**Project Proposal: Credit Card Fraud Detection Using Machine Learning**

**Objective:**  
The aim of our project is to develop and evaluate a machine learning model capable of detecting fraudulent credit card transactions. By leveraging historical transaction data and advanced algorithms, we intend to provide insights into fraud patterns and create a fraud detection system. This solution addresses a critical issue for financial institutions by reducing potential losses and safeguarding user trust.

**Project Scope**

**Dataset:**  
We will use the **Credit Card Fraud Detection Datasets 2023, 2013**

**Dataset** available on Kaggle. This dataset contains 284,807 transactions with the following features:

* **Time**: Timestamp of each transaction, useful for identifying temporal patterns.
* **Amount**: The transaction value, indicating spending trends.
* **V1–V28**: Anonymized features derived from Principal Component Analysis (PCA) for privacy.
* **Class**: A binary label indicating fraud (1) or legitimate (0).

 **Data Exploration and Preprocessing**:

* Analyze feature distributions and visualize differences between fraudulent and legitimate transactions.
* Normalize key features like amount and time to improve model accuracy.

 **Model Development**:

* Train and compare various machine learning models using **Scikit-learn**, including:
  + Logistic Regression (baseline).
  + Random Forest for non-linear relationships.

 **Visualization**:

* Leverage **Matplotlib** and **Plotly** to illustrate:
  + Feature importance.
  + Fraud distribution and trends over time.
  + Confusion matrix for model performance.

 **Presentation**:

* Use **Tableau** to create interactive dashboards showcasing:
  + Fraud trends over time.
  + Key predictive features and their significance.
  + Insights into fraud detection patterns.

 **Documentation**:

* Organize analysis and model implementation in Jupyter Notebooks.
* Provide clear documentation for reproducibility.

**Expected Outcomes**

* A high-performing fraud detection model with balanced precision and recall.
* Insights into the features most indicative of fraud.
* A Tableau dashboard providing an interactive summary of fraud detection insights for stakeholders.

**Technologies and Tools**

 **Programming Language**: Python.

 **Libraries**: Scikit-learn, Pandas, Matplotlib, Plotly.

 **Data Visualization**: Tableau for interactive dashboards.

 **Platform**: Kaggle for dataset and Jupyter Notebook for analysis.