**CarDekho Machine Learning Model for Used Car Price Prediction.**

**-By Shivani Nagar**

**Dataset Link**

The dataset used for this project can be accessed [here](https://drive.google.com/file/d/1WtxKHx5uQoFYmAEKNWs0Jdx4jkS-OXDq/view?usp=sharing).

**Project Overview**

This project focuses on developing a machine learning model to predict the price of used cars based on their features. By leveraging historical data from CarDekho.com, we aim to assist both buyers and sellers in making more informed decisions in the used car market.

**Objectives**

* To build a machine learning model that accurately predicts the price of a used car.
* To empower stakeholders with better pricing insights.

**Benefits**

1. **For Sellers**: Improved accuracy in pricing can lead to faster sales and higher profit margins.
2. **For Buyers**: Access to competitive pricing enhances decision-making.
3. **For the Market**: A more efficient and transparent used car marketplace in India.

**Approach**

1. **Dataset Analysis**: The dataset was analyzed for trends, patterns, and relationships between features.
2. **Data Preprocessing**: Steps included:
   * Handling missing values.
   * Encoding categorical features.
   * Scaling numerical features.
   * Removing outliers.
3. **Univariate Analysis**: Key insights from individual features such as mileage, year of manufacture, and price were visualized using histograms and boxplots to understand their distributions.
4. **Bivariate Analysis**: Relationships between features and the target variable (selling price) were explored using scatter plots and correlation matrices. For example, mileage vs. price and year vs. price were analyzed to identify trends.
5. **Multivariate Analysis**: Combined effects of multiple features on price were investigated using pair plots and regression plots to identify interactions and dependencies. (Using corr() and Heatmap())
6. **Feature Selection**: Identifying the most significant features influencing car prices.
7. **Model Development**: Implementing a regression-based machine learning model to predict prices.
8. **Evaluation**: Using appropriate metrics like R-squared and Mean Squared Error to measure model performance.

**Tools and Libraries**

* Python
* Pandas, NumPy (Data Manipulation)
* Matplotlib, Seaborn (Data Visualization)
* Scikit-learn (Machine Learning Model Development)

**Solution Details**

The complete implementation of this project is detailed in the provided Jupyter Notebook. It includes:

* Exploratory Data Analysis (EDA)
* Univariate, bivariate, and multivariate analyses
* Data preprocessing techniques
* Model building and evaluation

**Code File :** [**Github Link**](https://github.com/Nagar2nd/ML-RegressionModel---CarDekho-Price-Prediction)

**Results**

* The machine learning model was successfully trained and evaluated.
* Key insights:
  + Features such as age, mileage, fuel type, and brand significantly influence the price.
  + The model demonstrated satisfactory performance on test data.

**Conclusion**

This machine learning solution provides a reliable method to predict used car prices in the Indian market. With further fine-tuning and additional data, the model's accuracy and applicability can be enhanced.