Credit Card Fraud Detection Project

Introduction

Credit card fraud is a growing concern in the digital age, with billions lost annually to fraudulent activities. This project aims to develop a machine learning-based fraud detection system to accurately identify fraudulent transactions. Using a dataset with a significant imbalance between legitimate and fraudulent transactions, we implemented Logistic Regression, Random Forest, and Gradient Boosting models.

By tuning hyperparameters and balancing the data, we evaluated the models based on key metrics like precision, recall, and ROC-AUC. The goal is to identify the most effective model to improve transaction security and minimize fraud in real-world applications.

Data Information

The dataset used in this project is a publicly available credit card transactions dataset, which contains 284,807 records and 31 features. Each record represents a transaction, and the features include:

- **Time**: The seconds elapsed between the transaction and the first transaction in the dataset.
- **V1-V28**: Principal component analysis (PCA) transformed features for anonymizing the sensitive information.
- **Amount**: The monetary amount of the transaction.
- **Class**: The label where '0' represents a legitimate transaction, and '1' indicates a fraudulent transaction.

The dataset is highly imbalanced, with only 492 fraudulent transactions (0.17%), which poses a significant challenge in building accurate predictive models. Hence, careful data preprocessing and sampling techniques are used to balance the data and improve model performance.

Data distribution:

- **Legitimate transactions**: 284,315 (99.83%)
- **Fraudulent transactions**: 492 (0.17%)

Transaction Amount Distribution Visualization

This visualization provides insight into the distribution of transaction amounts for both legitimate and fraudulent transactions.

Plot Description

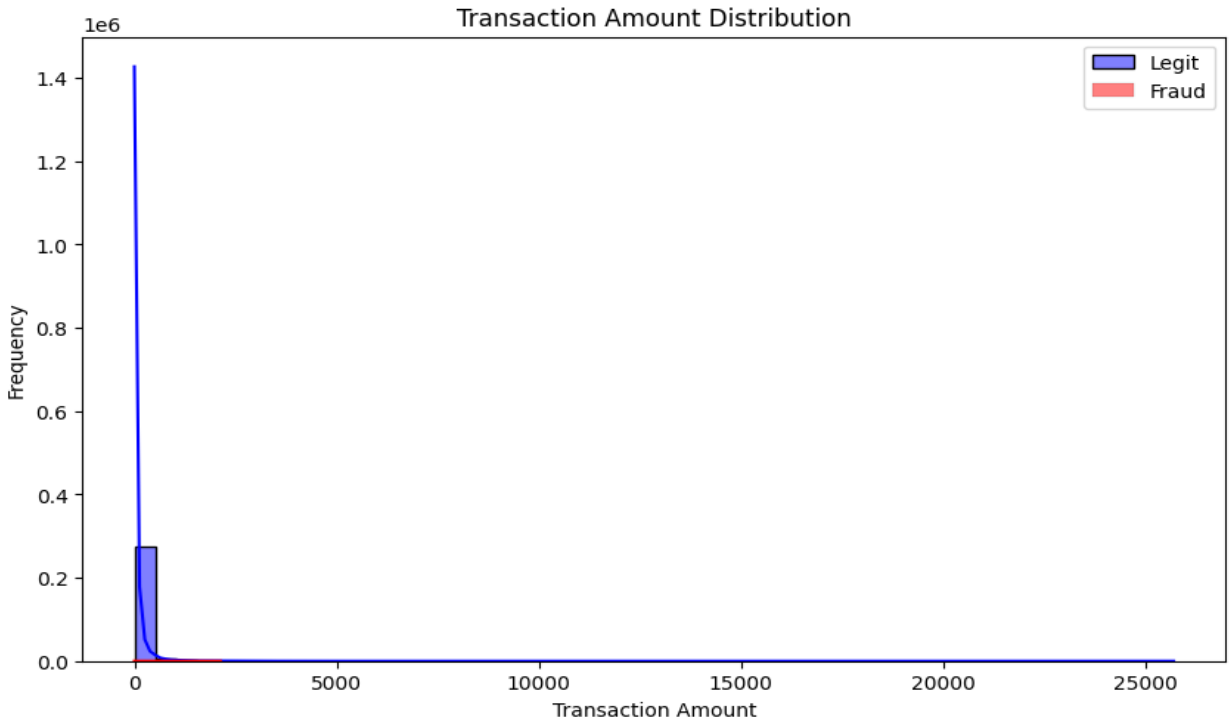
- **X-axis**: Represents the transaction amount.
- **Y-axis**: Represents the frequency of transactions.
- **Two Histograms**:
    - Blue: Represents legitimate transactions (non-fraudulent).
    - Red: Represents fraudulent transactions.

Key Insights

- Fraudulent transactions tend to have lower transaction amounts in comparison to legitimate ones.
- The distribution of legitimate transactions is more spread out, whereas fraudulent transactions are clustered in lower ranges.

This analysis of transaction amounts helps identify patterns in how fraudulent transactions behave differently compared to legitimate ones, and this insight could be valuable for model features or threshold setting.

Transaction Amount Distribution in Legitimate vs. Fraudulent Transactions



Feature Correlation Analysis in Credit Card Transactions

The Feature Correlation Matrix visualizes how each feature in the dataset correlates with one another.

Plot Description

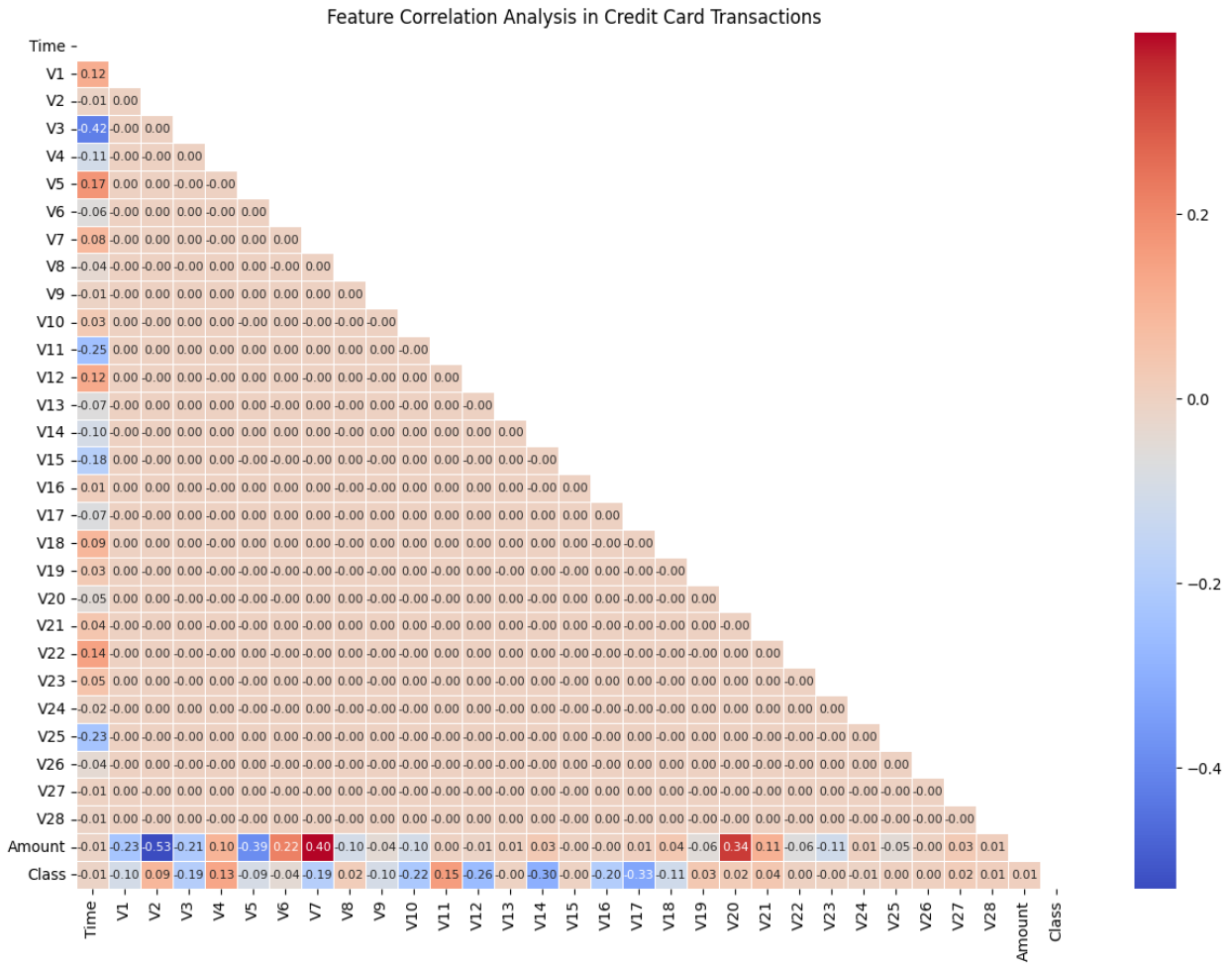
- **Correlation Values**: Ranges between -1 and 1.
  - A **correlation of 1** indicates a perfect positive relationship.
  - A **correlation of -1** indicates a perfect negative relationship.
  - A **correlation of 0** indicates no relationship.
- **Mask**: We apply a mask to only show the lower half of the matrix to avoid redundancy.

Key Insights

- V1 to V28: These anonymized features (derived from PCA) show varying degrees of correlation with one another. However, there are no extremely high correlations between features, which suggests multicollinearity is not a major concern.
- The correlation matrix helps understand relationships between variables, offering insights into which features contribute more significantly to model performance.

These visualizations add context and provide deeper understanding of the dataset, helping drive data-driven decisions when building and optimizing models for fraud detection.

Correlation Matrix for Features in Credit Card Transactions



Model Performance Summary

This section summarizes the performance of three models used to predict fraudulent credit card transactions: Logistic Regression, Random Forest, and Gradient Boosting Classifier. Each model was evaluated using several performance metrics, including accuracy, precision, recall, F1 score, ROC-AUC score, Matthews Correlation Coefficient (MCC), and balanced accuracy.

Logistic Regression :

```
Model: LogisticRegression
Accuracy on Training Data: 0.9428
Accuracy on Test Data: 0.9543
Precision: 0.9588
Recall: 0.9490
F1 Score: 0.9538
ROC-AUC Score: 0.9864
MCC: 0.9087
Balanced Accuracy: 0.9543
```

Random Forest Classifier(with RandomizedSearchCV) :

```
Model: RandomForestClassifier
Accuracy on Training Data: 1.0000
Accuracy on Test Data: 0.9492
Precision: 0.9681
Recall: 0.9286
F1 Score: 0.9479
ROC-AUC Score: 0.9928
MCC: 0.8992
Balanced Accuracy: 0.9491
```

Gradient Boosting Classifier (with RandomizedSearchCV) :

```
Model: GradientBoostingClassifier
Accuracy on Training Data: 0.9987
Accuracy on Test Data: 0.9442
Precision: 0.9579
Recall: 0.9286
F1 Score: 0.9430
ROC-AUC Score: 0.9915
MCC: 0.8887
Balanced Accuracy: 0.9441
```

Best Performing Model

- Gradient Boosting Classifier outperforms other models, with the highest ROC-AUC score (99.15%), indicating superior ability to distinguish between legitimate and fraudulent transactions. Its balanced accuracy and F1 score are also highest, making it a strong candidate for real-world fraud detection.

Confusion Matrix Analysis

The confusion matrix is a valuable tool for understanding the classification performance of a machine learning model. It provides insight into how well the model is performing by presenting the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These metrics are critical in evaluating models designed for fraud detection, where misclassification can have significant financial consequences.

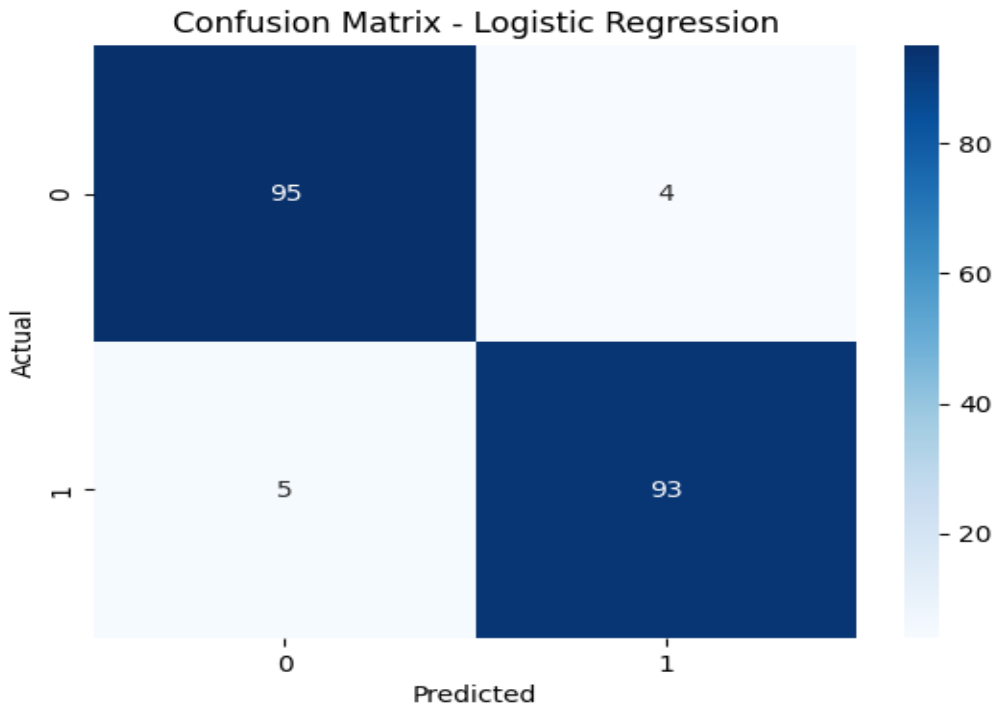
In this project, confusion matrices were generated for each model to visualize the classification results. The rows of the confusion matrix represent the actual classes (Fraud or Legitimate), and the columns represent the predicted classes. The four elements of a confusion matrix are:

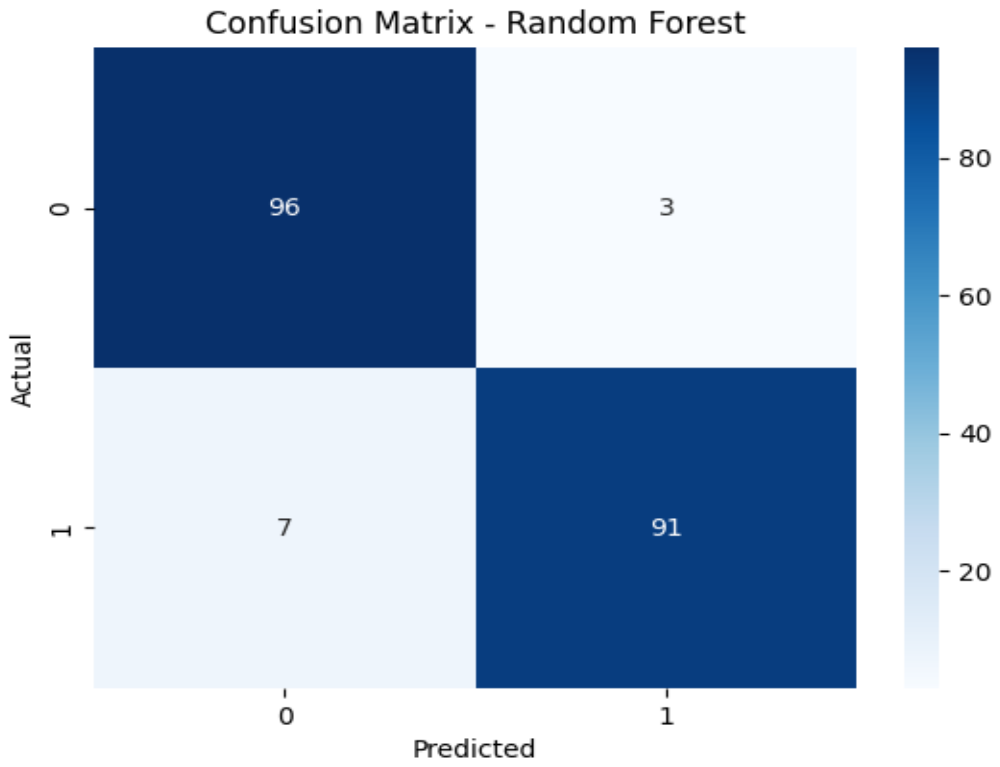
- **True Positive (TP) :** The number of fraudulent transactions correctly identified as fraud.
- **True Negative (TN) :** The number of legitimate transactions correctly identified as legitimate.
- **False Positive (FP) :** The number of legitimate transactions incorrectly identified as fraud (Type I error).
- **False Negative (FN) :** The number of fraudulent transactions incorrectly identified as legitimate (Type II error).

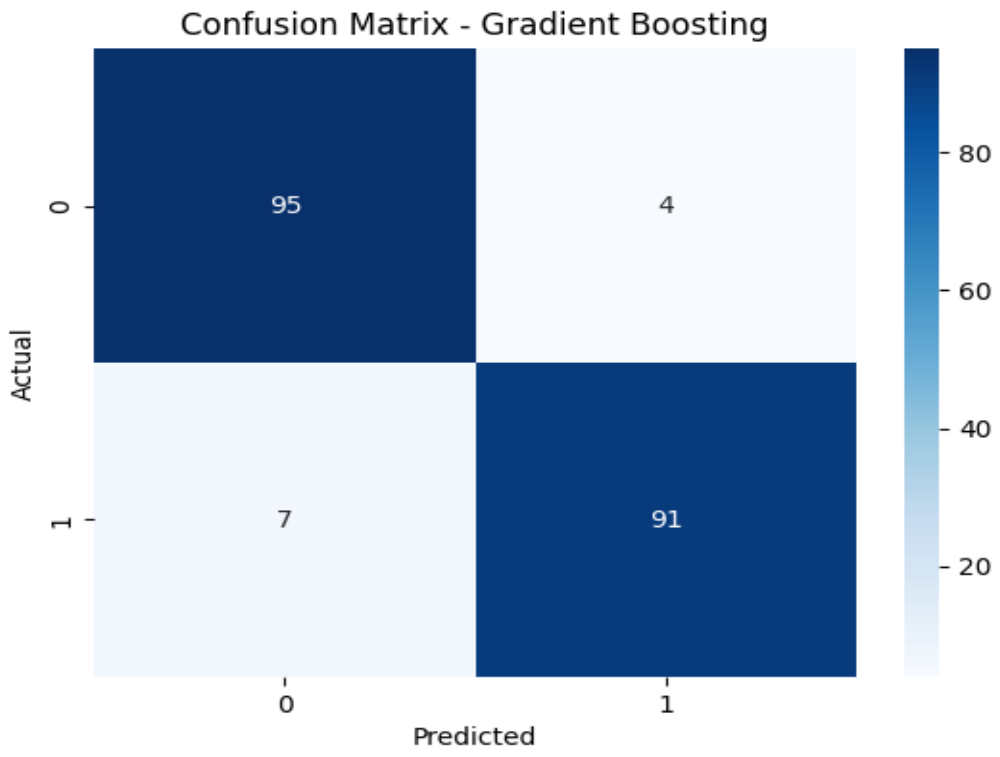
Interpreting the Results

- **True Positives (TP) :** These are the cases where the model correctly identifies fraudulent transactions. Maximizing this value is crucial in fraud detection to catch as much fraud as possible.
- **False Positives (FP) :** If too high, this value indicates that legitimate transactions are frequently flagged as fraud, which could lead to a poor user experience and unnecessary security interventions.
- **False Negatives (FN) :** This is particularly dangerous in fraud detection as it means fraudulent transactions are being missed, allowing potential financial loss.
- **True Negatives (TN) :** These reflect how well the model identifies legitimate transactions, ensuring smooth operations for genuine customers.

Below are the confusion matrices for the three models:




Confusion Matrix - Random Forest
96
3
7
91
0
1
Actual
0
1
Predicted
80
60
40
20


Confusion Matrix - Gradient Boosting
95
4
7
91
0
1
Actual
0
1
Predicted
80
60
40
20

Feature Importance Analysis

Feature importance helps in understanding which features (variables) play a significant role in the decision-making process of a machine learning model. By evaluating the importance of each feature, we can identify the most influential factors contributing to the detection of fraudulent transactions.

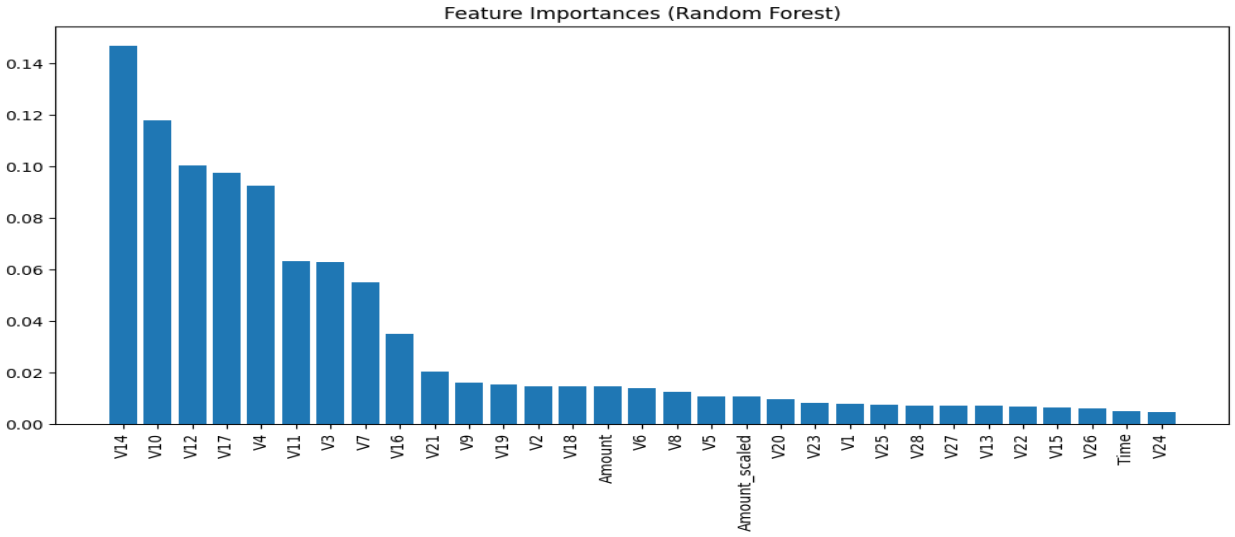
For this project, feature importance was analyzed using Random Forest and Gradient Boosting, two ensemble models that naturally provide this functionality.

Random Forest Feature Importance

Random Forest assigns importance to features based on how much they improve the purity of nodes in the decision trees. The higher the contribution to reducing impurity (e.g., reducing Gini index or increasing information gain), the more important the feature is considered.

The bar plot below shows the feature importance for the Random Forest model, highlighting the most critical features:

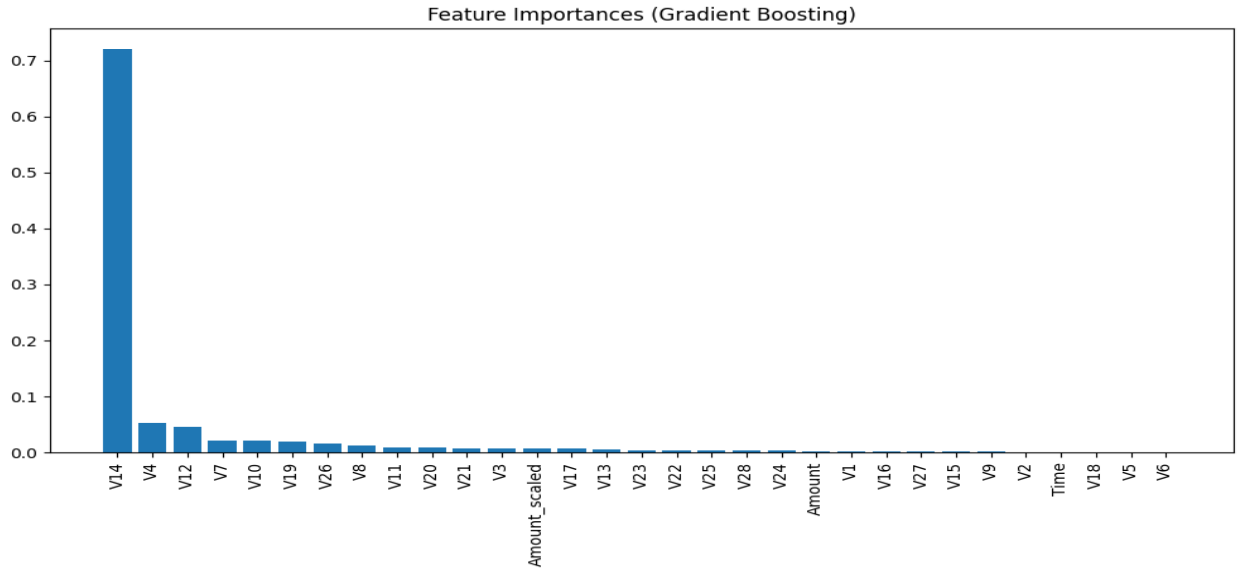
Bar plot of Random Forest Feature Importance



Gradient Boosting Feature Importance

Gradient Boosting focuses on reducing errors in each iteration, and it similarly ranks feature importance based on their contribution to reducing the overall loss (error). The following plot shows the most critical features for the Gradient Boosting model:.

Bar plot of Gradient Boosting Feature Importance



ROC Curve Analysis

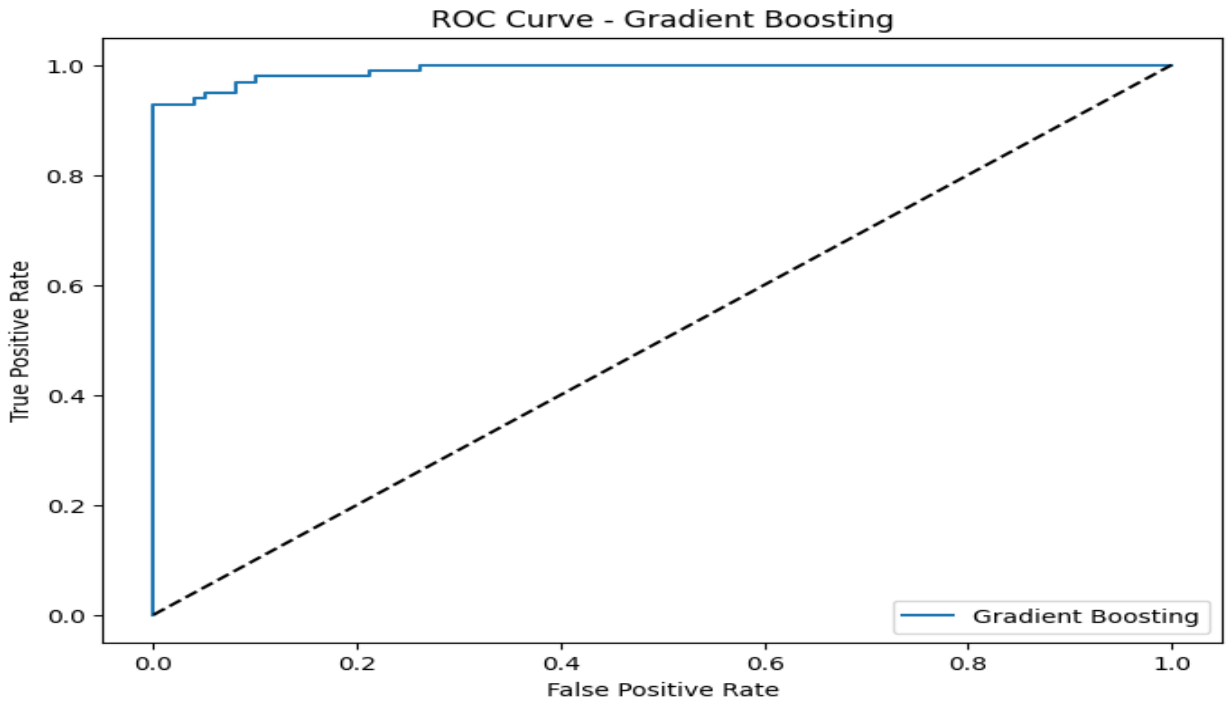
The ROC Curve (Receiver Operating Characteristic Curve) is a graphical representation used to evaluate the performance of a classification model. It plots the True Positive Rate (Sensitivity or Recall) against the False Positive Rate (1 - Specificity) at various threshold settings. The closer the curve is to the top-left corner, the better the model is at distinguishing between classes.

For this project, we evaluated the ROC curves for the Gradient Boosting model, which had the best overall performance.

ROC Curve for Gradient Boosting Model

- X-axis (False Positive Rate): The rate of incorrectly predicted frauds.
- Y-axis (True Positive Rate): The rate of correctly predicted frauds.
- Diagonal Line: Represents random guessing (AUC = 0.5).

Plot of ROC Curve for Gradient Boosting Model



Model Recommendation and Application

Best Model Selection

After evaluating various machine learning models, the Gradient Boosting Classifier emerged as the top performer in this credit card fraud detection project. With an impressive accuracy of 94..42% on the test set, a precision of 95.79%, and an F1 Score of 94.30%, it demonstrates excellent capability in identifying fraudulent transactions while minimizing false positives.

Solving the Problem

The Gradient Boosting Classifier effectively addresses the problem of credit card fraud detection by:

1. **High Accuracy**: Its robust performance ensures that a significant percentage of fraudulent transactions are accurately identified, thus protecting financial institutions and consumers.
2. **Balanced Performance**: With a high recall rate, the model minimizes the risk of missed fraudulent transactions, which is critical in financial operations.

3. **Feature Importance**: The model's interpretation of feature importance allows stakeholders to understand which variables (such as transaction amount and frequency) significantly contribute to fraud detection. This insight can guide further strategies for fraud prevention.

Implementation Strategy

To utilize this model effectively in a real-world scenario, consider the following steps:

1. **Integration**: Implement the model within the existing transaction processing system of financial institutions. This would enable real-time fraud detection and alerting mechanisms.
2. **Continuous Monitoring**: Regularly update and retrain the model with new transaction data to ensure its effectiveness, especially as fraud patterns evolve.
3. **Collaboration with Fraud Analysts**: Provide insights generated by the model to fraud detection teams. Analysts can further investigate flagged transactions and take appropriate action.
4. **Feedback Loop**: Create a feedback mechanism to continuously improve the model based on new fraudulent activity detected and feedback from analysts.

Conclusion

By deploying the Gradient Boosting Classifier, organizations can significantly enhance their fraud detection capabilities, thereby protecting their assets and customers. This model serves not only as a reactive measure but also as a proactive strategy in the ongoing fight against credit card fraud.