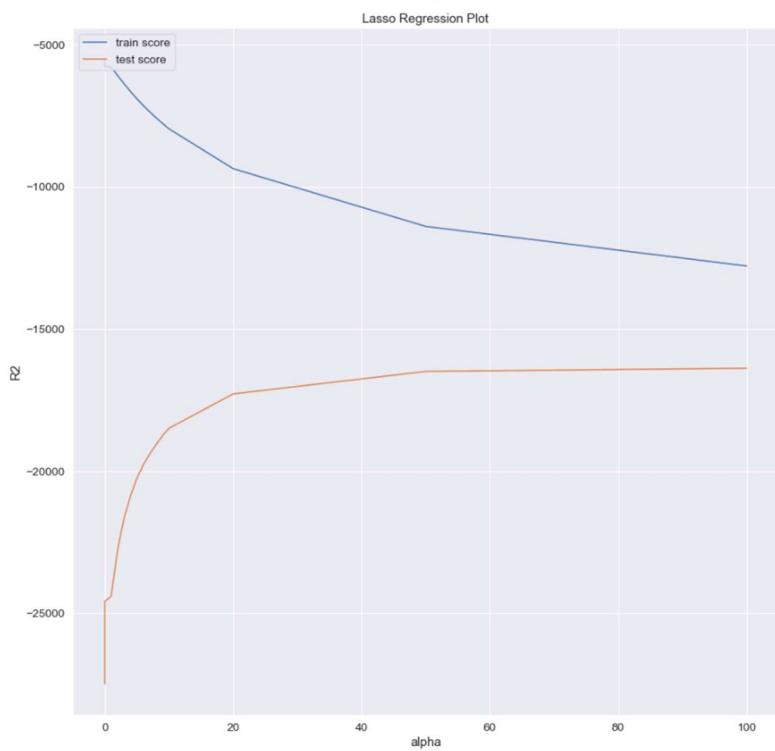
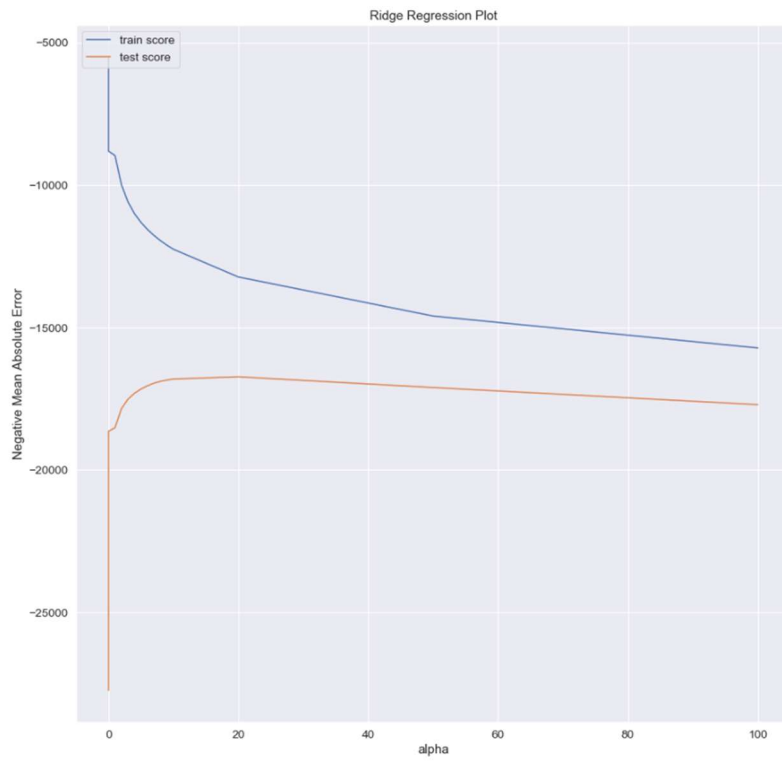


Assignment-based Subjective Questions:

Question 1: What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Answer 1: -

1. The optimal value of alpha for ridge and lasso are 20 and 100 respectively, based on the graph plotted with Alpha and Negative mean. Refer to the graph below.
2. Both models have changed the R2 value if we double the alpha value.
 - a. The R2 train data value for Ridge changed from 0.9280 to 0.9159.
 - b. The R2 train data value for Lasso changed from 0.9452 to 0.9265.



3. The most important variable after the change is implemented are

Based on the above table, leaving the categorial parameters

The top 10 important predictor variables are

1. GrLivArea
2. TotalBsmiSF
3. OverallQual
4. OverallCond
5. BSMIFinSF1
6. GarageCars
7. LotArea
8. PoolArea
9. Fireplaces
10. Garage Area

The top 5 predictor variables that are having negative impact are

1. Age
2. MSSubclass
3. BedroomAbvGr
4. BsmthalfBath
5. Mosold

Question 2: You have determined the optimal value of lambda for ridge and lasso regression during the assignment. Now, which one will you choose to apply and why?

Answer 2: In comparison with Lasso, ridge regression's R2 was greater. However, Lasso allows zero values to be assigned to irrelevant variables, so Lasso was selected for the second calculation.

Question 3: After building the model, you realised that the five most important predictor variables in the lasso model are not available in the incoming data. You will now have to create another model excluding the five most important predictor variables. Which are the five most important predictor variables now?

Answer 3:

The five most important variables to be excluded are:

1. GrLivArea
2. TotalBsmISF
3. OverallQual
4. OverallCond
5. BSMIFinSF1

The next five variables to be considered are:

1. GarageCars
2. LotArea
3. PoolArea
4. Fireplaces
5. Garage Area

Question 4: How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?

Answer 4:

1. To achieve a robust and generalizable model, we need to have a diverse data set with relevant features, proper regularization technique to be used, and cross-validation to check the unseen data with the developed model.
 2. The accuracy needs to be balanced with generalizability, not just focusing on high accuracy, the high accuracy model may not work well when applied to unseen data.
-