Minor Project
ReportSubmitted
For

ARTIFICIAL INTELLIGENCE-(UCS411)

Submitted by:

Sanyam Singhal -102003125
Karamveer Singh - 102003139

B.E Second Year

Submitted To

Mr Niyaz Ahmad Wani

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Computer Science and Engineering Department

TIET,Patiala

Feb-June 2022

TABLE OF CONTENT

ABSTRACT

Deciding whether a used car is worth the posted price when you see listings online can be difficult. Several factors, including mileage, make, model, year, etc. can influence the actual worth of a car. From the perspective of a seller, it is also a dilemma to price a used car appropriately[2-3]. Based on existing data, the aim is to use machine learning algorithms to develop models for predicting used car prices.

.

DECLARATION

I,the undersigned solemnly declare that the project report CAR PRICE PREDICTION USING ML is based on my own work.

I assert the statements made and conclusions drawn are an outcome of myresearch work. I further certify that

I. The work contained in the report is original and has been done by meunder the general supervision of my supervisor.
II. The work has not been submitted to any other Institution for any other degree/diploma/certificate in this university or any other University of Indiaor abroad.
III. We have followed the guidelines provided by the university in writing thereport.

Sanyam Singhal-102003125
Karamveer Singh-102003139

INTRODUCTION

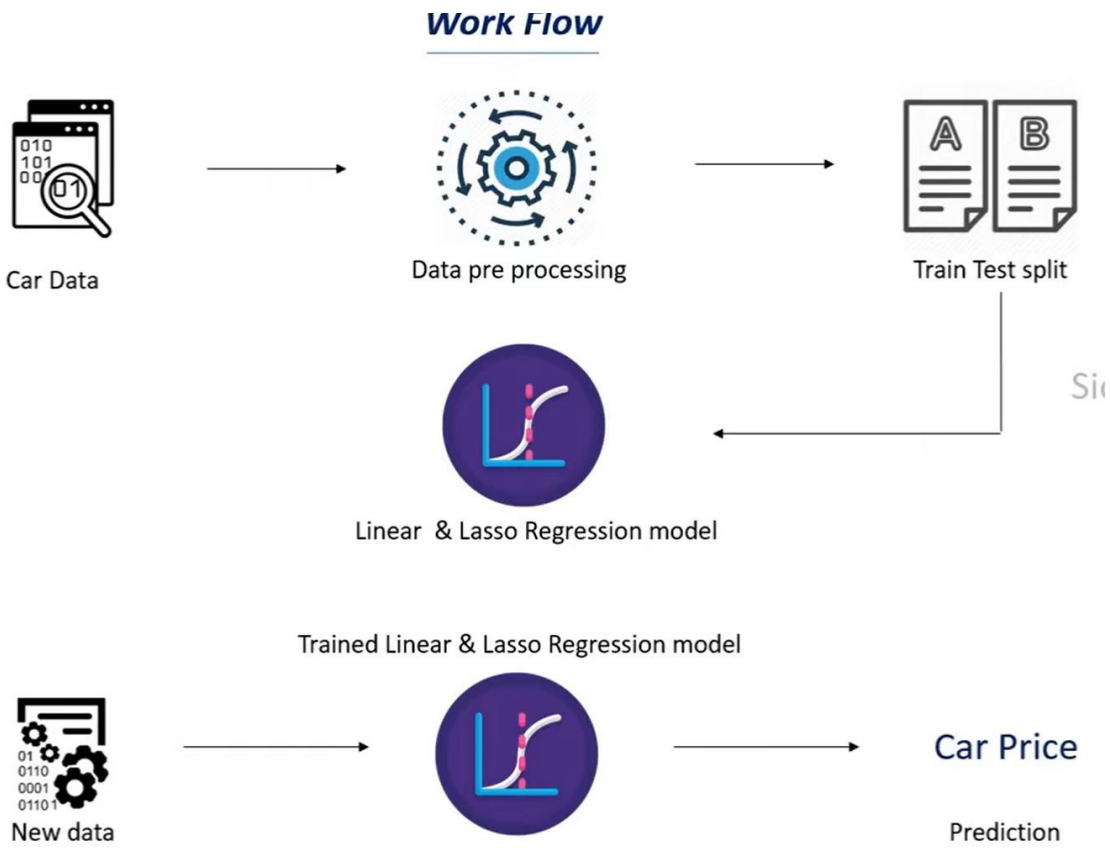
This project focuses on improving the accuracy of existing Machine Learning Data Models using the Linear Regression Model.

A linear regression model describes the relationship between a dependent variable, y, and one or more independent variables, X. The dependent variable also called as response variable. Independent variables also known as predictor or explanatory variables.

The biggest advantage of linear regression models is linearity: It makes the estimation procedure simple and, most importantly, these linear equations have an easy to understand interpretation on a modular level.

The price of a new car in the industry is fixed by the manufacturer with some additional costs incurred by the Government in the form of taxes. So, customers buying a new car can be assured of the money they invest to be worthy. But, due to the increased prices of new cars and the financial incapability of the customers to buy them, Used Car sales are on a global increase. Therefore, there is an urgent need for a Car Price Prediction system which effectively determines the worthiness of the car using a variety of features. Existing System includes a process where a seller decides a price randomly and buyer has no idea about the car and it's value in the

present day scenario. In fact, seller also has no idea about the car's existing value or the price he should be selling the car at. To overcome this problem we have developed a model which will be highly effective. Regression Algorithms are used because they provide us with continuous value as an output and not a categorized value. Because of which it will be possible to predict the actual price a car rather than the price range of a car.


Work Flow
Car Data
Data pre processing
Train Test split
Linear & Lasso Regression model
Trained Linear & Lasso Regression model
New data
Car Price
Prediction

Problem Statement

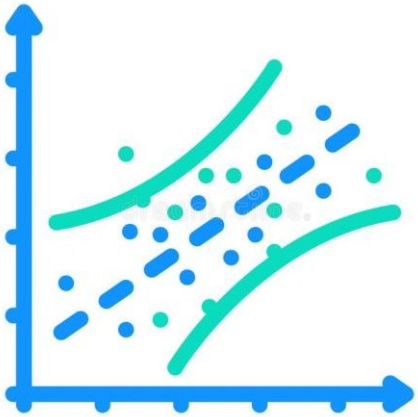
The problem statement is that we have to predict the prices of used cars based on the provided information that includes various fields such as Car brand , year of purchase , fuel type etc .

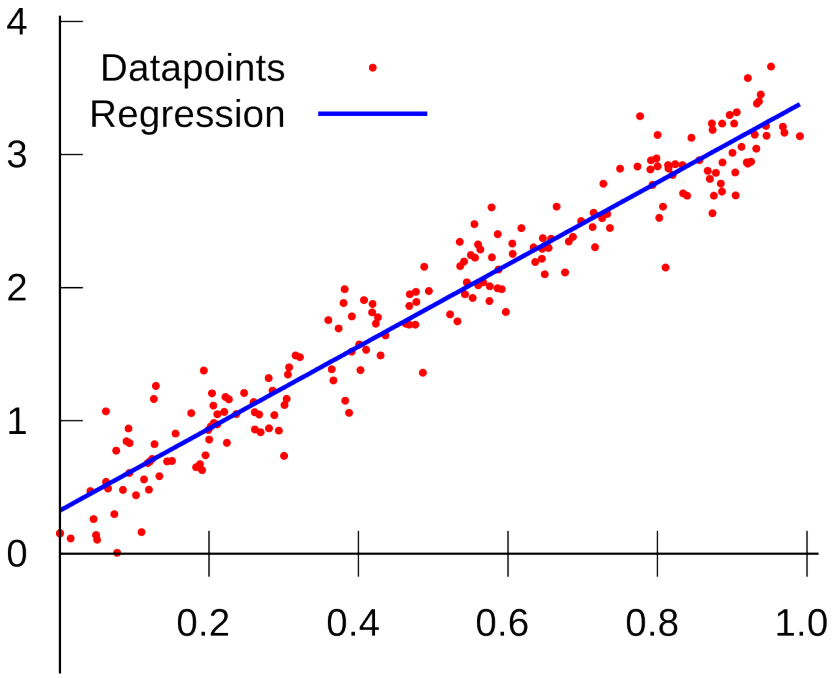
. Here we are using Linear regression simpler model to predict the price of the used cars.

We need to build a machine learning system that can learn from data and predict what this price can be. For that we will be dividing the data into test and train mode.

Linear Regression

Linear Regression is a machine learning algorithm based on **supervised learning**. It performs a **regression task**. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on – the kind of relationship between dependent and independent variables they are considering, and the number of independent variables getting used.



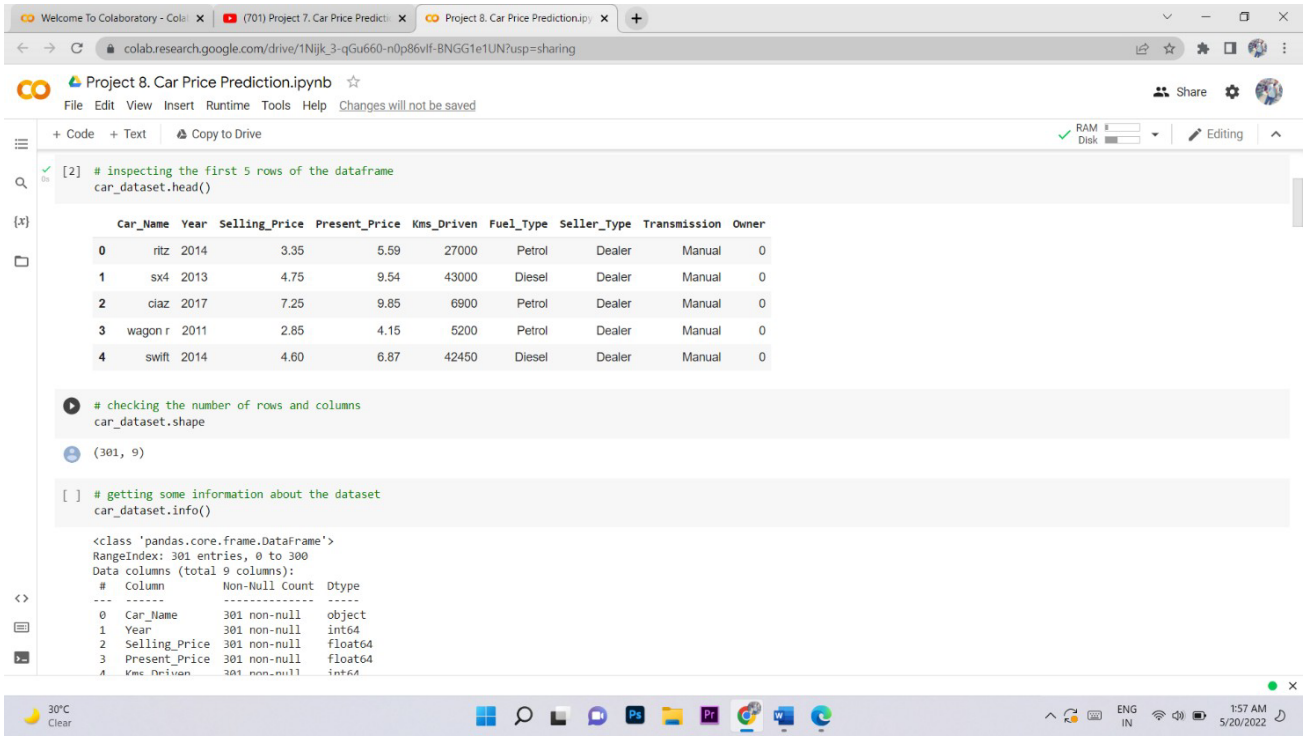


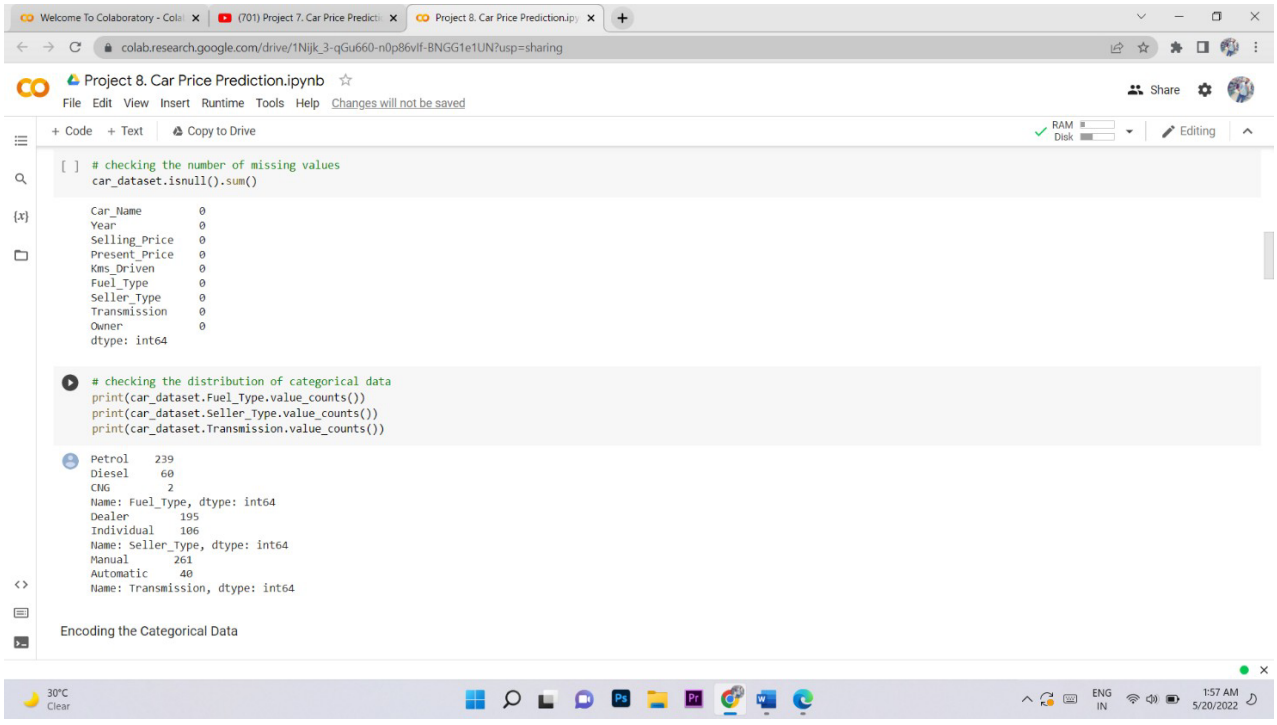
RESULT ANALYSIS

We analysed learning algorithms on our chosen data set. By using the linear regression model we aimed to maintain the linearity of the database so as to predict the data very closely. The biggest advantage of linear regression models is linearity.

Through this project, we found that most suited and, therefore, most common algorithm used to work on big data with linearity, this is very important classifier. It makes the estimation procedure simple and, most importantly, these linear equations have an easy to understand interpretation on a modular level.

Working Screenshots





Project 8. Car Price Prediction.ipynb

File Edit View Insert Runtime Tools Help Changes will not be saved

Splitting the data and Target

```
[ ] X = car_dataset.drop(['Car_Name','Selling_Price'],axis=1)
    Y = car_dataset['Selling_Price']
```

```
[ ] print(X)
```

```
     Year  Present_Price  Kms_Driven  ...  Seller_Type  Transmission  Owner
0    2014           5.59       27000  ...            0             0      0
1    2013           9.54       43000  ...            0             0      0
2    2017           9.85        6900  ...            0             0      0
3    2011           4.15        5200  ...            0             0      0
4    2014           6.87       42450  ...            0             0      0
..    ...            ...         ...  ...          ...           ...    ...
296  2016          11.60       33988  ...            0             0      0
297  2015           5.90       60000  ...            0             0      0
298  2009          11.00       87934  ...            0             0      0
299  2017          12.50        9000  ...            0             0      0
300  2016           5.90        5464  ...            0             0      0

[301 rows x 7 columns]
```

```
[ ] print(Y)
```

```
0      3.35
1      4.75
2      7.25
3      2.85
4      4.60
      ...
296    9.50
297    4.00
```

Project 8. Car Price Prediction.ipynb

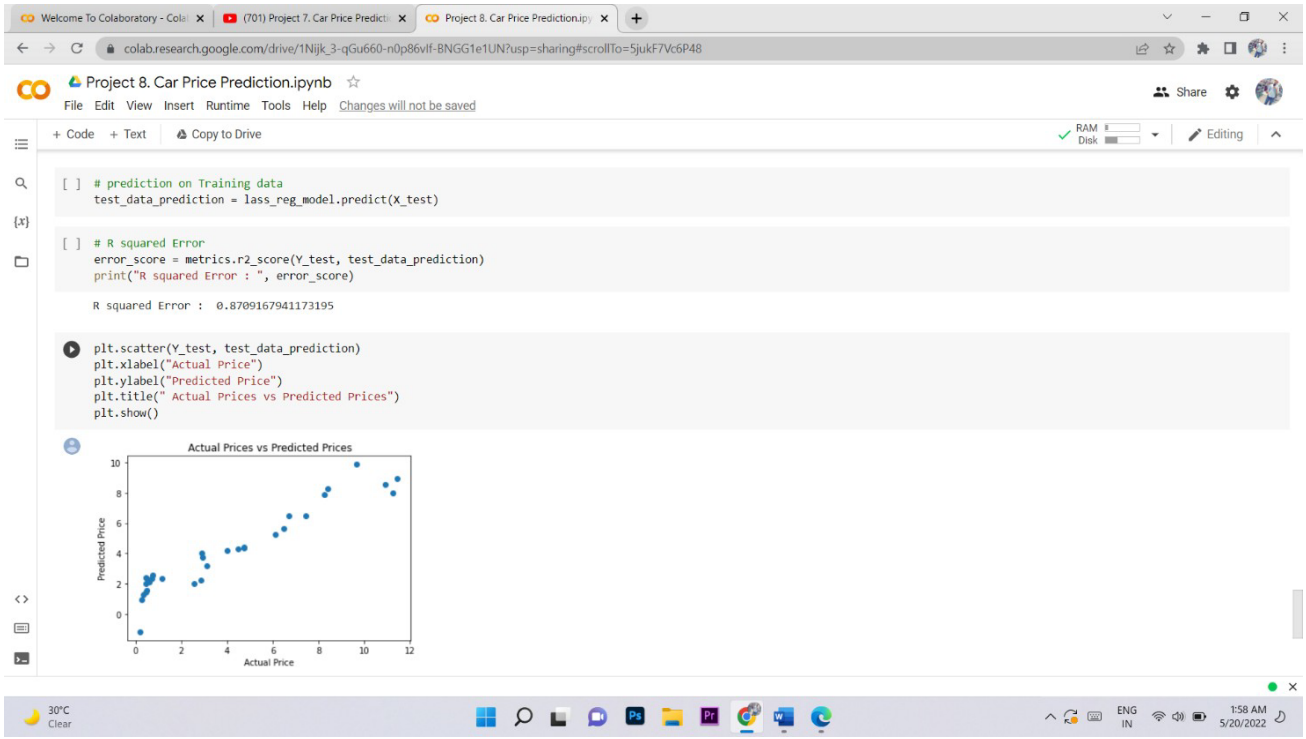
File Edit View Insert Runtime Tools Help Changes will not be saved

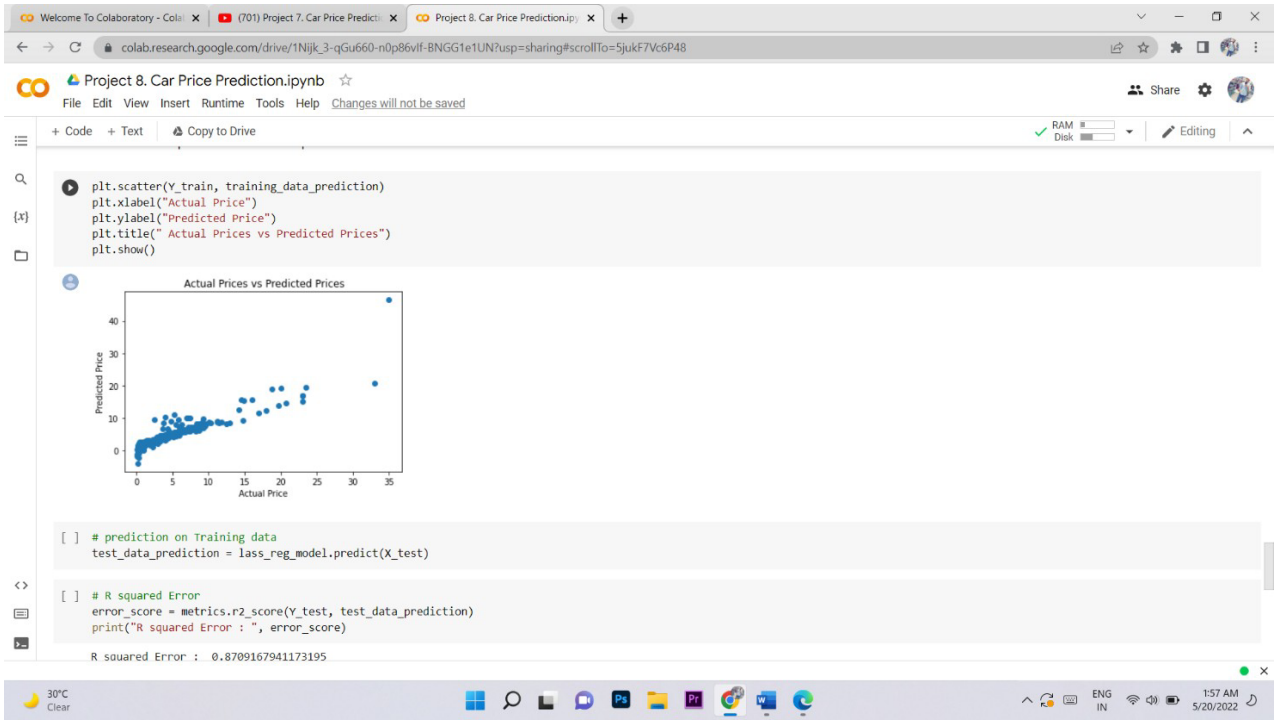
```
plt.scatter(Y_train, training_data_prediction)
plt.xlabel("Actual Price")
plt.ylabel("Predicted Price")
plt.title(" Actual Prices vs Predicted Prices")
plt.show()
```


```
[ ] # prediction on Training data
    test_data_prediction = lin_reg_model.predict(X_test)
```

```
[ ] # R squared Error
    error_score = metrics.r2_score(Y_test, test_data_prediction)
    print("R squared Error : ", error_score)
```

```
R squared Error :  0.8365766715027051
```


Welcome To Colaboratory - Cola
(701) Project 7. Car Price Predicti
Project 8. Car Price Prediction.ipy
colab.research.google.com/drive/1Nijk_3-qGu660-n0p86vIf-BNGG1e1UN?usp=sharing#scrollTo=5jukF7Vc6P48
Project 8. Car Price Prediction.ipynb
File Edit View Insert Runtime Tools Help Changes will not be saved
Share
+ Code + Text Copy to Drive
RAM Disk
Editing
[ ] # prediction on Training data
test_data_prediction = lass_reg_model.predict(X_test)
[ ] # R squared Error
error_score = metrics.r2_score(Y_test, test_data_prediction)
print("R squared Error : ", error_score)
R squared Error : 0.8709167941173195
plt.scatter(Y_test, test_data_prediction)
plt.xlabel("Actual Price")
plt.ylabel("Predicted Price")
plt.title(" Actual Prices vs Predicted Prices")
plt.show()
Actual Prices vs Predicted Prices
Predicted Price
Actual Price
30°C Clear
ENG IN
1:58 AM 5/20/2022


Welcome To Colaboratory - Cola
(701) Project 7. Car Price Predicti
Project 8. Car Price Prediction.ipy
colab.research.google.com/drive/1Nijk_3-qGu660-n0p86vIf-BNGG1e1UN?usp=sharing#scrollTo=5jukF7Vc6P48
Project 8. Car Price Prediction.ipynb
File Edit View Insert Runtime Tools Help Changes will not be saved
Share
+ Code + Text Copy to Drive
RAM Disk
Editing
plt.scatter(Y_train, training_data_prediction)
plt.xlabel("Actual Price")
plt.ylabel("Predicted Price")
plt.title(" Actual Prices vs Predicted Prices")
plt.show()
Actual Prices vs Predicted Prices
Predicted Price
Actual Price
[ ] # prediction on Training data
test_data_prediction = lass_reg_model.predict(X_test)
[ ] # R squared Error
error_score = metrics.r2_score(Y_test, test_data_prediction)
print("R squared Error : ", error_score)
R squared Error : 0.8709167941173195
30°C Clear
ENG IN
1:57 AM 5/20/2022