Credit Card Fraud Detection using Machine Learning

Aryan Gahlaut (3rd Year)
Delhi Technological University
Machine Learning Intern, Cantilever

August 3, 2025

Abstract

A machine learning method for identifying fraudulent transactions in credit card data is presented in this report. Our goal is to create a reliable and scalable fraud detection model by addressing class imbalance with neural networks and resampling techniques. The findings show that detecting infrequent fraudulent activity can be done with high accuracy and dependability.

Introduction

With serious financial consequences, credit card fraud is becoming a bigger problem. Since fraudulent transactions make up a very small portion of all transactions, the main issue is the stark class disparity. An efficient machine learning pipeline to manage such imbalance and precisely identify fraud is described in this report.

Methodology

Dataset

The dataset contains transactions made by credit cards in September 2013 by European cardholders. It includes 284,807 transactions, out of which 492 are frauds—only 0.172%. All features (except `Time` and `Amount`) have been PCA-transformed due to confidentiality. The target feature `Class` is 1 for fraud and 0 for legitimate.

Preprocessing

- Dropped the `Time` feature as it was not useful.
- Scaled the `Amount` feature using `StandardScaler`.
- Performed a stratified train-validation-test split.
- Applied **SMOTE (Synthetic Minority Oversampling Technique)** to balance the training data.

Model Architecture

A feedforward neural network was implemented using TensorFlow and Keras:

- Input layer matching the number of features
- Two hidden layers with 64 and 32 units, ReLU activations, batch normalization, and dropout
- Output layer with sigmoid activation

Training

The model was trained using the Adam optimizer and binary cross-entropy loss for 10 epochs with a batch size of 512. Validation was monitored to detect overfitting.

Experiments and Results

Performance Metrics

- Accuracy: 99.93%
- Precision: 78.43%
- Recall (TPR): 81.63%
- F1 Score: 80.00%
- False Positive Rate (FPR): 0.04%
- True Negative Rate (TNR): 99.96%
- False Negative Rate (FNR): 18.37%
- ROC AUC: 97.38%
- Precision-Recall AUC: 75.33%

Visualizations

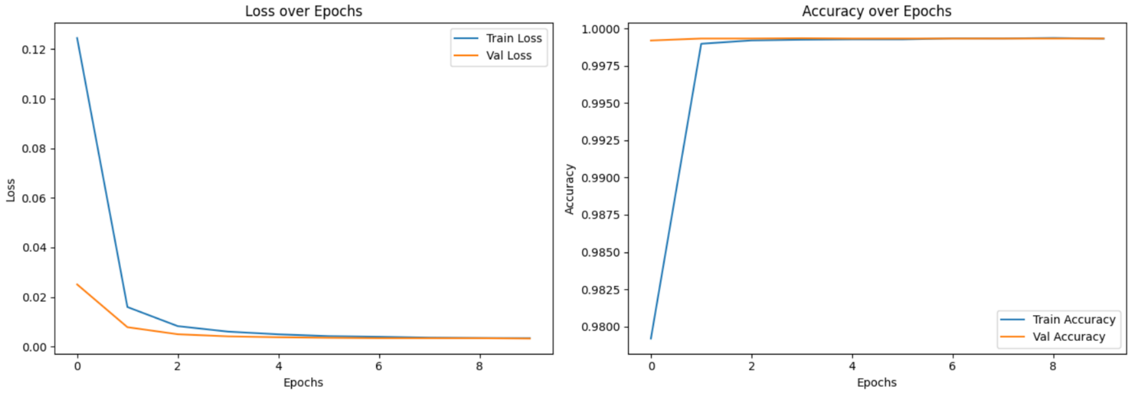
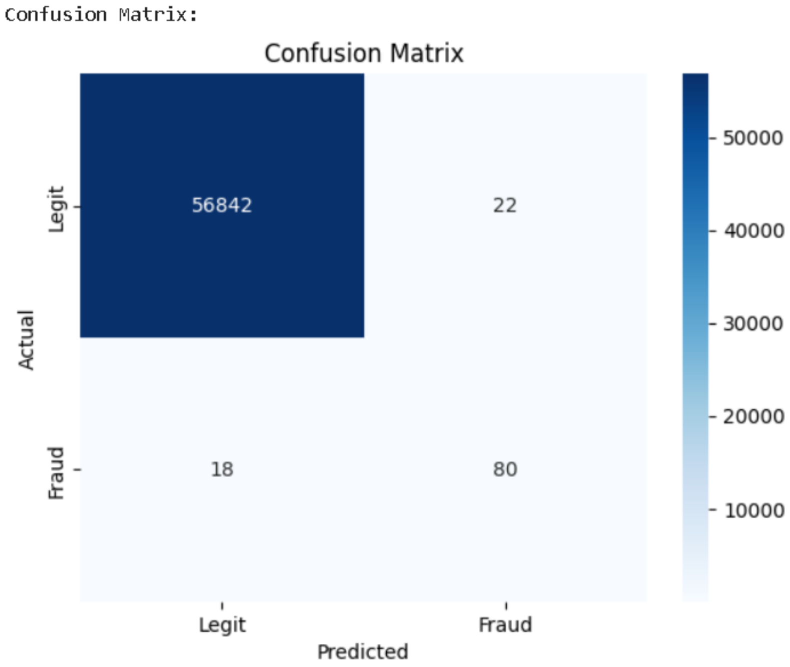


Figure 1: Training and Validation Loss/Accuracy



```
Confusion Matrix:
```

Figure 2: Confusion Matrix

Conclusion

This project shows how neural networks and resampling methods like SMOTE can effectively handle class imbalance for fraud detection. For UI, Gradio is used to easily enter input values for the model to predict. Real-time systems integration, anomaly detection, and ensemble approaches may be investigated in future research.

References

- Kaggle Credit Card Dataset: `https://www.kaggle.com/mlg-ulb/creditcardfraud`
- Chawla, N. V., et al. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research.
- TensorFlow/Keras Documentation: `https://www.tensorflow.org/`