# Action Plan: Anomaly Detection in Financial Transactions

## 1. Project Goal

The objective of this project is to develop models that can detect anomalous or fraudulent financial transactions using unsupervised and semi-supervised machine learning techniques. The models used include Isolation Forest, One-Class SVM, and Autoencoders, all of which are designed to identify data points that deviate from the norm in a highly imbalanced dataset.

## 2. Dataset Description

The dataset used is a credit card fraud detection dataset, commonly available from Kaggle. It contains transaction records with anonymized features, along with a binary label indicating whether a transaction is fraudulent (1) or not (0).

## 3. Tools and Libraries

• Python  
• Pandas and NumPy for data manipulation  
• Matplotlib and Seaborn for visualization  
• Scikit-learn for machine learning (Isolation Forest, One-Class SVM)  
• TensorFlow and Keras for Autoencoders  
• StandardScaler for feature scaling

## 4. Exploratory Data Analysis and Preprocessing

The EDA phase included inspecting the dataset, visualizing feature distributions, analyzing class imbalance, and checking for missing values. Preprocessing steps included:  
• Handling missing data  
• Feature scaling using StandardScaler  
• Class imbalance handling using undersampling and class weighting  
• Dimensionality checks and correlations analysis

## 5. Model Implementation

Three anomaly detection models were implemented:  
• Isolation Forest: Built using Scikit-learn to identify outliers based on recursive partitioning.  
• One-Class SVM: Trained to learn the boundary of normal transactions.  
• Autoencoder: Constructed using Keras for unsupervised learning of compressed transaction representations.  
Each model was trained and tested using the preprocessed data.

## 6. Model Evaluation

The models were evaluated using metrics suitable for imbalanced datasets:  
• Precision  
• Recall  
• F1-score  
• Area Under the Curve (AUC)  
These metrics provided a comprehensive view of model performance, especially in identifying rare fraudulent transactions.

## 7. Timeline

Week 1: Research and dataset acquisition  
Week 2: EDA and preprocessing  
Week 3: Model development (Isolation Forest and One-Class SVM)  
Week 4: Autoencoder implementation and evaluation  
Week 5: Analysis, documentation, and reporting

## 8. Conclusion

This action plan outlines the structured process used to build anomaly detection models for financial fraud detection. By leveraging various machine learning techniques and addressing the challenge of imbalanced datasets, this project aims to contribute to secure and intelligent financial systems.