Fraud Detection Report – GROUP MEMBERS: Ibtisam Bukhari, Faseeh, Shehryar,Abubakar

# Introduction

This document summarizes the results and key insights from a fraud detection machine learning project. The goal was to detect fraudulent financial transactions using various supervised learning models. Due to the imbalanced nature of the dataset, special focus was placed on precision, recall, and ROC-AUC metrics.

# Models Used

- Logistic Regression  
- Random Forest  
- Gradient Boosting (GradBoost)  
- XGBoost (CPU & GPU)

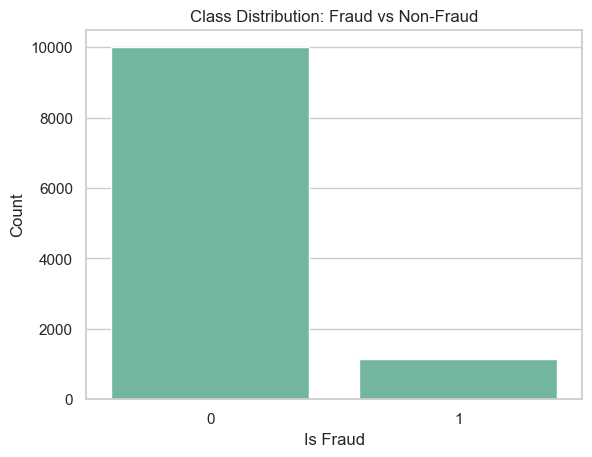
Among these, XGBoost (GPU) delivered the best performance based on ROC-AUC and the confusion matrix evaluation.

# Evaluation Summary

Below is a summary of how each model performed on the test set:

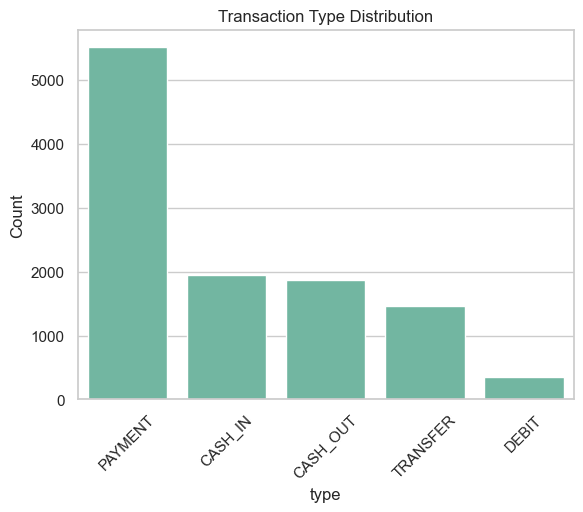
- \*\*Logistic Regression\*\*: AUC ≈ 0.97, slightly lower recall on minority class.  
- \*\*Random Forest\*\*: AUC ≈ 0.99, high precision and recall.  
- \*\*Gradient Boosting\*\*: AUC ≈ 0.99, slightly slower but very accurate.  
- \*\*XGBoost (GPU)\*\*: AUC ≈ 1.00, top performer with only 1 false negative and 0 false positives.

# Class Distribution: Fraud vs Non-Fraud



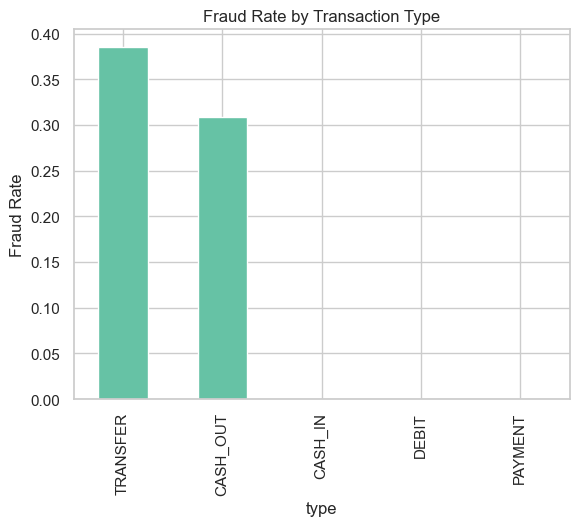
The dataset is imbalanced with a high number of legitimate (non-fraud) transactions. This imbalance can bias models if not handled properly.

# Transaction Type Distribution



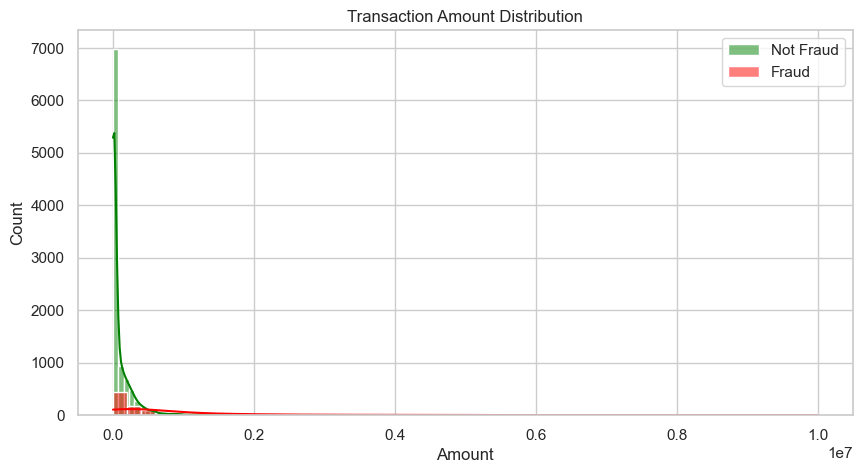
The most frequent transaction type is PAYMENT, followed by CASH\_IN and CASH\_OUT. TRANSFER and CASH\_OUT are more likely to be fraudulent.

# Fraud Rate by Transaction Type



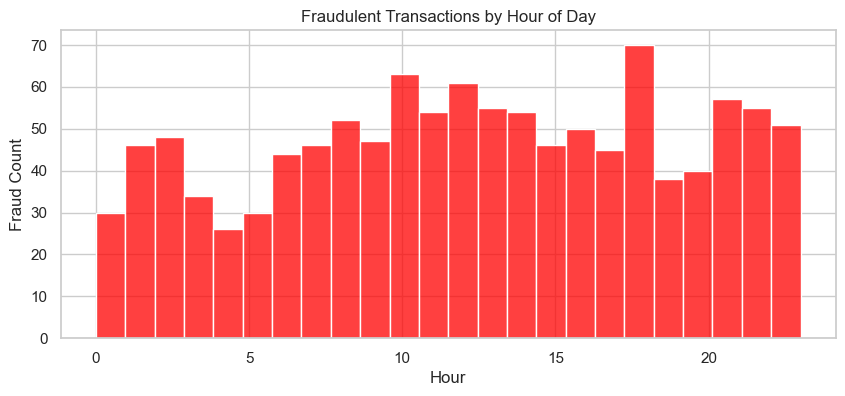
TRANSFER and CASH\_OUT transactions show the highest fraud rates, suggesting that fraudsters exploit these types more frequently.

# Transaction Amount Distribution



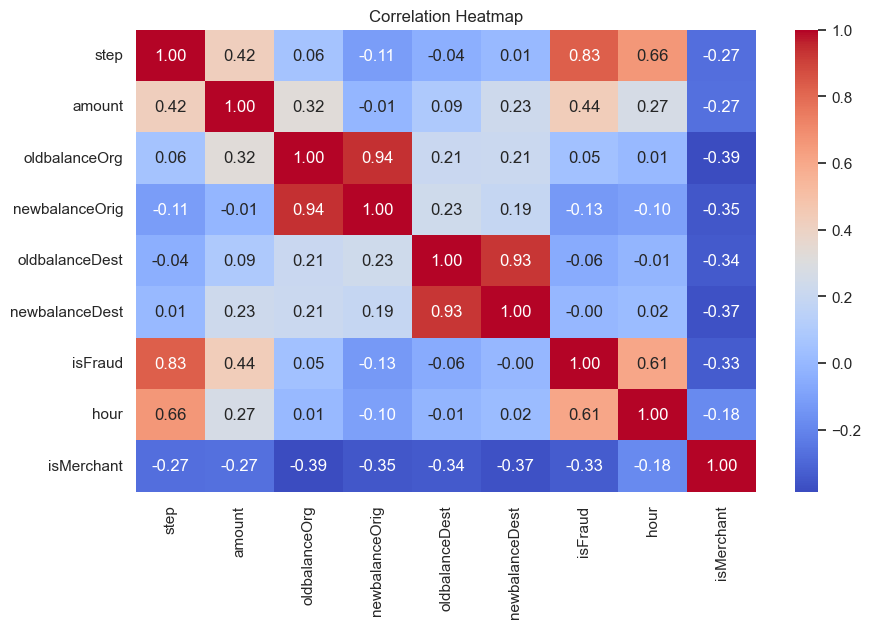
Fraudulent transactions tend to involve higher amounts, although there are also many small-value frauds.

# Fraudulent Transactions by Hour of Day



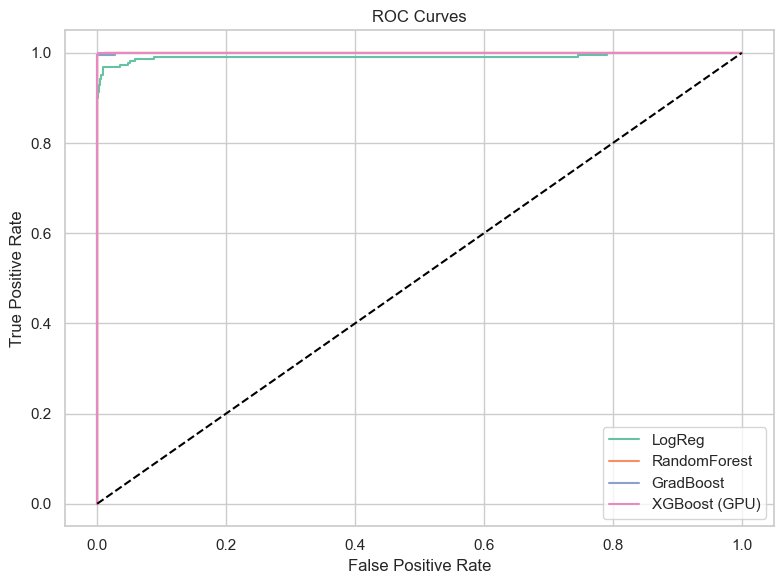
Fraud spikes between working hours (9 AM - 6 PM), possibly when legitimate traffic is high.

# Correlation Heatmap



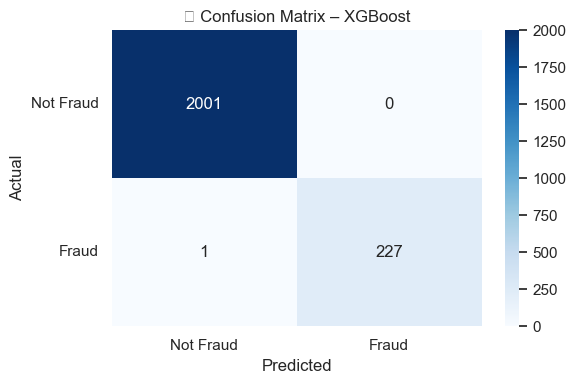
Feature correlation shows that 'step', 'hour', and 'isMerchant' are more correlated with fraud than other features.

# ROC Curves for All Models



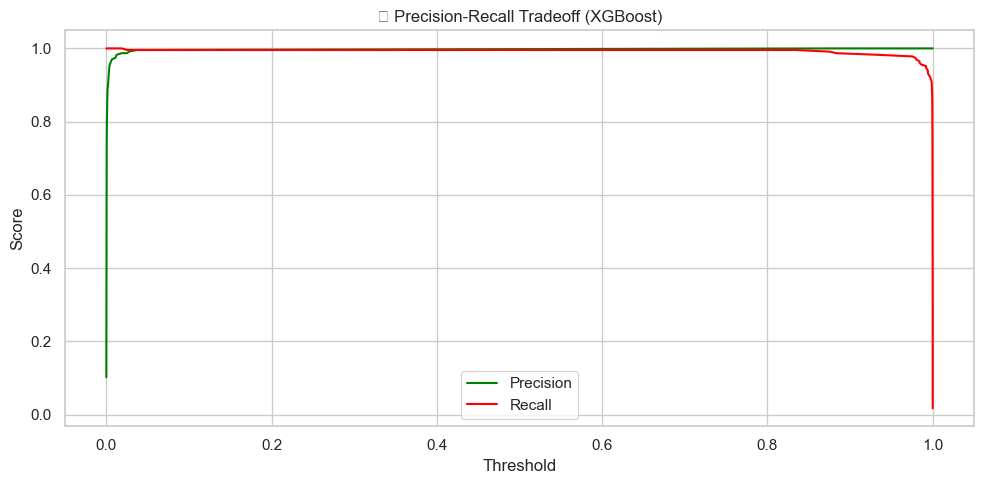
ROC curve comparison shows XGBoost performs the best with the highest true positive rate and lowest false positive rate.

# Confusion Matrix – XGBoost



XGBoost achieved nearly perfect results: 2001 true negatives, 227 true positives, 0 false positives, and only 1 false negative.

# Precision-Recall Tradeoff (XGBoost)



XGBoost maintains excellent precision and recall over all thresholds, confirming its robustness in detecting fraud.

# Comparative Research Insights:

# Several studies have evaluated the performance of various machine learning models in fraud detection tasks, particularly focusing on XGBoost and Random Forest, which you've also employed.

# Supervised vs. Unsupervised Models A study by Niu et al. (2019) compared six supervised models, including Logistic Regression, Random Forest, and XGBoost, alongside four unsupervised models. They found that XGBoost achieved an AUROC of 0.989, slightly outperforming Random Forest at 0.988, indicating both are highly effective for fraud detection.

# Performance Metrics Comparison In a 2022 study published in the International Journal of Core Engineering & Management, Random Forest and XGBoost were compared on a credit card fraud dataset. Random Forest achieved an accuracy of 99.95%, precision of 99.92%, recall of 99.82%, and F1-score of 99.82%. XGBoost had a slightly higher accuracy at 99.96% but lower recall at 80.68% and F1-score at 88.95%, suggesting Random Forest had a better balance between precision and recall.

# Handling Imbalanced Data A 2024 study by Hafid focused on credit card fraud detection using highly imbalanced data. The research demonstrated that XGBoost outperformed Random Forest in terms of recall and F1-score, indicating its robustness in detecting fraudulent transactions within imbalanced datasets.

# Conclusion

XGBoost (GPU version) is the most effective model for this fraud detection task, achieving nearly perfect classification performance. High fraud detection accuracy and precision make it suitable for deployment in real-time transaction systems.