**DOCUMENTATION Team-19**

**Used Car Price Prediction Using ML AND DL.**

**The Documentation file structure follows the following pattern:**

Model Evaluation

Model

Creation

Feature

Engineering

EDA

Data

Preparation

Data

**Data:**

Data Set Link:- https://drive.google.com/drive/folders/1WVg8RwErjepGrhXHJ5jo4XLdeUaolAfj?usp=sharing

Collection of Data Sets:

1. CarPrice\_Assignment.csv

**Data Preparation:**

1. Data Cleaning: Identifying and correcting mistakes or errors in the data. In data sets, Neither Missing nor duplicate rows present.
2. Identifying input variables that are more relevant to the task.
3. Adding new features and attributes to the data sets
4. Finding correlation between all the columns and remove unwanted rows.
5. Data structuring:The data needs to be structured ,modeled and organized into a modified format.
6. Handling data set CarPrice\_Assignment.csv

**EDA:**

1. Importing Libraries
2. Loading the Raw data
3. Preprocessing
4. To check if there is any outliers or not
5. Here we conclude that we don’t have any outliers as the values are gradually increasing
6. Information of the data
7. Checking the missing value
8. Data Handling
9. In the data column make,we can observe following dependencies
10. Data Visualisation
11. Feature Selection for model building
12. Data Preparation
13. Splitting the Data Feature scaling
14. Building the linear regression model
15. Adopting RFE technique to select the features
16. Building model using statsmodel, for the detailed statistics
17. Residual Analysis of the train data
18. Making Predictions of the Final Model
19. Model Evaluation
20. Conclusions

**Feature Engineering:**

1. Imputation: If missing values are present ,impute them.
2. Encoding categorical features and Standardization of data.
3. Scaling: For symmetric dataset scaling is required and used in normalization and standardization.

**Model Creation:**

1. It is iterative phase where a data scientist continually train and test machine learning models to discover the best one for the task.
2. Linear Regression, SVM algorithm and some other regression techniques are applied to old car price prediction
3. Adding a constant variable
4. Calculate the VIFs for the model
5. None of the VIF's are above 5 and from the summary we can see that none of them have significant p-value (> 0.05). So we can conclude that this is our final model.
6. Making Predict Now that we have fitted the model and checked the normality of error terms, let us now make predictions using our final model.
7. Applying the scaling on the test setsions Using the Final Model

**Model Evaluation:**

1. Model Evaluation aims to estimate the generalization accuracy of model on future.
2. Let's now plot the graph for actual versus predicted values.
3. There are 3 error metrics are used for evaluating and reporting the performance of a regression model.
4. It seems like the predicted values are really good. Actual scores and predicted scores have almost perfect linearity. WE might achieve a more better relation if we treat the outliers but since our dataset is small we choose not to deal with them as we might end up loosing vital information.

* Mean Squared Error (MSE)
* Root Squared Error (RMSE)
* Mean Absolute Error(MAE)

Depending on the model evaluation we can give best model for our problem.

**Conclusions**

1. Following are the conclusions based on the model we built.
2. R-sqaured and Adjusted R-squared (extent of fit) are 0.875 and 0.870 respectively - 87% variance explained.
3. F-stats and Prob(F-stats) (overall model fit) is 192.0 and 6.44e-60(approx. 0.0) respectively - Model fit is significant and explained 87% variance is just not by chance.
4. p-values - p-values for all the coefficients seem to be less than the significance level of 0.05. This indicates that all the predictors are statistically significant.
5. Thus, we say that themodel will be a good way for management to understand the pricing dynamics of a new market.
6. Following are the variables that are significant in predicting the price of a car:
7. Horsepower
8. Wheelbase
9. carbody (hatchback)
10. carbody (wagon)
11. Luxury cars (Cars that are very expensive or high-end)