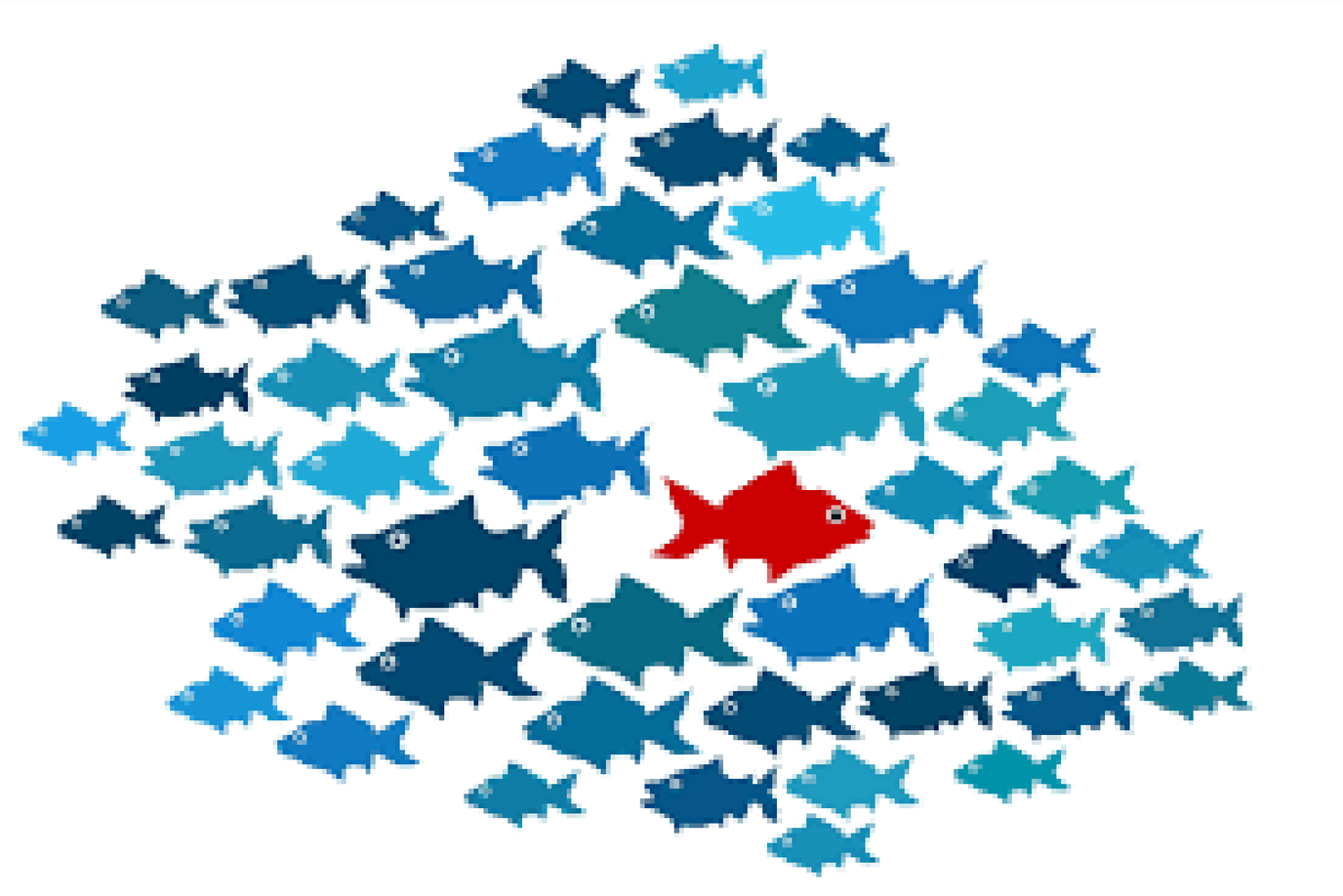

# Credit Card Fraud Detection

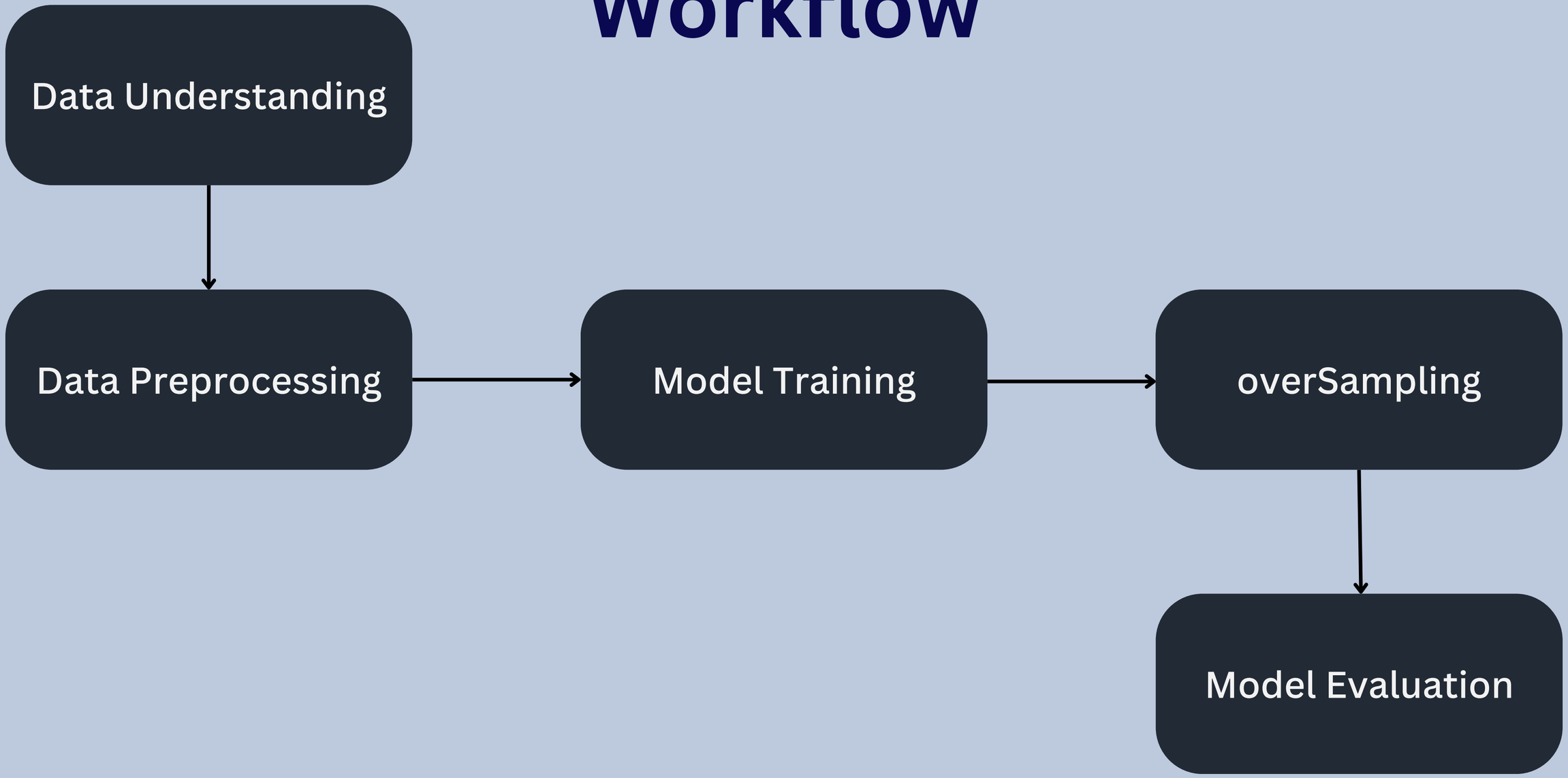
Named:  Passant Adel Farouq

# Analyzing Fraudulent Activities Using Anomaly Detection



# Business Understanding

# Workflow


Data Understanding
Data Preprocessing
Model Training
overSampling
Model Evaluation

# Data Understanding

## Dataset Overview:

- Number of records: **284 807**
- Features: **31**
- Class Distribution:

  Non-Fraud - (284315) **99.8%**

  Fraud - (492) **0.17%**

# Methodology

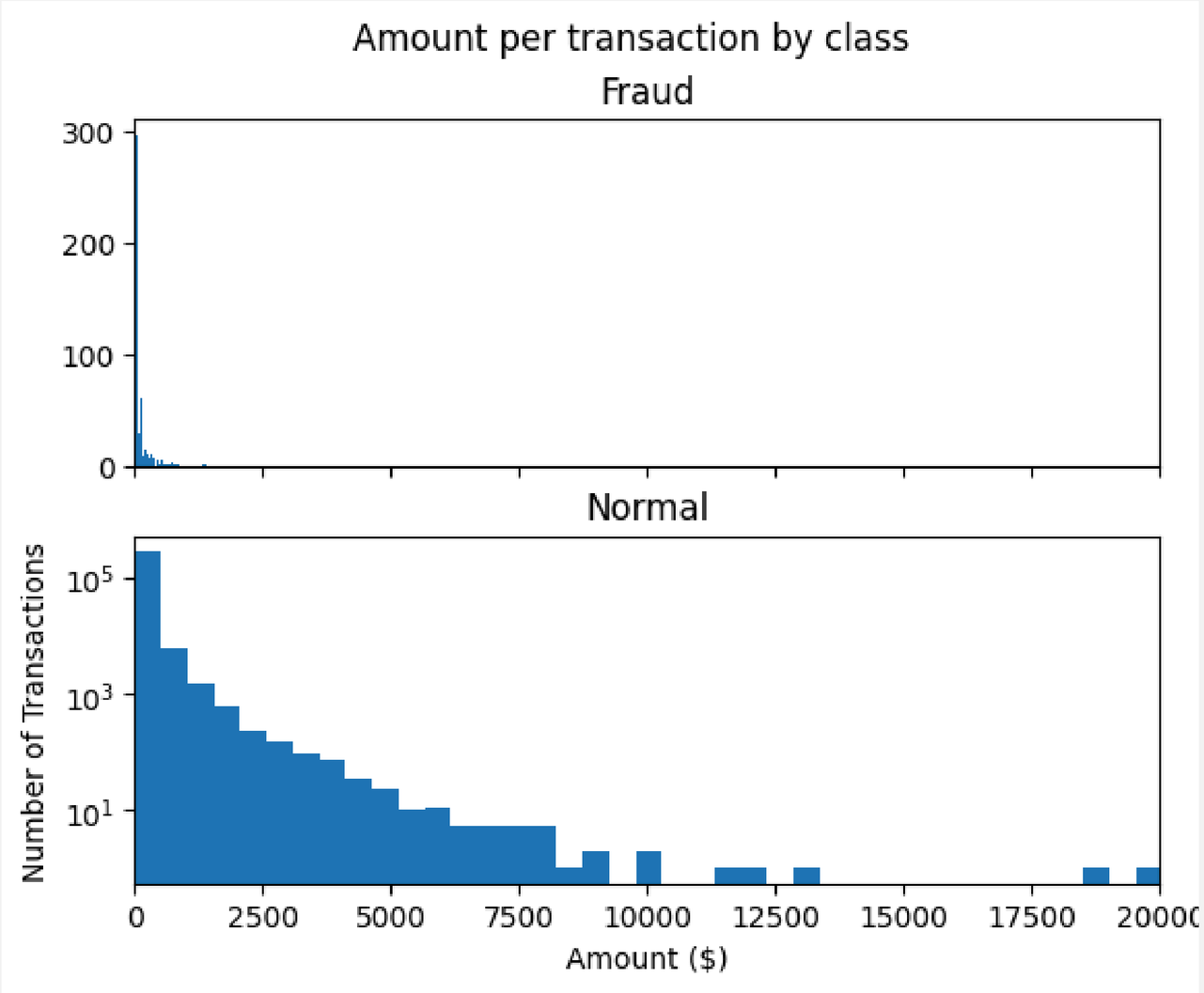
- **Standard scaling for 'Amount' and 'Time'.**

- **Models Employed:**
  Decision Tree.
  Random Forest.
  Logistic Regression.

- **Sampling Technique:**
  SMOTE for handling class imbalance.

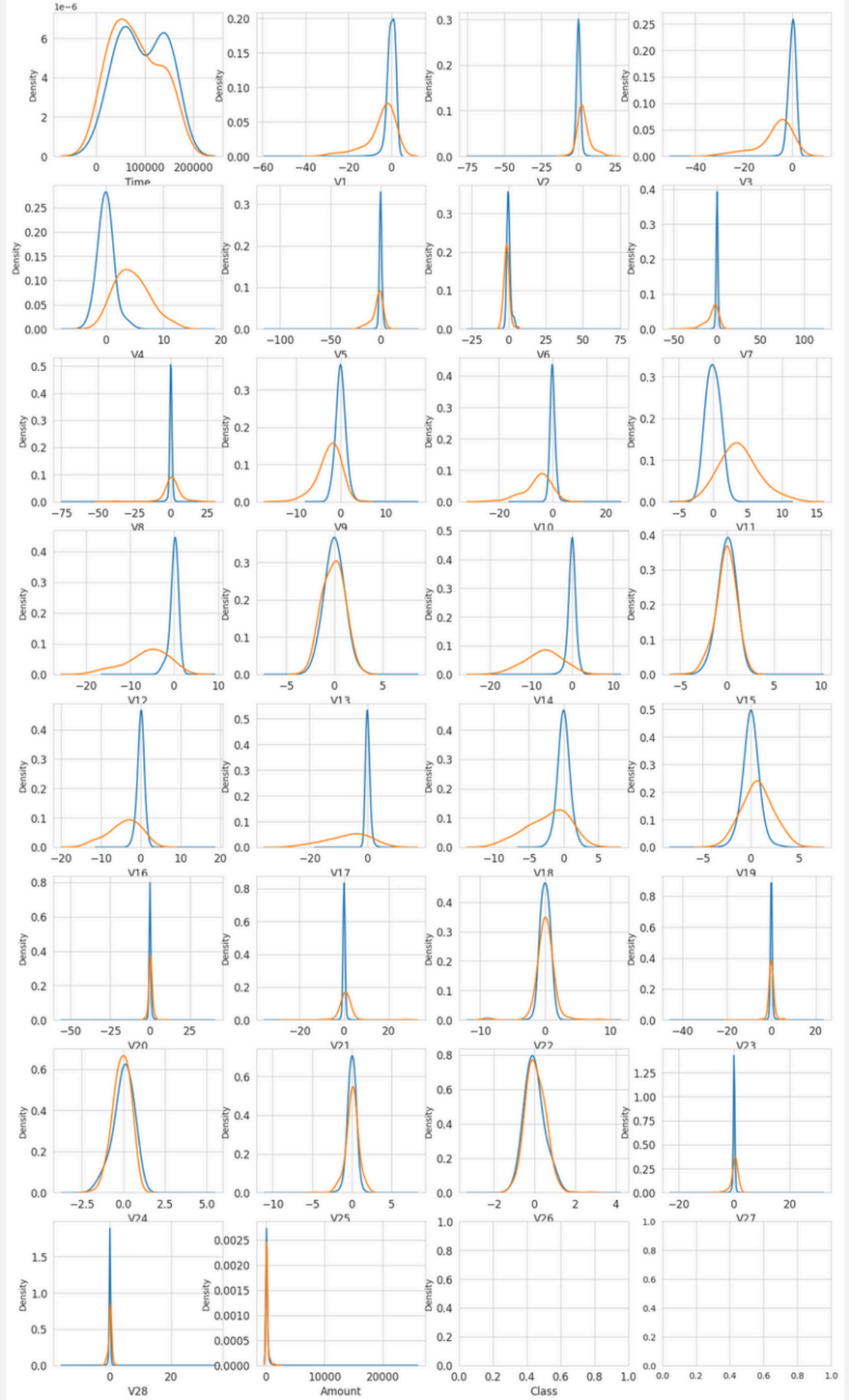
# EDA Charts

Transaction Amount Distribution:



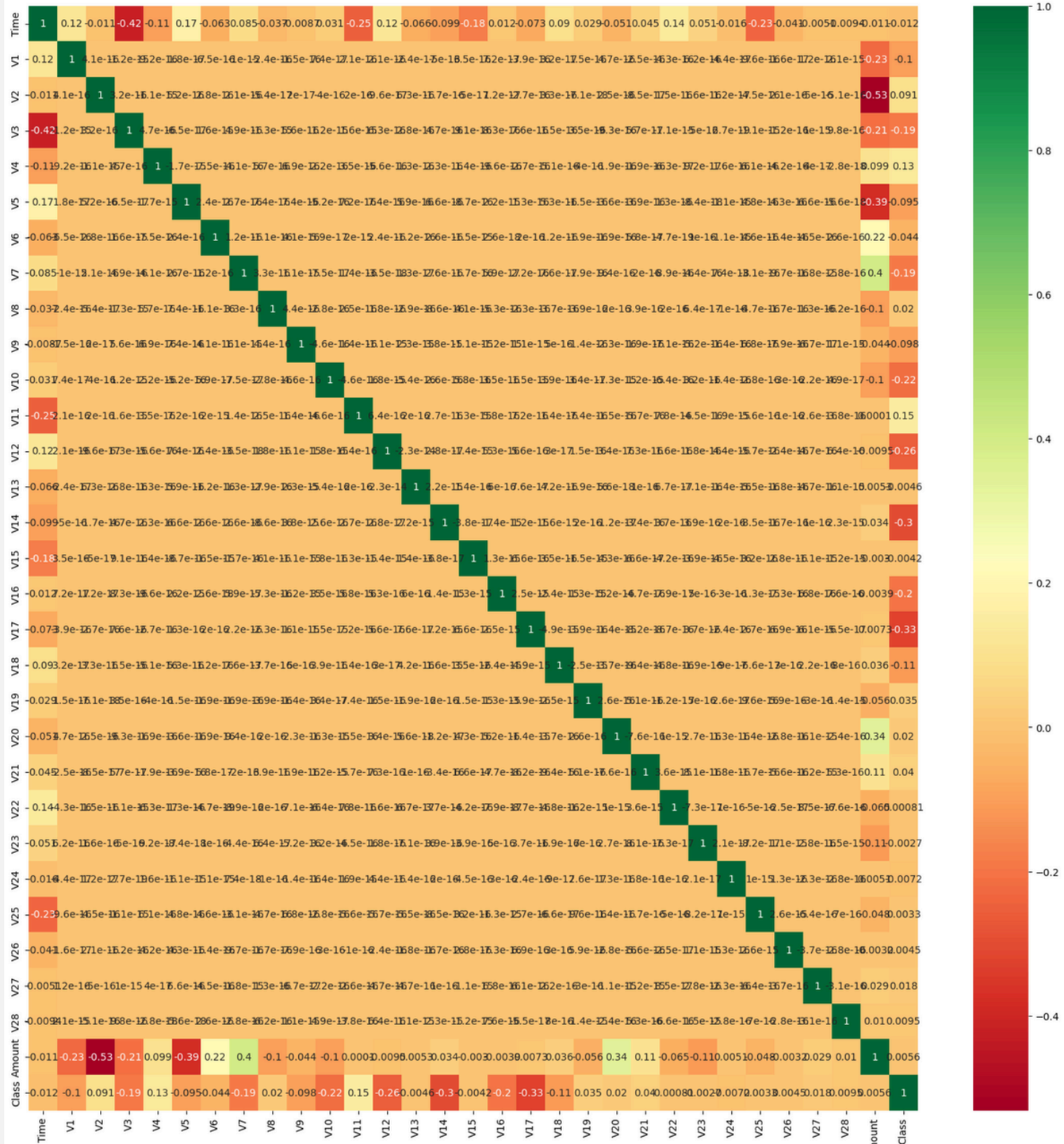
# EDA Charts

Normal vs Fraud Distribution



# EDA Charts

Correlation Matrix

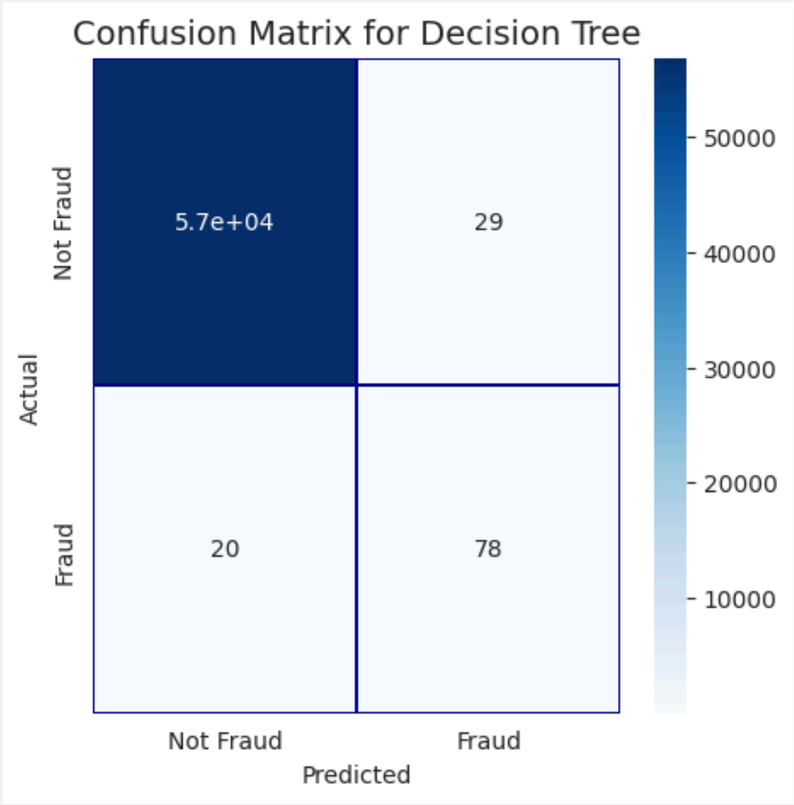


# Model Evaluation
# Decision Tree

- Evaluation Metrics:
  - Accuracy: 99.9%
  - Precision: 68.1%
  - Recall: 79.5%
  - F1 Score:76%

  - **ROC AUC:89.77%**

- Confusion Matrix:

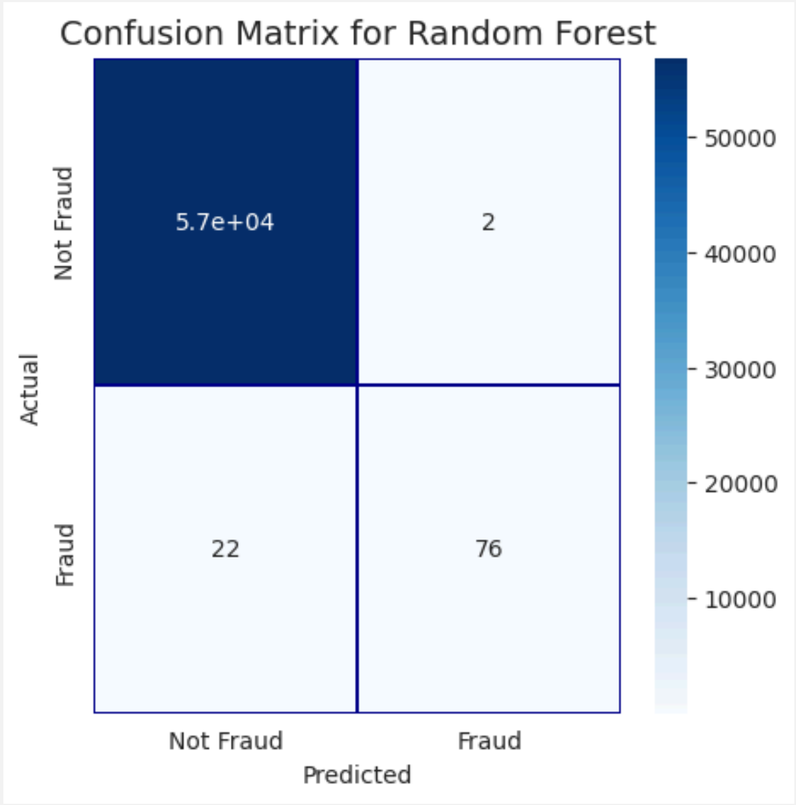


# Model Evaluation
# Random Forest

- Evaluation Metrics:
  - Accuracy: 99.9%
  - Precision: 97.4%
  - Recall: 77.5%
  - F1 Score:86.3%

  - **ROC AUC:88.7%**

- Confusion Matrix:

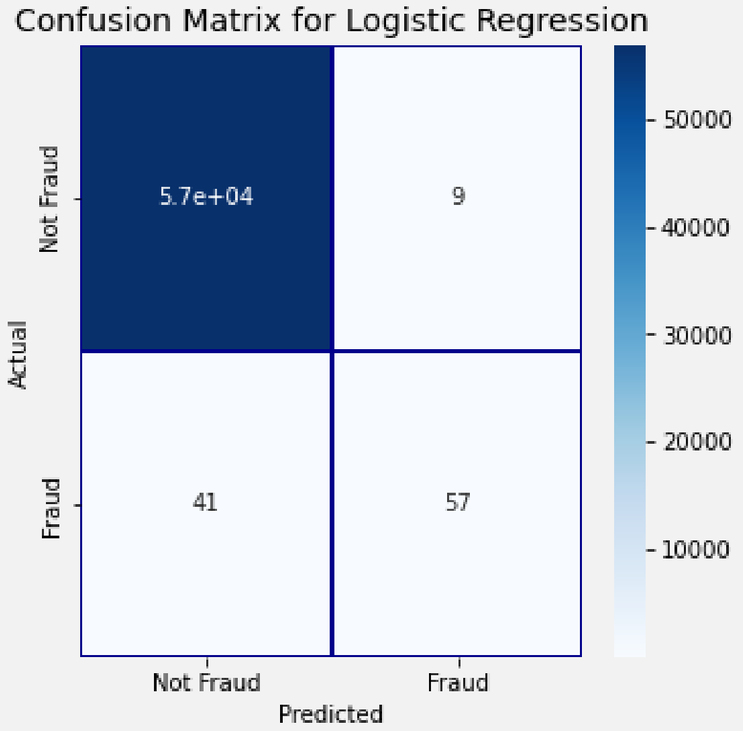


# Model Evaluation
# Logistic Regression

- Evaluation Metrics:
  - Accuracy: 99.9%
  - Precision: 86%
  - Recall: 58%
  - F1 Score:69.5%

  - **ROC AUC:79%**

- Confusion Matrix:



# Oversampling

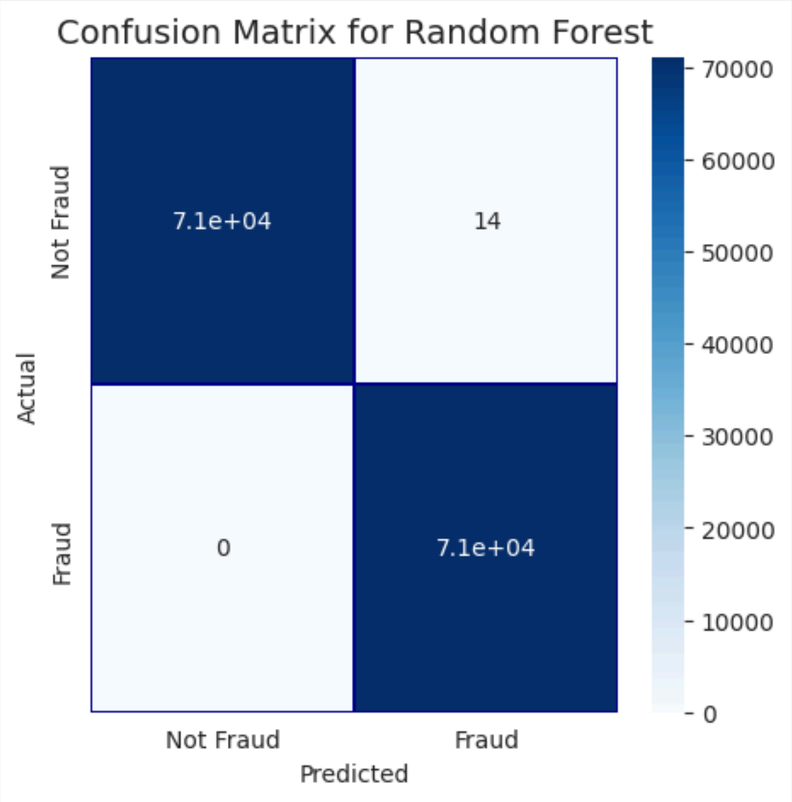
- Oversample the minority class
- The Synthetic Minority Oversampling Technique (SMOTE)
- Data augmentation for the minority class.
- Resampled  our data

# Model Selection
# Random Forest

- Evaluation Metrics:
  - Accuracy: 99.9%
  - Precision: 99.9%
  - Recall: 100%
  - F1 Score:99.9%

  - **ROC AUC:99.99%**

- Confusion Matrix:



Fixed Solutions