**Problem Statement**

**Objective:**

Develop a machine learning model to accurately identify fraudulent credit card transactions among a dataset of European cardholder transactions from the year 2023.

**Background:**

Credit card fraud is a major concern for financial institutions and cardholders alike, leading to significant financial losses and security issues. With the increasing volume of digital transactions, traditional rule-based systems are becoming less effective, necessitating the need for advanced, scalable, and efficient fraud detection solutions.

**Dataset Description:**

The dataset comprises over 550,000 records of credit card transactions made by European cardholders in 2023. It includes the following features:

id: Unique identifier for each transaction.

V1-V28: Anonymized features representing various aspects of each transaction, likely encompassing time, location, amount, and other transactional details.

Amount: The transaction amount.

Class: Binary label (1 for fraudulent, 0 for non-fraudulent transactions).

#### **Expected Challenges:**

Imbalanced Data, Anonymized Features, Model Interpretability.

**Tasks:**   
Data Preprocessing: Address class imbalance using techniques like oversampling, undersampling, or SMOTE (Synthetic Minority Over-sampling Technique).

Feature Engineering: Derive new features or transform existing ones to improve model performance.

Model Selection: Experiment with various algorithms like Decision Trees, Random Forest, Gradient Boosting, Support Vector Machines, and Neural Networks.

Hyperparameter Tuning: Optimize model parameters for better performance.

Cross-Validation: Implement cross-validation to ensure the model’s robustness and generalizability.

**Evaluate Model using various evaluation parameters.**