**Prediction of Selling Price of Used Cars**

1. **Objective:**

To develop a predictive model to accurately estimate the selling prices of used cars based on various features such as make, model, year, mileage, and other relevant attributes.

1. **Scope:**

This project covers data preprocessing, exploratory data analysis (EDA), model training, prediction, evaluation, and visualization. The performance of linear regression and logistic regression models are compared, with an emphasis on achieving high accuracy in price prediction.

1. **Business Problem:**

Estimating the selling price of a used car can be challenging due to numerous influencing factors. An accurate prediction model can help buyers and sellers make informed decisions and facilitate fair transactions.

1. **Methodology:**

**4.1 Data Collection:**

- Importing the necessary libraries.

- Loading the dataset, which includes features such as make, model, year, mileage, and price.

- Source for the dataset: https://www.google.com/url?q=https%3A%2F%2Fwww.kaggle.com%2Fnehalbirla%2Fvehicle-dataset-from-cardekho%3Fselect%3Dcar%2Bdata.csv

**4.2 Data Preprocessing:**

1. First Look at the MPG dataset.

2. Data Cleaning:

- Standardizing column names and text columns.

- Removing unnecessary spaces from text columns.

- Removing duplicated records.

- Handling missing values.

- Extracting information from specific columns.

- Handling outlier records.

**4.3 Exploratory Data Analysis (EDA):**

- Univariate Analysis: Examining the distribution of individual features.

- Bivariate Analysis: Exploring relationships between pairs of features.

- Multivariate Analysis: Investigating interactions among multiple features.

**4.4 Model Development:**

- Training both linear regression and logistic regression models.

- Comparing their performance based on evaluation metrics.

**4.5 Evaluation Metrics:**

- Mean Squared Error (MSE)

- Mean Absolute Error (MAE)

- R2 Score

**5. Key Findings:**

- The logistic regression model outperformed the linear regression model with an R2 score of 0.986, indicating a very high level of accuracy in predicting car prices.

**6. Recommendations:**

- Utilizing the logistic regression model for price prediction due to its superior performance.

- Continuous updating and retraining of the model with new data to maintain accuracy over time.

- Further refining the model by exploring additional features or advanced techniques such as ensemble methods.