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 to double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Answer:

Alpha value for Ridge = 100, by doubling alpha not seen significant impact to predict vale for train and test and R2score very close to original value

Alpha value for Lasso = 0.01, by doubling alpha r2 score reduced and variation predict value

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:

I'll use Lambda of 0.01, as it will result in less computation compared to ridge value.

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:

1. Neighborhood_Crawfor
2. BsmtFullBath
3. TotRmsAbvGrd
4. GrLivArea
5. Condition1_Norm

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:

Depending on the curve of the predict we can able to identify the how generalised the model is.

And based on Test and train score, if both are comparably good rather than having large delta.

Then we can assume model is good and generic.