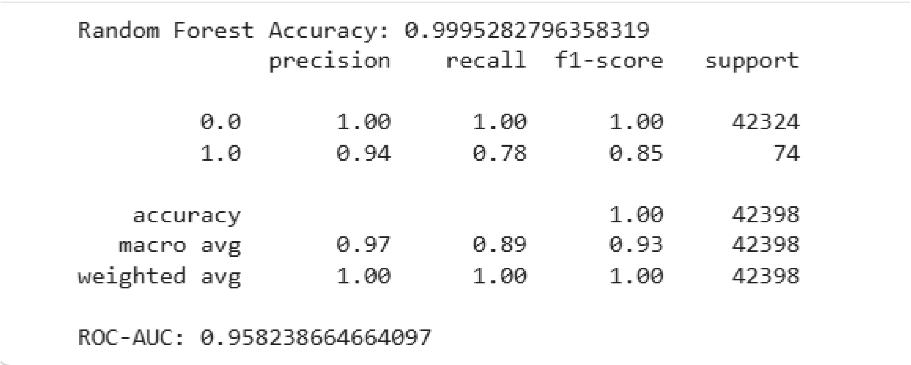
- Precision: 0.94

- Recall: 0.78

- F1-score: 0.85

- Support: 74

These results show that the random forest model outperforms logistic regression in detecting fraudulent transactions. When predicting fraud, the model is correct **94%** of the time and captures **78%** of all fraud cases. The **ROC-AUC score of 95.82%** is slightly lower than that of logistic regression but still indicates strong overall performance.

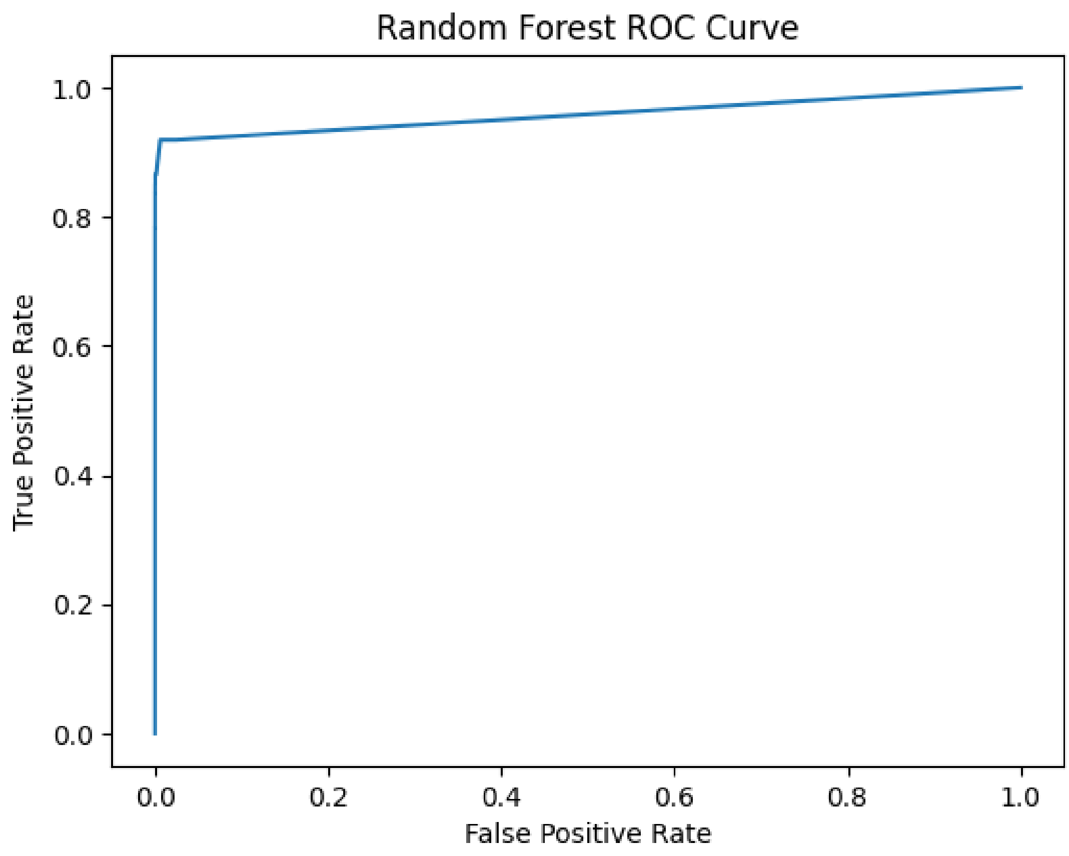


```
Random Forest Accuracy: 0.9995282796358319
              precision    recall  f1-score   support

         0.0       1.00      1.00      1.00     42324
         1.0       0.94      0.78      0.85        74

    accuracy                           1.00     42398
   macro avg       0.97      0.89      0.93     42398
weighted avg       1.00      1.00      1.00     42398

ROC-AUC: 0.958238664664097
```

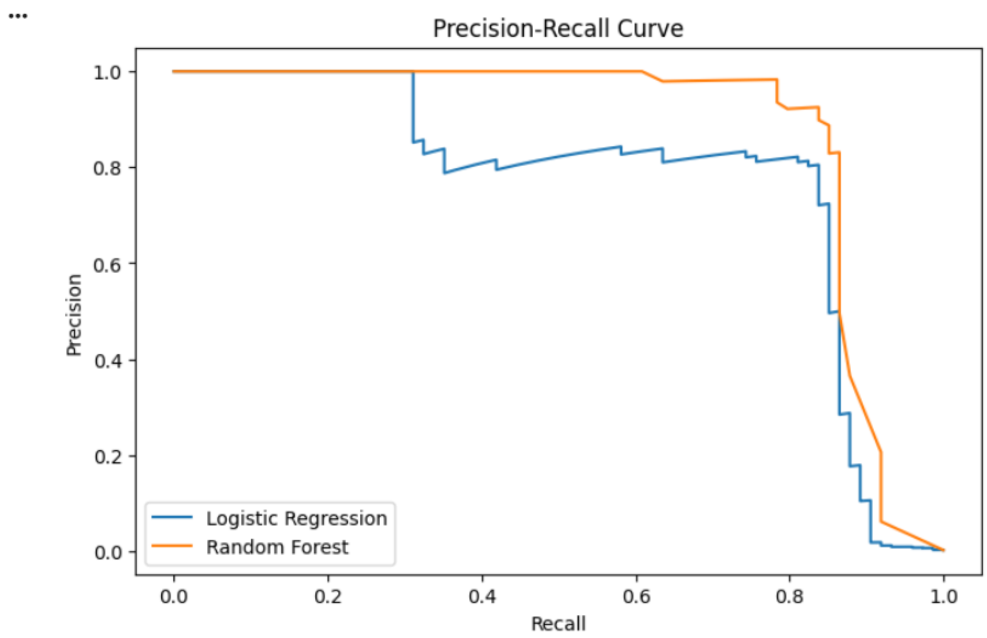


Precision-Recall curve analysis

Figure X presents the precision–recall (PR) curves for both the logistic regression and random forest models. The PR curve is particularly informative for highly imbalanced datasets, such as credit card fraud detection, because it focuses on the trade-off between precision and recall for the minority (fraudulent) class.

The random forest model consistently maintains higher precision across a broader range of recall values compared to logistic regression. This indicates that as the random forest model identifies more fraudulent transactions, it does so while producing fewer false positives. In contrast, the logistic regression model experiences a more rapid decline in precision as recall increases, suggesting weaker performance when attempting to capture a larger proportion of fraud cases.

Overall, the area under the precision–recall curve further supports earlier findings: the random forest model demonstrates superior performance in identifying fraudulent transactions within an imbalanced dataset, making it a more reliable choice for real-world fraud detection systems.



Conclusion

In this project, we explored credit card fraud detection using machine learning models applied to a large, imbalanced dataset. Through this analysis, we gained valuable insights into data preprocessing, model selection, and evaluation in real-world fraud detection scenarios.

One key takeaway is that imbalanced datasets require special handling, as many models can produce misleadingly high accuracy while performing poorly in minority classes. This highlights the importance of using evaluation metrics such as precision, recall, and F1-score rather than relying solely on accuracy. Additionally, proper scaling and preprocessing were shown to significantly improve model performance, particularly when working with features that vary widely in scale.

Overall, this project demonstrated how machine learning models can effectively detect fraudulent activity when applied carefully and evaluated appropriately. The findings reinforce best practices in handling imbalanced data and illustrate the practical value of machine learning in enhancing financial security systems.

Resources

Experian. (n.d.). *Credit card fraud: What to do if you are a victim*. https://www.experian.com/blogs/ask-experian/credit-education/preventing-fraud/credit-card-fraud-what-to-do-if-you-are-a-victim/

GeeksforGeeks. (n.d.). *Random forest regression in Python*. https://www.geeksforgeeks.org/machine-learning/random-forest-regression-in-python/

GeeksforGeeks. (n.d.). *Understanding logistic regression*. https://www.geeksforgeeks.org/machine-learning/understanding-logistic-regression/

Transparency Note: Parts of this assignment (partial graph creation, Overall Writing Structure) were developed with the assistance of ChatGPT to improve clarity and structure.