Credit Card Fraud Detection

This project uses machine learning to detect fraudulent credit card transactions. The Random Forest Classifier is used to classify transactions as fraudulent or legitimate based on the dataset.

Steps in the Code

1. Import Libraries

The project uses the following libraries:

- `numpy` : For numerical operations.
- `pandas` : For data handling and manipulation.
- `matplotlib` & `seaborn` : For data visualization.
- `scikit-learn` : For machine learning and model evaluation.

2. Load Dataset

The dataset `creditcard.csv` is loaded using `pandas.read_csv` .

```
data = pd.read_csv('creditcard.csv')
```

3. Explore the Data

- View the first few rows of the dataset to understand its structure:

```
print(data.head())
```

- Check for missing values:

```
print(data.isnull().sum())
```

4. Analyze Class Distribution

Count and visualize the class distribution to check for imbalance:

```
print(data['Class'].value_counts())
sns.countplot(x='Class', data=data)
plt.show()
```

5. Prepare the Data

Split the dataset into features (x) and target (y), then split into training and testing sets:

```
X = data.drop('Class', axis=1)
y = data['Class']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)
```

6. Train the Model

Train a Random Forest Classifier on the training data:

```
model = RandomForestClassifier(random_state=42)
model.fit(X_train, y_train)
```

7. Test the Model

Predict the class (fraud or legitimate) on the test data:

```
y_pred = model.predict(X_test)
```

8. Evaluate the Model

Evaluate the performance of the model using a confusion matrix and classification report:

```
print(confusion_matrix(y_test, y_pred))
print(classification_report(y_test, y_pred))
```

How to Run

1. **Install Requirements**:
   Ensure you have the necessary libraries installed:

   ```
   pip install numpy pandas matplotlib seaborn scikit-learn
   ```

2. **Run the Notebook**:
   Open the Jupyter Notebook and execute the cells step by step.

3. **Input Dataset**:
   Place the `creditcard.csv` file in the same directory as the notebook.

Outputs

- **Confusion Matrix**: Shows correct and incorrect predictions.
- **Classification Report**: Displays accuracy, precision, recall, and F1-score.