

# Fraud Detection & Anomaly Analytics Engine

## An Unsupervised Learning System for Financial Transaction Surveillance

### Executive Summary

This project focuses on developing a robust fraud detection engine using unsupervised machine learning techniques. Using anonymized credit card transaction data, we identified fraudulent patterns in highly imbalanced datasets through anomaly detection algorithms like Isolation Forest, Autoencoders, and Local Outlier Factor. The system is designed to assist financial institutions in proactively flagging suspicious behavior, improving fraud response time, and reducing losses.

### Tools & Technologies

- **Python Libraries:** NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, TensorFlow, Keras
- **Machine Learning:** Isolation Forest, Local Outlier Factor (LOF), Deep Learning Autoencoder
- **Evaluation Metrics:** Confusion Matrix, Precision, Recall, F1-Score
- **Data:** Anonymized Credit Card Transactions Dataset (284,807 records)
- **Additional Stack:** Scalers (StandardScaler), Model Export (Joblib), Streamlit (for deployment)

### Project Highlights

- **Data Preprocessing & Exploration**
  - Scaled **Amount** and **Time** features using StandardScaler.
  - Performed extensive EDA on class imbalance and transaction behavior.
- **Anomaly Detection Models**
  - **Isolation Forest:** Ensemble-based model achieved strong recall on minority class.
  - **Local Outlier Factor:** Effective for identifying point anomalies based on local density deviation.
  - **Autoencoder:** Neural network trained to reconstruct normal transactions; higher reconstruction error implied potential fraud.
- **Model Comparison**

- Evaluated models on precision, recall, and F1-score specific to the fraud class.
- Autoencoder yielded the most balanced performance with improved fraud detection without sacrificing precision.

## Key Results

Model	Precision (Class 1)	Recall (Class 1)	F1-Score (Class 1)
Isolation Forest	0.31	0.34	0.33
Local Outlier Factor	0.00	0.00	0.00
Autoencoder	<b>0.62</b>	<b>0.75</b>	<b>0.68</b>

## Final Insight

By using deep learning-based Autoencoders and ensemble techniques, the system intelligently identifies suspicious transactions with higher accuracy. This solution is scalable, adaptable to real-time pipelines, and forms the backbone of a deployable fraud detection framework.