# Fraud Detection in Credit Card Transactions using Machine Learning


## Submitted by:

Pavan Kumar
B.Tech (CSE - IT)
Mallareddy University


## 1. Introduction

Credit card fraud is a significant financial issue worldwide. Fraudulent activities can cause major financial losses to banks and customers. This project aims to build a Machine Learning model that can detect fraudulent credit card transactions and help prevent misuse. The app is built using Python, Machine Learning, and Streamlit, providing both manual and bulk testing functionalities.

## 2. Dataset

The dataset used in this project is the popular Credit Card Fraud Detection Dataset available on Kaggle.
- Total transactions: 284,807
- Fraudulent transactions: 492 (0.172%)
- Legitimate transactions: 284,315
- Features: V1–V28 (anonymized PCA features), Amount, and Time
- Target: Class (0 = Legitimate, 1 = Fraudulent)

## 3. Methodology

1. Data Preprocessing:
- Standardized the 'Amount' column.
- Removed irrelevant features ('Time').
- Created a new column 'normAmount'.

2. Model Training:
- Used Logistic Regression for fraud classification.
- Trained on features V1–V28 and normAmount.

3. Streamlit Application:
- Manual Input: User can enter transaction details manually.
- CSV Upload: Upload multiple transactions for batch prediction.
- Dataset & Charts: View dataset, fraud vs legitimate distribution, and pie/bar charts.

## 4. Results

- The model predicts whether a transaction is Fraudulent (1) or Legitimate (0).
- In the Streamlit app, predictions are shown instantly with clear labels.
- Visualizations: Fraud vs Legitimate transactions are shown using bar charts and pie charts.

## 5. Screenshots

- App Home Page
- Manual Input Prediction
- CSV Upload Prediction
- Fraud vs Legitimate Charts

## 6. Conclusion

This project demonstrates the use of Machine Learning for fraud detection in real-world financial data. The app provides an interactive way to detect frauds and visualize results. It can be further improved using advanced models such as Random Forests, XGBoost, or Deep Learning techniques.

## 7. Future Scope

- Implement Deep Learning models like Autoencoders and LSTMs.
- Improve accuracy with ensemble models.
- Deploy on cloud platforms for large-scale real-time fraud detection.