CAR PRICE PREDICTION

We have to scrape at least 5000 used cars data. we scraped more data as well. more the data better the model .In this section we need to scrape the data of used cars from websites (Olx, cardekho, Cars24 etc.) We fetched data using web scraping for this. You have to fetch data for different locations. Generally, these columns are Brand, model, variant, manufacturing year, driven kilometers, fuel, number of owners, location and at last target variable Price of the car.

after collecting the data, you need to build a machine learning model. Before model building do all data pre-processing steps. Try different models with different hyper parameters and select the best model. Follow the complete life cycle of data science. Include all the steps like. 1. Data Cleaning 2. Exploratory Data Analysis 3. Data Pre-processing 4. Model Building 5. Model Evaluation 6. Selecting the best model.

1)Data cleaning:

There are no nullvalues and duplicates in our dataset

2)EDA:we plot different plots to check the distribution of the different variables..

We used the count plot,histplot,barplot..etc

We checked whether the outliers are present in our dataframe.

We removed the outliers using the iqr method

3)Data-preprocessing:we fetch the categorical variable from the dataframe and convert it into the continuous variables

We use the different techniques such as label encoding,ordinal encoding and one hot encoding techniques..

4)Model building:we split the data into train and test split.

Then we standardize the data using the standard scaler.

Then we train the model using the different models such as ridge regression,gradientboosting regression,kneighborsregressor,svr and RandomforestRegressor..etc

We find the best model using the metric such as r2_score,MAE and MSE.

5)ModelEvaluation: we test the best ml model with test data and record the metrics.

6)hyperparameter tuning is used with the best model and try the best parameters for ml model i.e Randomforest and this will improve the accuracy of best model.