**Credit Card Fraud Detection**

**Project Description:** The "Credit Card Fraud Detection using Logistic Regression" project aims to build a predictive model that can distinguish between legitimate and fraudulent credit card transactions. The dataset used in this project contains information about credit card transactions, including features like transaction amount and time. The project follows a step-by-step approach to achieve this goal:

**1. Data Exploration and Pre-processing:**

- The project begins with importing necessary libraries, loading the dataset, and performing initial data exploration.

- It displays the top and bottom 5 rows of the dataset, checks for null values, and performs feature scaling using StandardScaler.

- The 'Time' column is dropped, as it is not considered as a relevant feature for the model.

**2. Handling Duplicates:**

- Duplicate rows in the dataset are identified and removed to ensure data integrity.

**3. Addressing Class Imbalance:**

- The dataset is highly imbalanced, with a significant difference between the number of legitimate and fraudulent transactions.

- A visualization is provided to show the distribution of both classes.

- To address this, the project uses an under-sampling technique to create a balanced sample dataset.

**4. Statistical Measures:**

- Statistical measures of transaction amounts for both legitimate and fraudulent transactions are presented, giving insights into their distributions.

**5. Comparing Transaction Values:**

- The mean values of different features for both legitimate and fraudulent transactions are compared, highlighting the differences between the two classes.

**6. Model Training and Evaluation:**

- The project uses Logistic Regression, a binary classification algorithm, to build the fraud detection model.

- The data is split into features (X) and the target variable (Y), and further divided into training and testing sets.

- The model is trained on the training data and evaluated on both training and testing sets.

**7. Results and Insights:**

- The accuracy of the model on both training and testing data is reported, providing an indication of its performance.

- This project serves as a foundation for more advanced fraud detection techniques and models.

**Conclusion:**

The "Credit Card Fraud Detection using Logistic Regression" project successfully demonstrates the process of building a simple yet effective fraud detection model. The logistic regression algorithm, coupled with appropriate data pre-processing techniques, proves to be a valuable tool in identifying fraudulent credit card transactions. This project lays the groundwork for more sophisticated machine learning approaches in the field of fraud detection.