

ARTIFICIAL INTELLIGENCE

ASSIGNMENT – IV

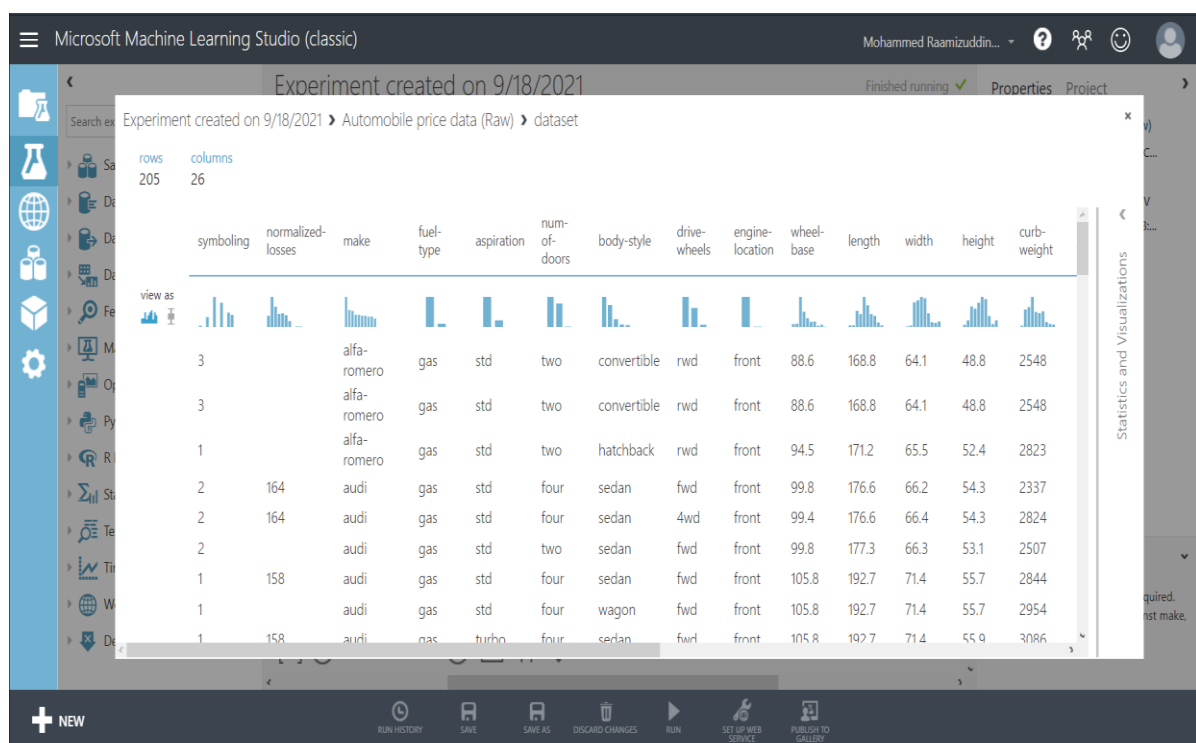
Software Used: Microsoft Azure Machine Learning Studio.

About the Assignment: The model used in this assignment is Linear Regression Model. The trained model finds the price of the automobile based on certain parameters. The Steps Followed in Workflow are:

1. Load the data.
2. Study the data and find out the necessary variables.
3. Pre-Process the data.
4. Choose the model to be used (Linear Regression Model in this case)
5. Split the data for training and testing.
6. Train the model with the training data.
7. Score the model using the testing data.
8. Evaluate the model based on the results.

Workflow:

- Study the data:



- Remove the missing data:

The screenshot shows the Microsoft Machine Learning Studio (classic) interface. The workflow consists of the following steps: 'Automobile price data (Raw)', 'Clean Missing Data' (highlighted), 'Split Data', 'Linear Regression', 'Train Model', 'Score Model', and 'Evaluate Model'. The 'Clean Missing Data' step is currently selected, and its properties are displayed on the right. The 'Columns to be cleaned' are set to 'All columns'. The 'Minimum missing value' is 0, and the 'Maximum missing value' is 1. The 'Cleaning mode' is set to 'Remove entire row'. The status of the experiment is 'Finished running'.

Properties Panel - Clean Missing Data

- Columns to be cleaned: **Selected columns: All columns**
- Launch column selector
- Minimum missing value: 0
- Maximum missing value: 1
- Cleaning mode: Remove entire row
- START TIME: 9/18/2021
- END TIME: 9/18/2021
- ELAPSED TIME: 0:00:01.6...
- STATUS CODE: Finished

- Splitting the data:

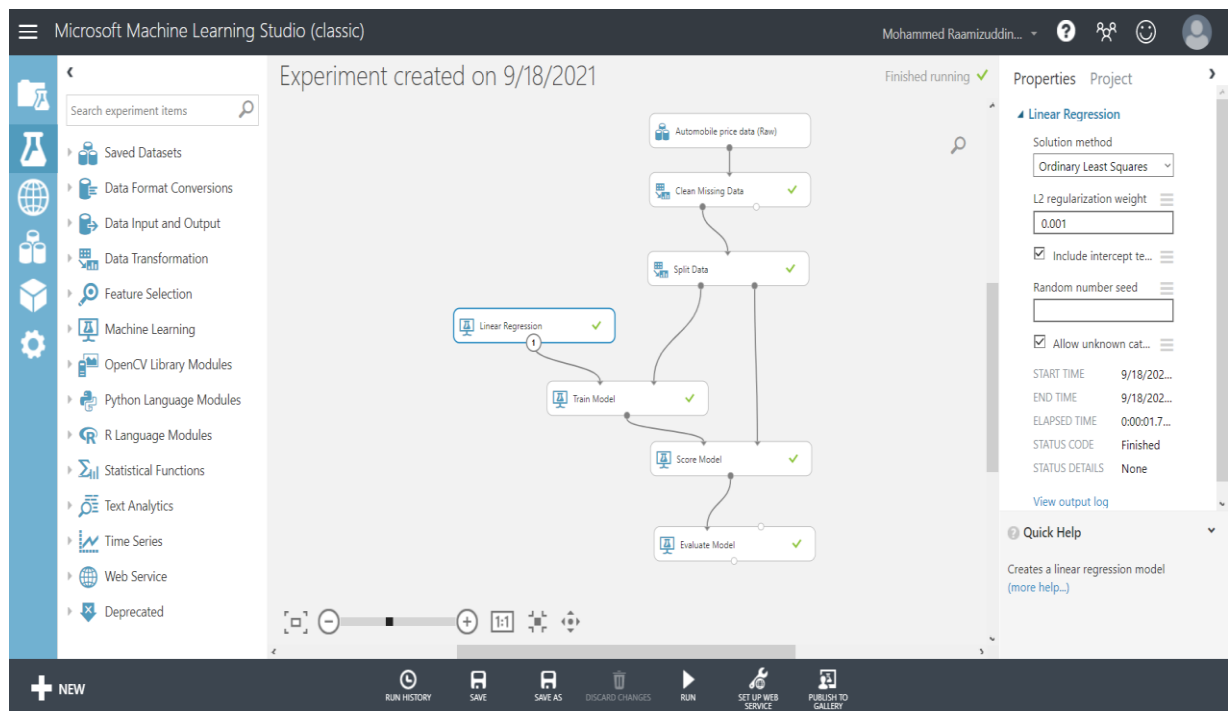
The screenshot shows the Microsoft Machine Learning Studio (classic) interface. The workflow consists of the following steps: 'Automobile price data (Raw)', 'Clean Missing Data', 'Split Data' (highlighted), 'Linear Regression', 'Train Model', 'Score Model', and 'Evaluate Model'. The 'Split Data' step is currently selected, and its properties are displayed on the right. The 'Splitting mode' is set to 'Split Rows'. The 'Fraction of rows in the f...' is 0.7. The 'Randomized split' checkbox is checked. The 'Random seed' is 0. The 'Stratified split' is set to 'False'. The status of the experiment is 'Finished running'.

Properties Panel - Split Data

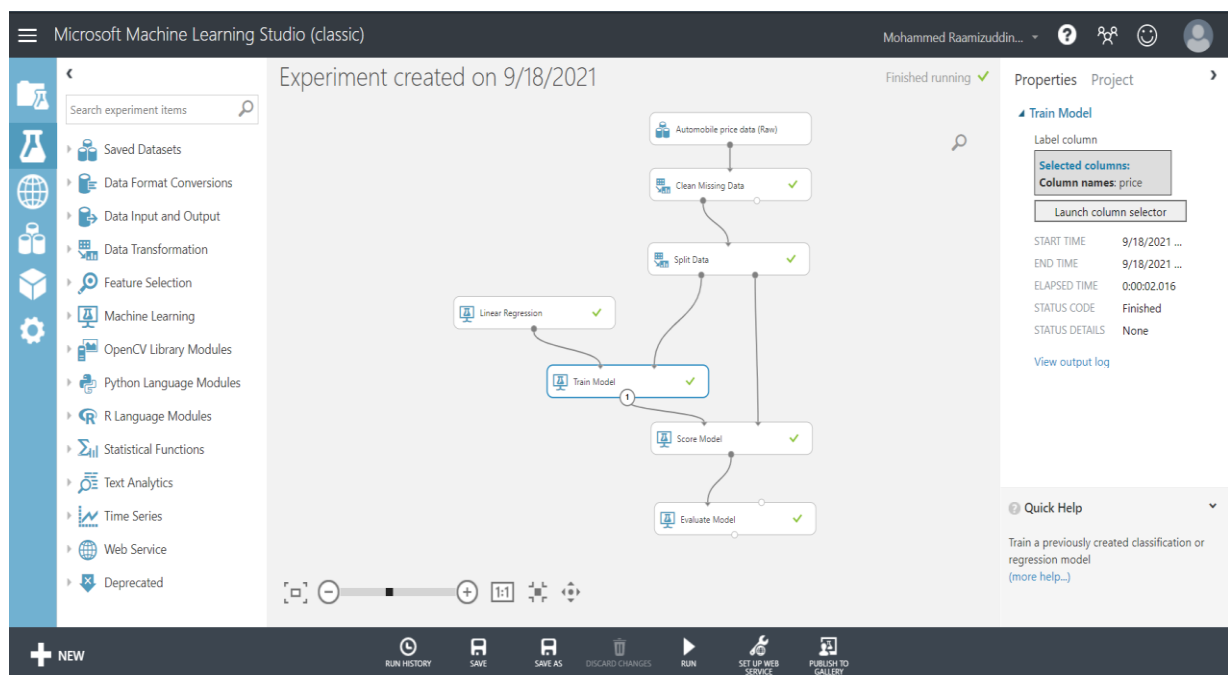
- Splitting mode: Split Rows
- Fraction of rows in the f...: 0.7
- Randomized split: ☒
- Random seed: 0
- Stratified split: False
- START TIME: 9/18/2021
- END TIME: 9/18/2021
- ELAPSED TIME: 0:00:01.6...
- STATUS CODE: Finished
- STATUS DETAILS: None

Next steps are to initialize the model and train the model for the prediction of prices depending on all the other data items in columns of the table.

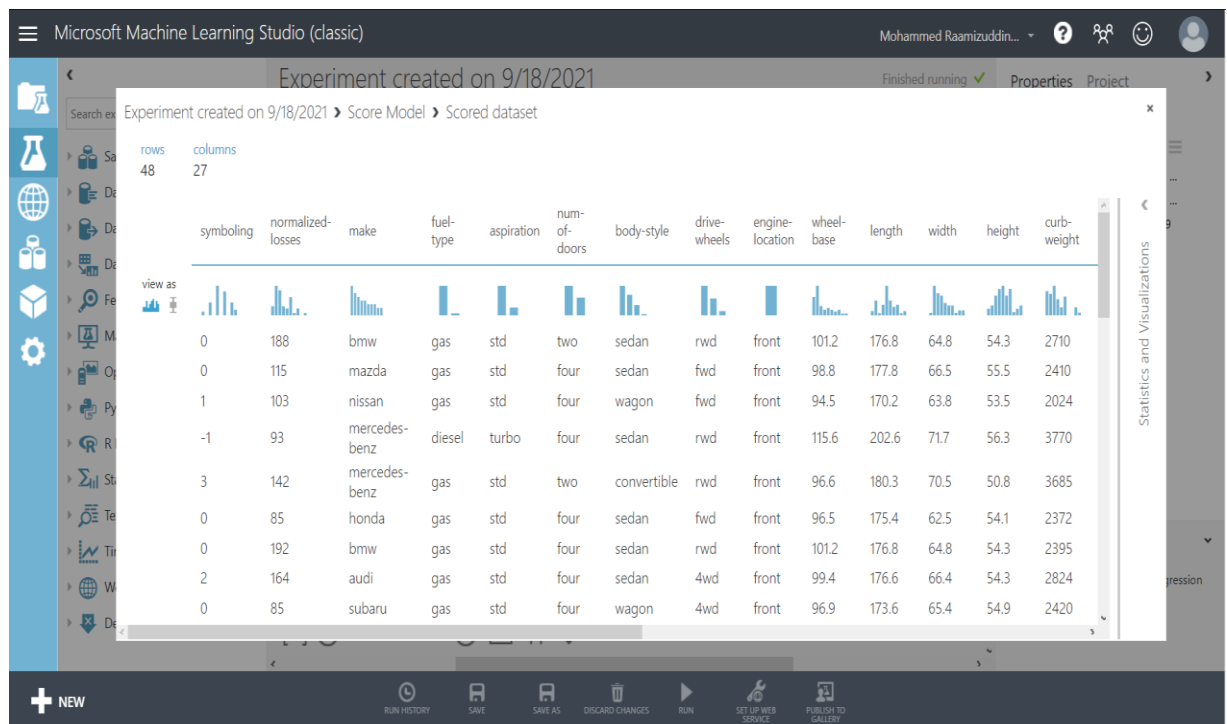
- Model initialization:



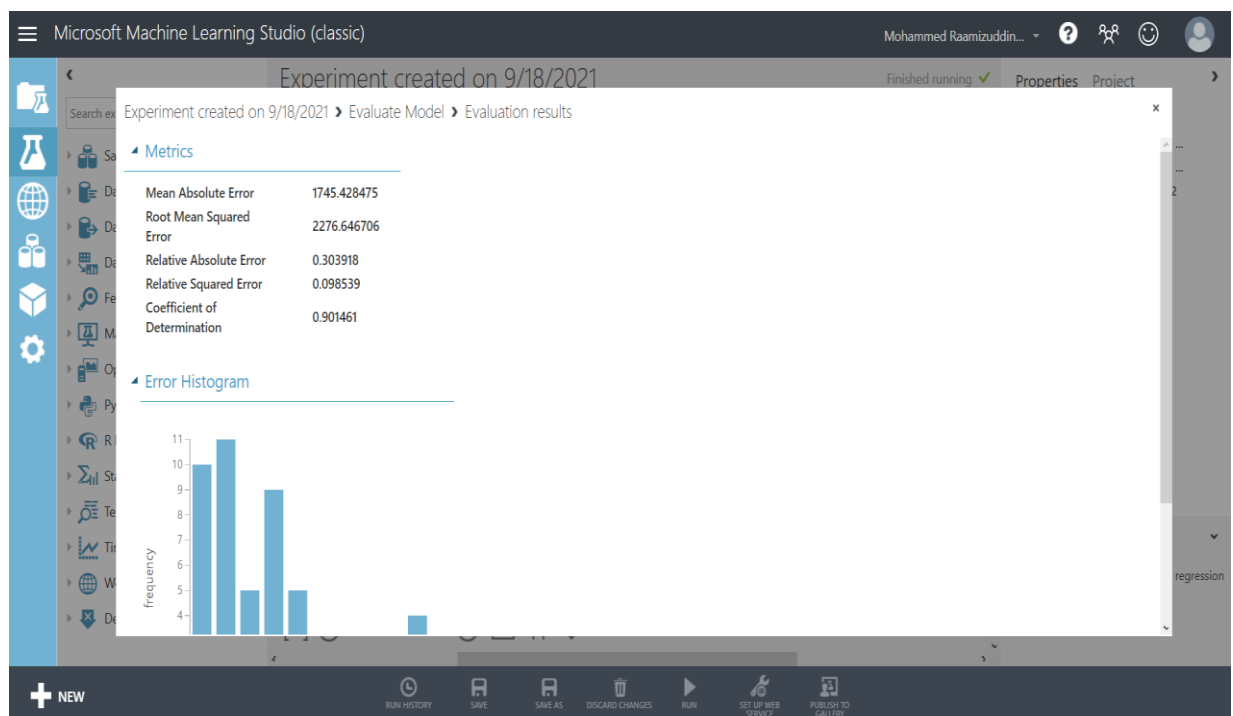
- Training model and Scored dataset:



The result of the prediction is that the price is predicted with an accuracy of 90.14 percent (%) as per the evaluation the model is able to predict outcome properly in most cases.



- Result:



The parameters used to determine price are:

All the parameters of 27 Columns are used to determine the price.

- Workflow:

