

- **What did you do to prepare the data?**

I downloaded the dataset from URL link but instead of using the .data file as default, I loaded the data into a CSV file so that the data type in the data file won't be all object type.

- **What insights did you get from your data preparation?**

During the data preparation, just by looking at the data type I would be able to know which data would be helpful for the classification and prediction.

Although I had an issue when I tried to convert horse power data type of object to the float since I got an error when I try to use .astype function however, ChatGPT did use two lines of code help me resolve the error. In addition, I used one hot-encoding on the origin variable as it's a categorical variable and drop the car name variable as a string categorical variable that's not necessary for the linear regression model prediction.

- **What procedure did you use to train the model?**

To train the data, I split the new dataset after conversion to test and train data by 30% as default. And then I fit the trained data into the fit the data to train model. Lastly, I use the mean\_squared\_error and r2\_score function to calculate the accuracy of train data and test data and find the fitting performance of the linear regression model.

- **How does the model perform to predict the fuel efficiency?**

In result of the calculations of the train data and test from the function, the accuracy score is about .8 and .82 which is very high that close to 1 and the mean squared error are 14.45 and 9.58 which is fine.

- **How confident are you in the model?**

In conclusion, the model fits very well with about 80% accuracy and shows that train data is slightly better than test data looks fair and reasonable. In conclusion, I am very confident in this model result.