Financial Risk Prediction Using DE

Team Name: BruteForce

Team Members: Khushi Saxena
Swayamshree Nanda
Abhishek Raj
Pranoy Jana
Ashish Anand

Introduction

Problem Statement:

- Financial institutions face challenges in managing risks across lending, investments, and fraud detection.
- Predicting financial risks using advanced data engineering is essential for proactive decision-making.

Challenges:

- Design a scalable data pipeline that ingests, processes, and analyzes financial data to predict risk levels.
- Choosing the right model for fraud detection is a critical challenge due to the nature of fraud detection tasks, which often involve **Imbalanced data** & **Non-linear patterns** in fraudulent behavior.

Goal:

- Provide actionable insights that help institutions make smarter, real-time decisions.
- Fraud Detection in Credit Card Transactions Using different Machine Learning Models.

Google   **https://colab.research.google.com/drive/1zb1Uyzwve9WDry_C9ufPoJMfy**

Collab Link: **IkGczko?usp=drive-dynamite#scrollTo=BvU_VwhHvO1J**
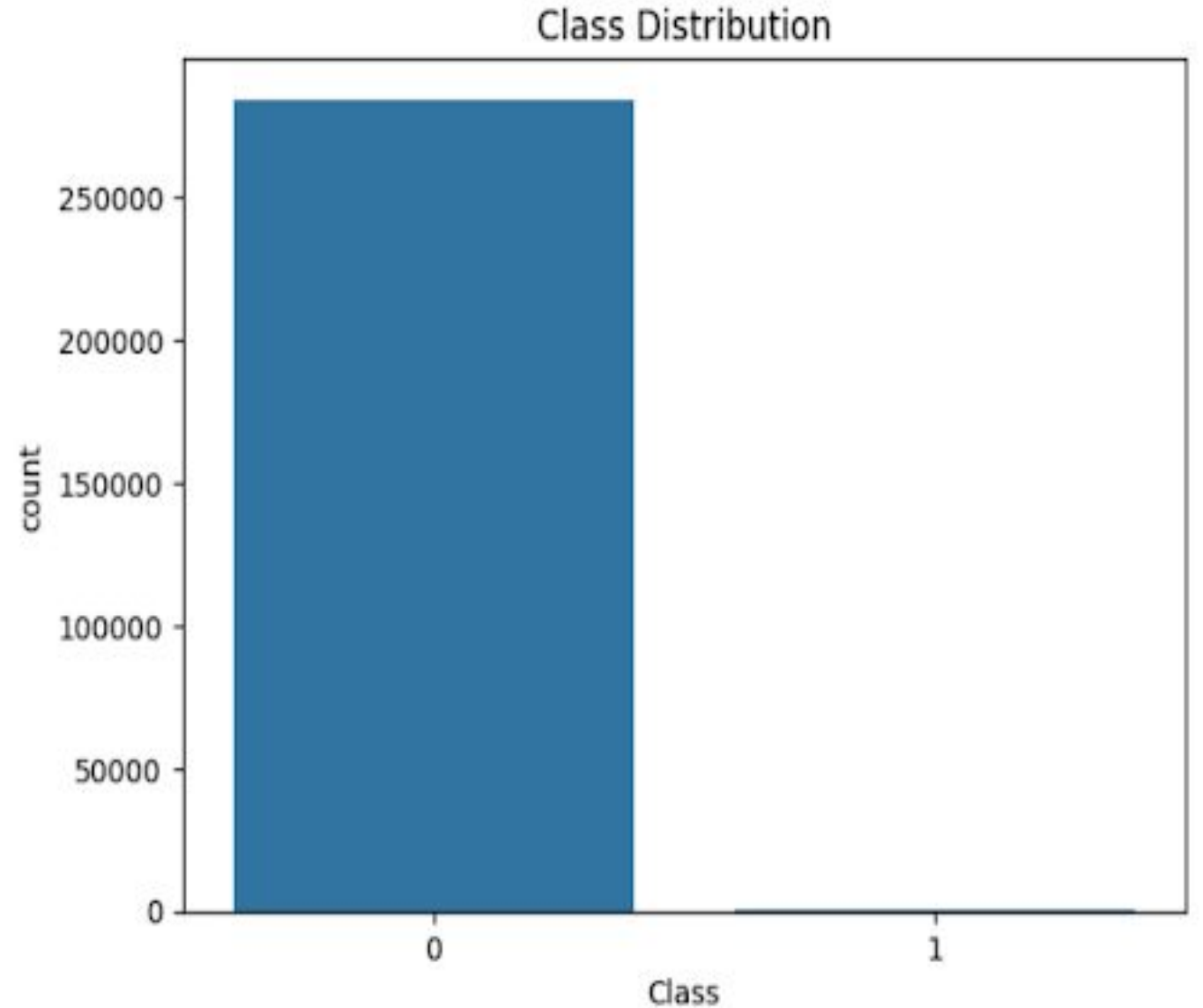
SIGMOID

Methodology

Data Preparation & Analysis

Import Libraries & Load Dataset

- NumPy, Pandas, Matplotlib, Scikit-learn, SciPy
- Import financial time-series data and indicators

Data Exploration

- Checking for missing values
- Data Visualisation

Class Distribution

Methodology

Feature Engineering & Preprocessing

Financial Feature Engineering

- Feature correlation
- Reduce dimensionality using PCA

Data Preprocessing

- Normalization of financial indicators
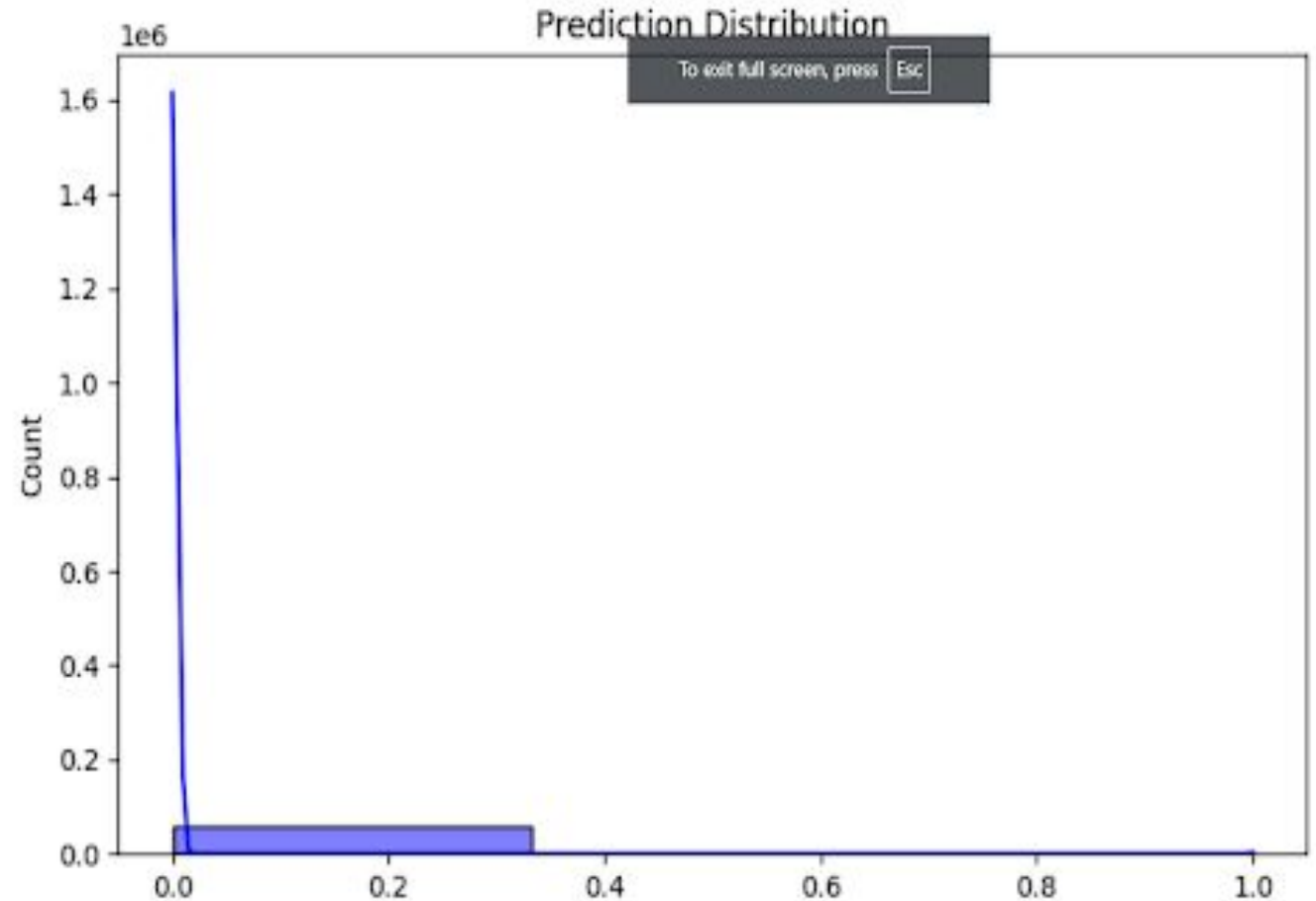- Handle class imbalance in risk events using SMOTE

Feature Correlation Matrix

PCA Visualization
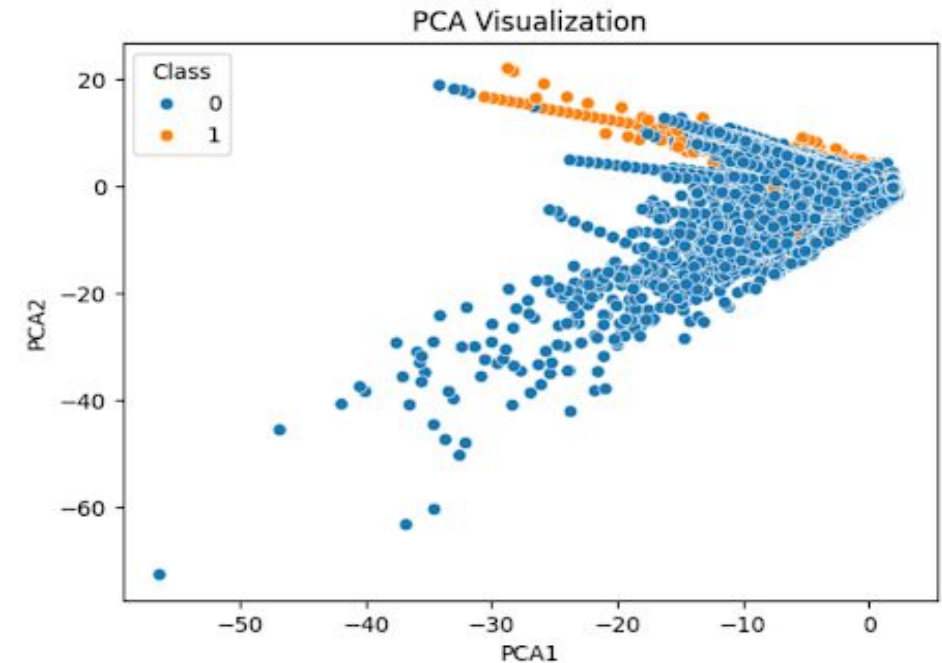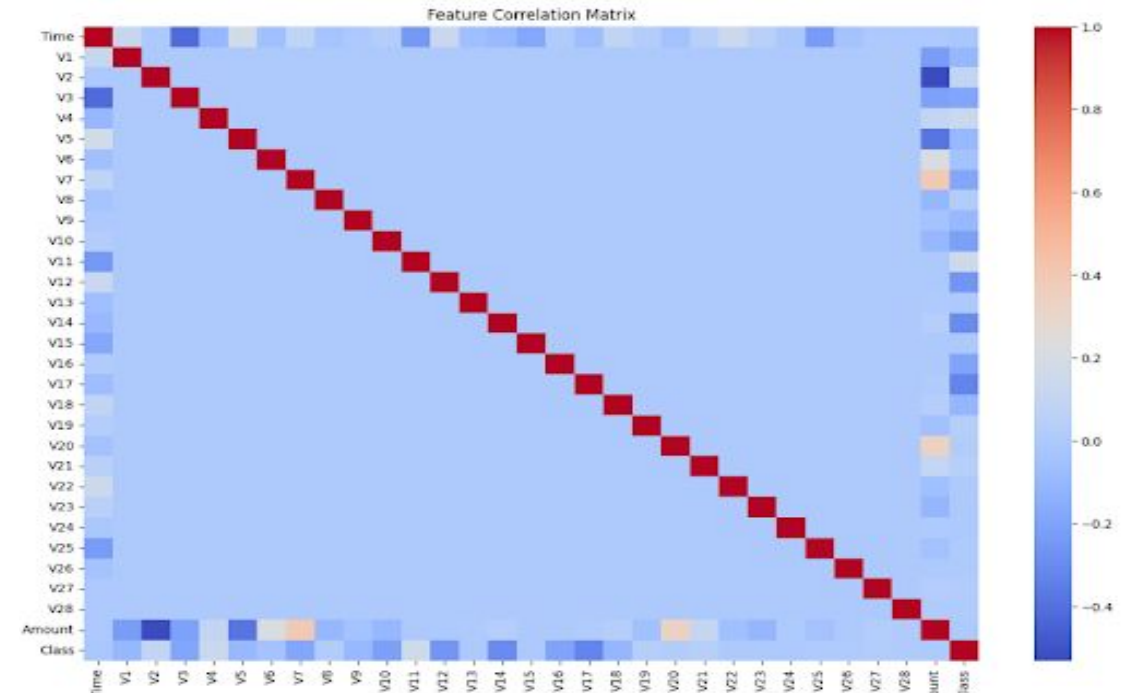
Methodology

## Model Development & Evaluation

Differential Evolution Implementation

- Model training with financial risk metrics
- Model Deployment
- Cross-validation with time-based splits
- Performance evaluation

Prediction Distribution

Methodology

Deployment & Insights

- Model backtesting and stress scenarios
- Risk score visualization and monitoring
- Confusion matrix visualization
- Feature Importance

Feature Importance

Confusion Matrix
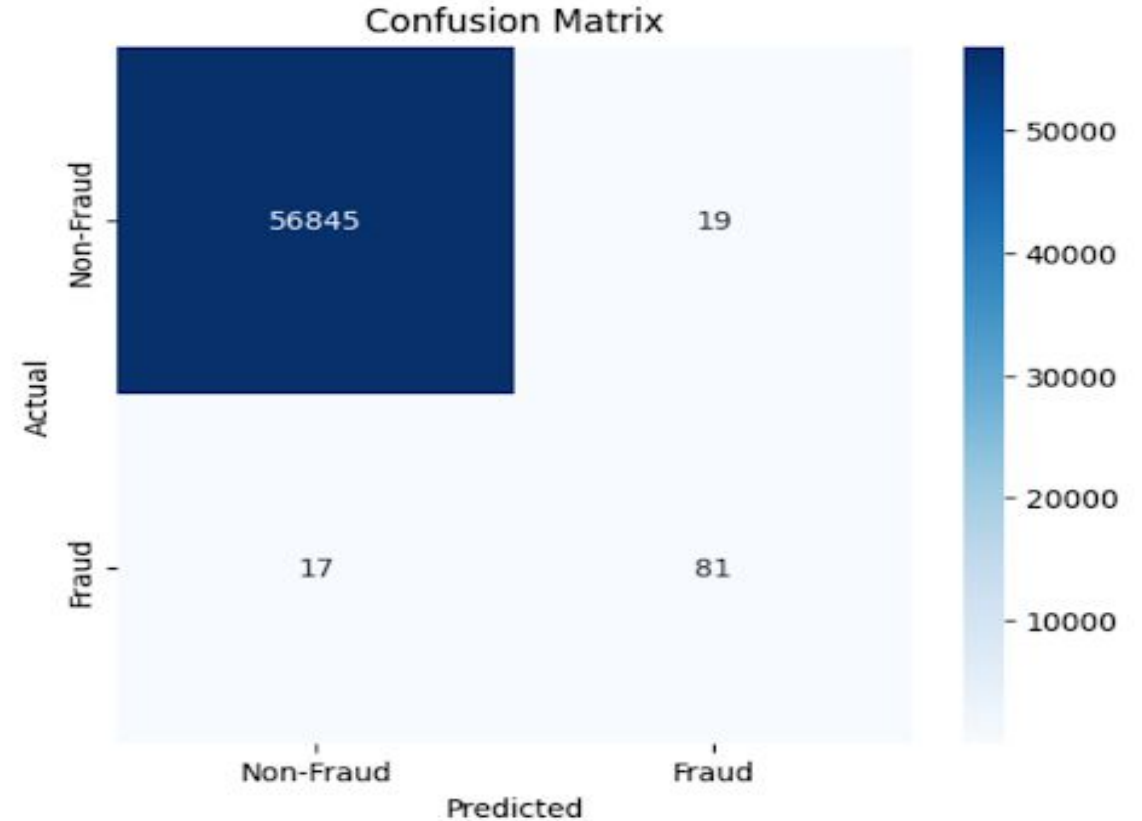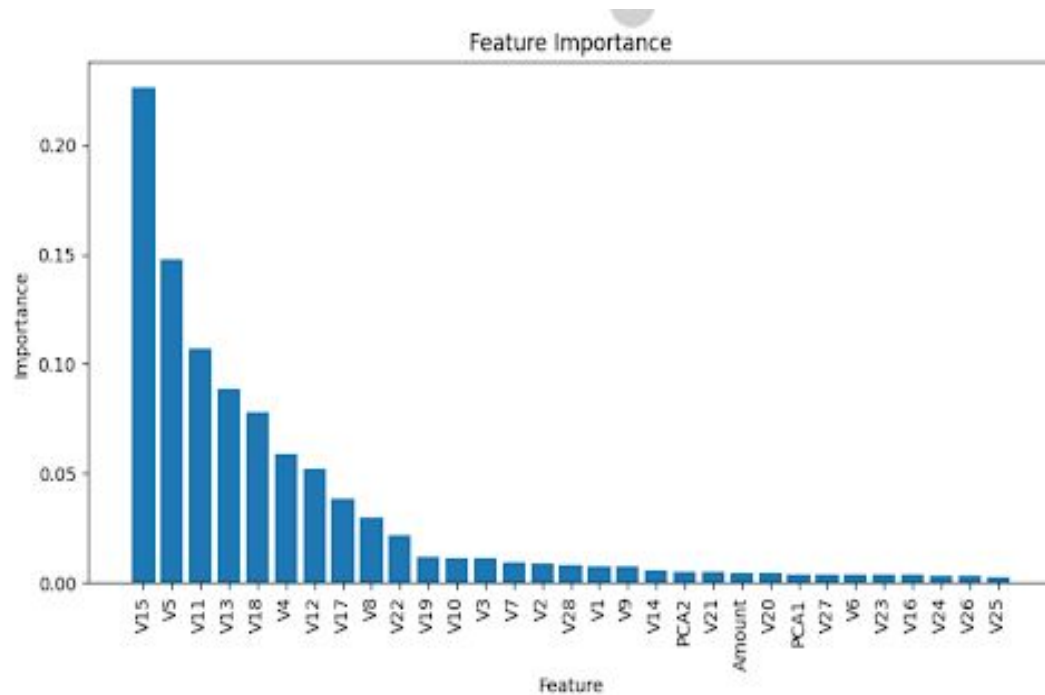
Tech Stacks

Tech Stack & Tools

"Our solution used the following tech stack:

- **Python** – Core programming

- **Pandas & NumPy** – Data manipulation

- **Matplotlib & Seaborn** – Visualization

- **Scikit-learn** – Model building & evaluation

- **Imbalanced-learn** – SMOTE for class imbalance

- **TensorFlow** – Planned for future deep learning integration

- **PCA / t-SNE** – Dimensionality reduction for visualization

- **Random Forest Classifier** – Final model

- **StandardScaler** – For normalization in real-time inference

Conclusion

To sum up, we built a full end-to-end system that ingests financial transaction data, cleans and transforms it, balances it using SMOTE, and applies machine learning to **predict fraud risk in real time**.

We're excited about the potential impact this could have in making financial systems safer and more resilient.

Thank you!

Google   https://colab.research.google.com/drive/1zb1U

Collab   yzwve9WDry_C9ufPoJMfyIkGczko?usp=drive-

Link: dynamite#scrollTo=BvU_VwhHvO1J

"

Sigmoid has been our preferred data & analytics partner. Their teams have an innovative approach to problem-solving and are highly invested to deliver tangible business results!

- VP Marketing of Digital & eCommerce,

F500 CPG firm