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

For lasso regression, value of alpha is 0.001, for ridge regression, value of alpha is 0.9.

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

After creating model in both Ridge and Lasso we can see that the r2 scores are almost same for both of them but as lasso will penalize more on the dataset and can also help in feature elemination i am gointg to consider that as my final model.

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.TotalBsmtSF

2.OverallQual

3. GrLivArea

4. GarageArea

5. OverallCond

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

More the model is robust, more it will learn. But to contradict this, the model should be as simple as possible. Accuracy may decrease though, but can be generalised easily.

We can make the model more robust by adding better data and doing proper EDA on it.