# 📌 Title: Credit Card Fraud Detection using Machine Learning

## 🔹 Introduction:

Fraudulent transactions pose serious risks in the financial sector. This project aims to identify suspicious credit card transactions using the Kaggle creditcard dataset. The goal is to accurately classify transactions as fraud or non-fraud using machine learning, particularly XGBoost.

## 🔹 Abstract:

We loaded and explored the dataset containing over 280,000 transactions. The dataset was imbalanced (only ~0.17% fraud), so we performed feature scaling and model training using XGBoost. Evaluation metrics included confusion matrix, classification report, and ROC curve. The final model was integrated into a Streamlit web application for real-time prediction.

## 🔹 Tools Used:

- Python (Pandas, NumPy, Matplotlib, Seaborn)
- Scikit-learn
- XGBoost
- Streamlit
- Jupyter Notebook

## 🔹 Steps Involved:

1. Data Preprocessing: Removed 'Time', scaled 'Amount', handled class imbalance analysis.
2. Exploratory Analysis: Class distribution and feature correlation.
3. Model Training: XGBoost Classifier trained using scaled inputs.
4. Evaluation: Confusion matrix, classification report, ROC-AUC score.
5. Deployment: Model and scaler saved using `joblib`. Streamlit app created to accept user-uploaded CSV files and display fraud predictions.
6. Optional (Advanced): Included potential use of Isolation Forest and Local Outlier Factor for future improvement.

## 🔹 Conclusion:

The XGBoost classifier achieved high accuracy and recall, making it suitable for identifying fraudulent transactions. The Streamlit app adds usability for practical deployment. Future work can include further balancing techniques and experiment with anomaly detection models like Isolation Forest or SMOTE.