



# Credit Card Fraud Detection & Prevention

Developers Institute

Rodgers Odondi



# The Business Problem

- Credit card fraud leads to billions in global losses.
- Fraudulent transactions are few making them hard to detect without many false positives.
- Businesses need accurate, real-time detection.

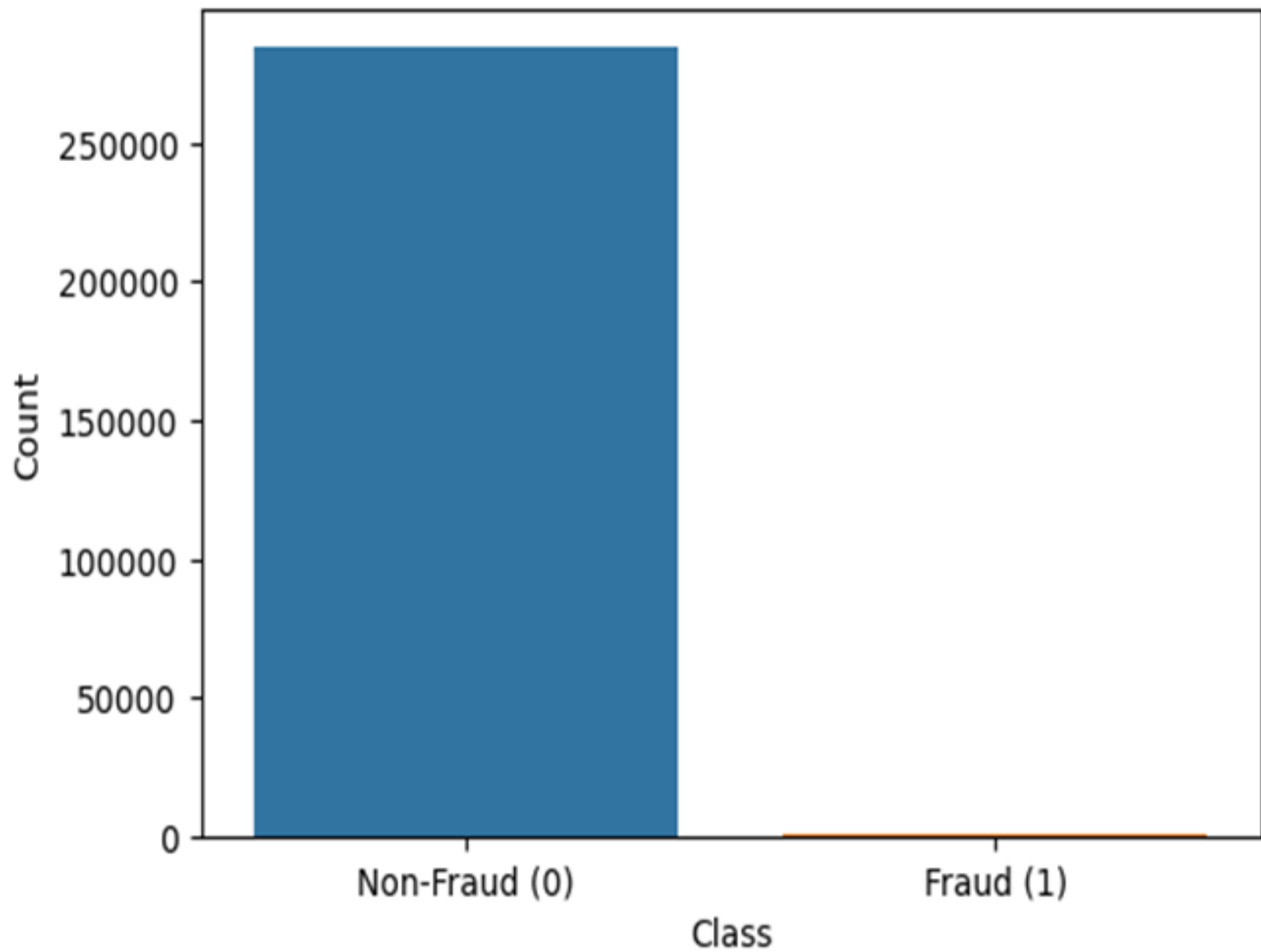
# Dataset Overview

- Source: Kaggle – Credit Card Fraud Detection Dataset
- Rows: 284,807 transactions
- Features: V1 to V28 (anonymized via PCA), Amount, Time, and Class (label)

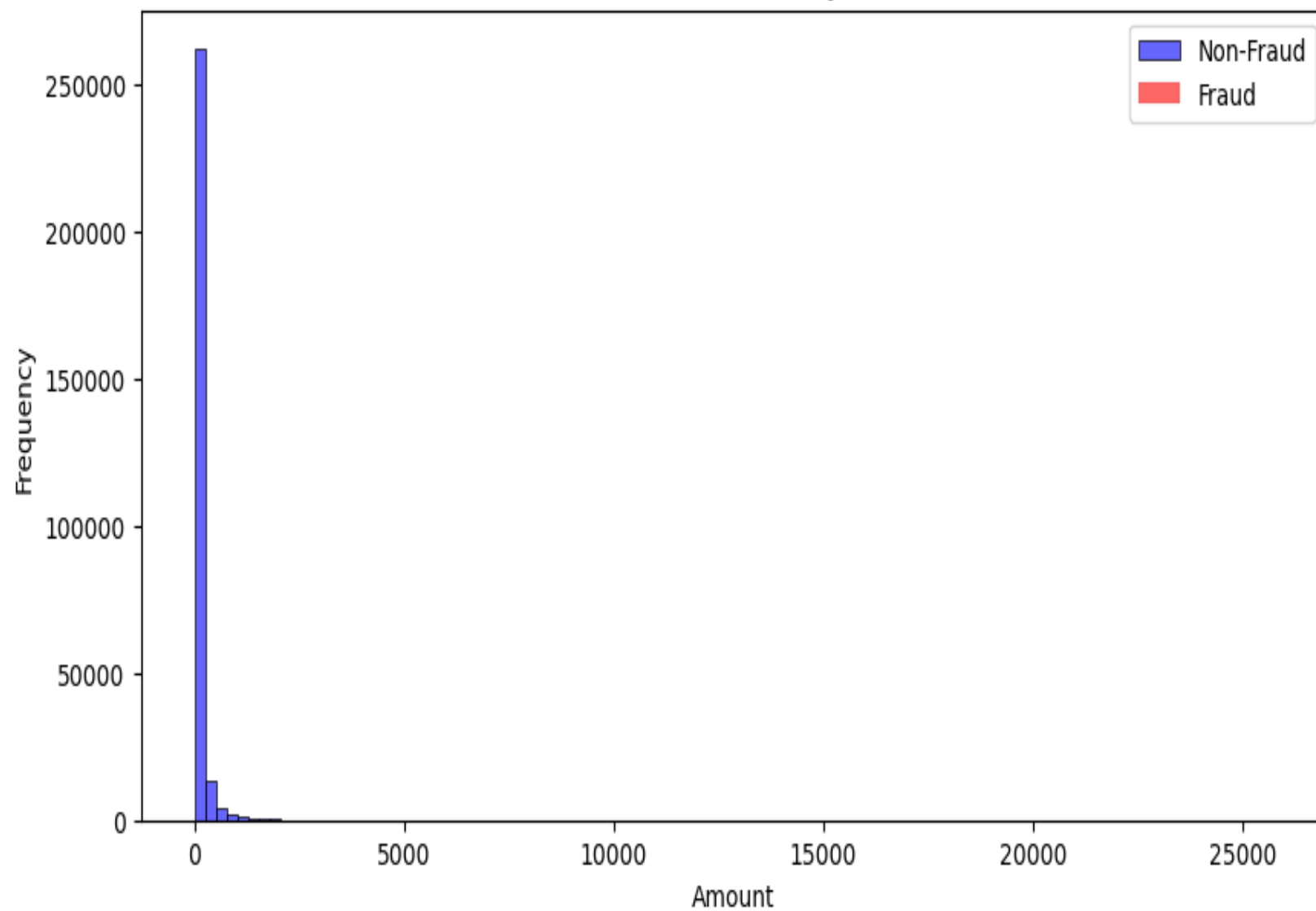
# Exploratory Data Analysis (EDA)

- Strong class imbalance: ~99.83% non-fraud, ~0.17% fraud
- Fraudulent transactions tend to have lower amounts
- Some features (e.g., V14, V10) show visible differences by class

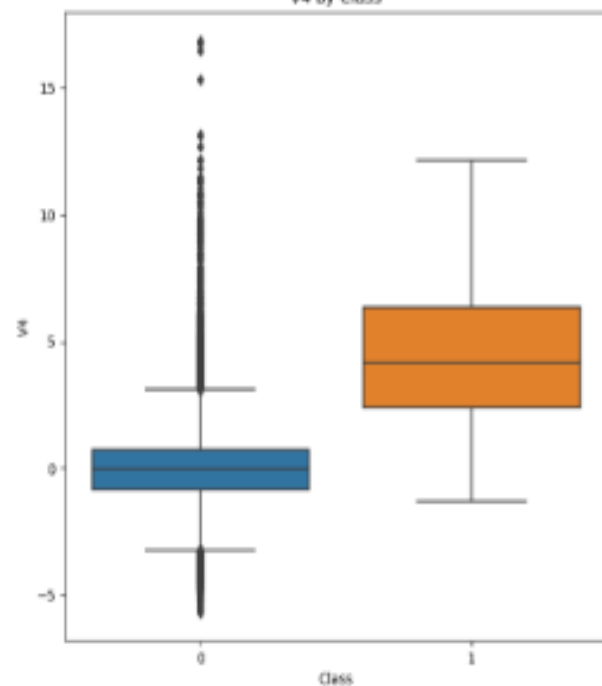
Class Distribution



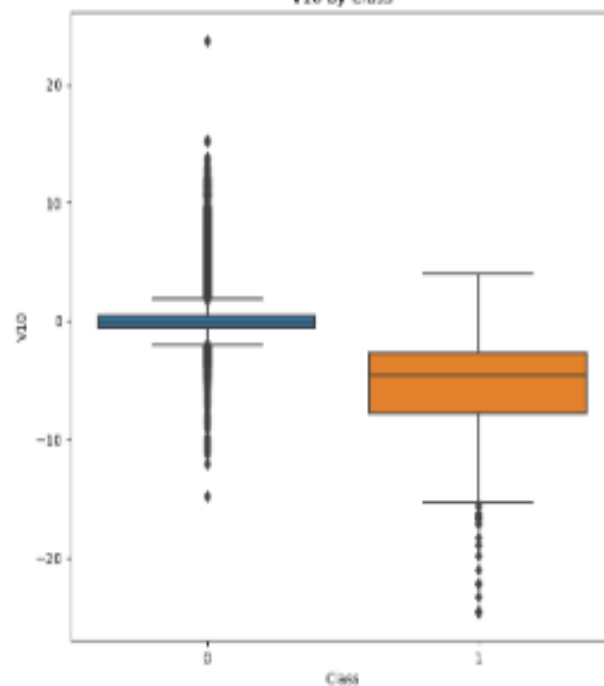
Transaction Amount by Class



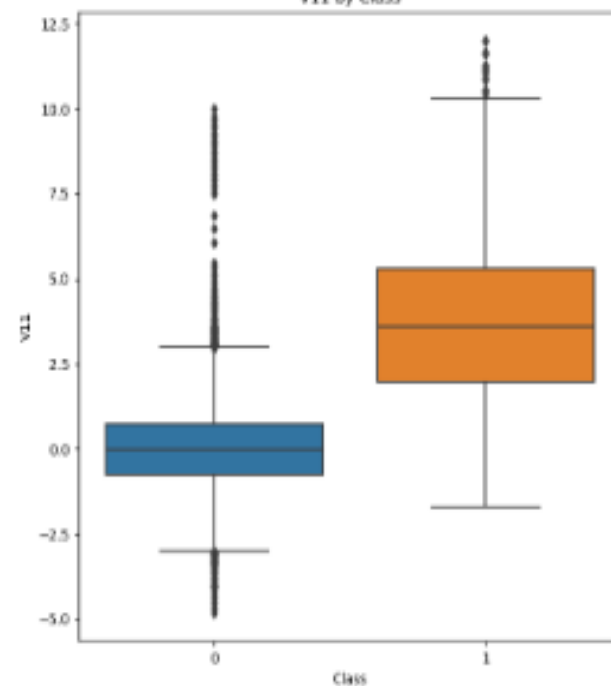
V4 by Class



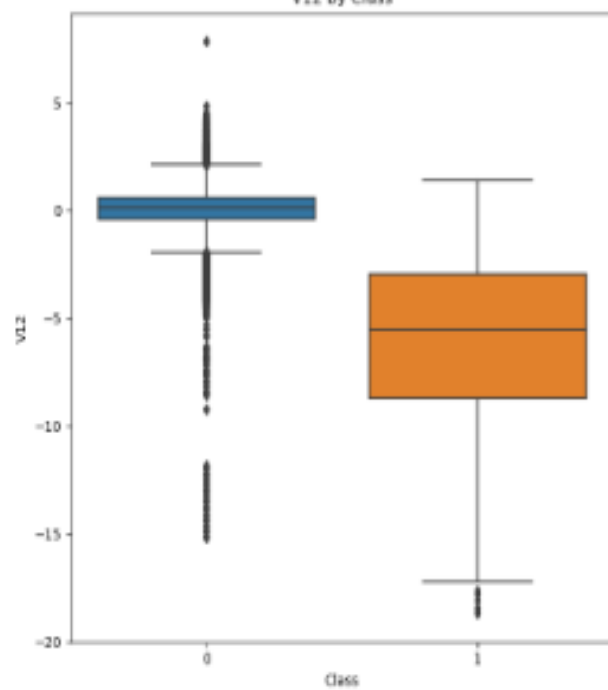
V10 by Class



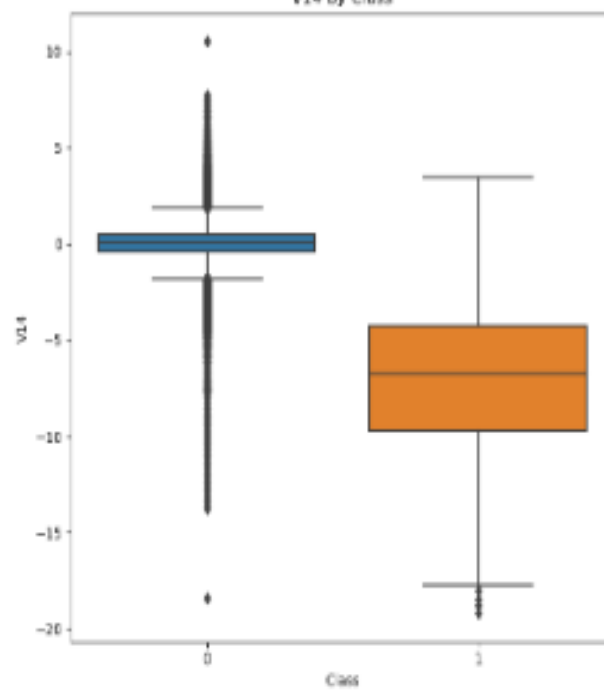
V11 by Class



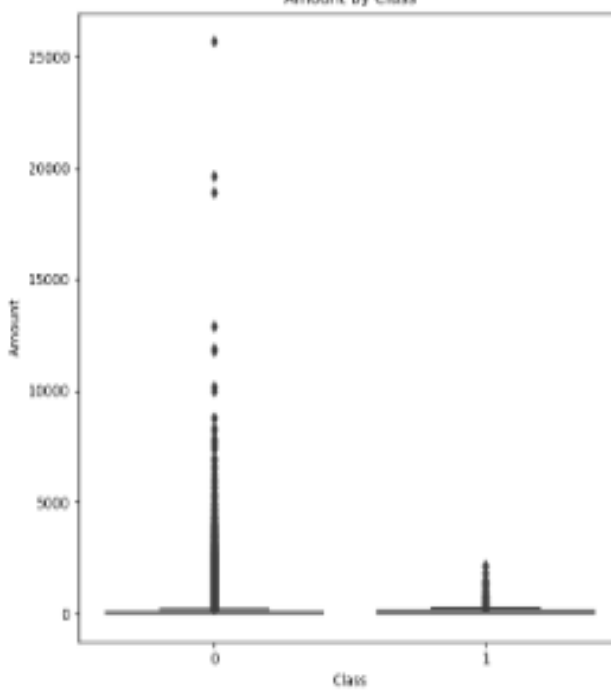
V12 by Class



V14 by Class



Amount by Class





# Our Solution

- Scaled 'Amount' and 'Time' using StandardScaler
- Applied SMOTE to balance classes in the training set.
- Split the dataset into 70% training and 30% testing
- Trained and evaluated Logistic Regression, Random Forest, XGBoost.



# Model Evaluation

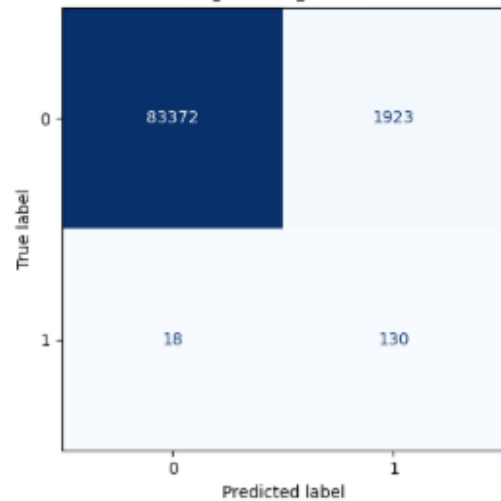
Precision, Recall, and F1-Score Comparison

Model	Precision, Recall, and F1-Score Comparison		
	Precision	Recall	F1-Score
Logistic Regression	0.0633	0.8784	0.1181
Random Forest	0.8519	0.7770	0.8127
XGBoost	0.7580	0.8041	0.7803

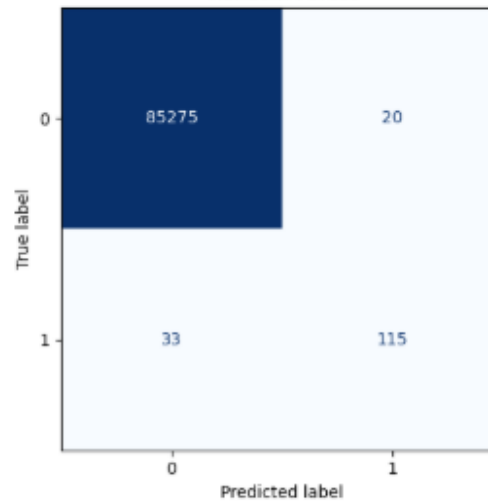
# Model Evaluation

Confusion Matrices: Logistic Regression vs Random Forest vs XGBoost

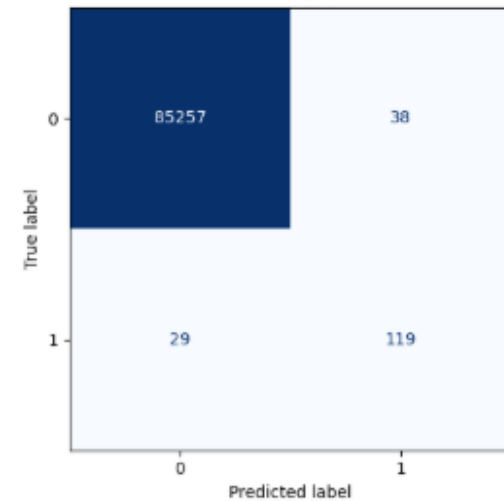
Logistic Regression



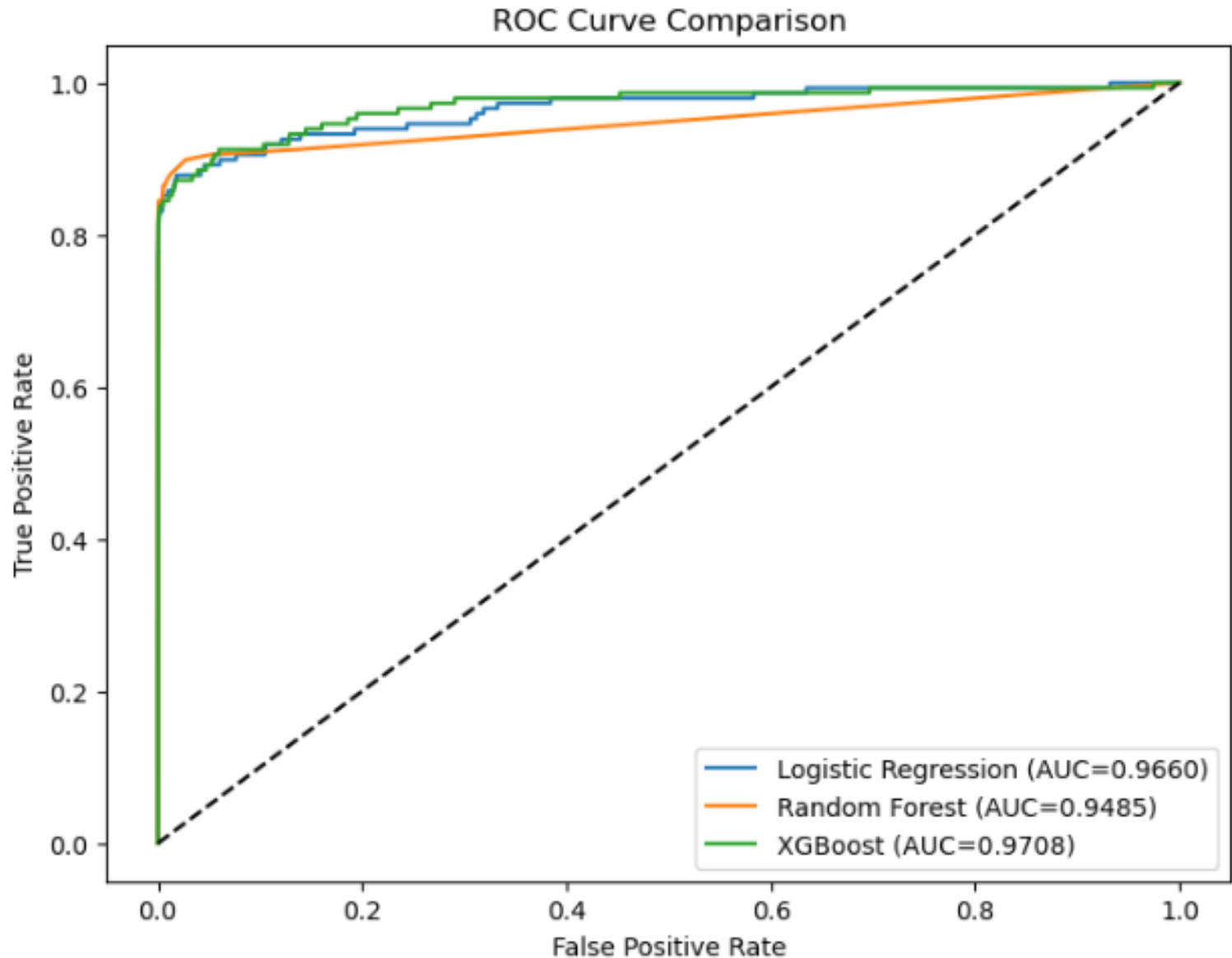
Random Forest



XGBoost



# Model Evaluation



# Model Evaluation

- XGBoost performed best in terms of AUC(~ 0.9708) and precision-recall.
- It handled the complex patterns in the data better, especially after class balancing with SMOTE.
- Confusion matrix shows low false positive rate



# Business Impact

- Reduces fraud-related losses and chargebacks.
- Improves customer trust and retention.
- Enables faster response and fewer false declines.
- Scalable model ready for real-time deployment.



# Fraud Prevention Strategy

- Deploy XGBoost model for real-time transaction scoring
- Use dynamic thresholds to flag or block high-risk events
- Integrate rule-based system for behavioral alerts
- Add new data sources: geo-location, device ID, merchant info
- Use dashboards to monitor fraud spikes and transaction heatmaps