**CREDIT CARD FRAUD DETECTION**

Overview:

This project focuses on fraud detection using machine learning techniques. It employs a Random Forest classifier to identify fraudulent transactions from a given dataset. The main goal is to provide accurate predictions for detecting potential fraud in financial transactions.

Key Steps:

1. Data Preprocessing: The project starts by loading the dataset, which is a CSV file, using the Pandas library. It preprocesses the data, including one-hot encoding categorical features ('merchant', 'category', 'job'), converting the 'gender' feature to binary values, and handling missing data.
2. Feature Selection: The project selects relevant features for model training and separates them into the feature matrix (X) and the target variable (y).
3. Model Training: A Random Forest classifier is utilized to train a machine learning model. It's configured with 100 trees (n\_estimators) and a random seed for reproducibility.
4. Model Evaluation: The model's performance is assessed using common classification metrics such as accuracy, the confusion matrix, and a classification report. These metrics help gauge the effectiveness of the fraud detection system.

Usage:

1. Download the dataset and place it in the project directory.
2. Ensure you have Python, Pandas, NumPy, and scikit-learn installed.
3. Run the fraud\_detection.py script to train the Random Forest classifier and make predictions.

Project Structure:

* fraud\_detection.py: The main Python script for data preprocessing, model training, and evaluation.
* fraudTest.csv: The dataset used for training and testing.