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:

Below is the optimal values for ridge and alpha is given below:

Ridge –Alpha= '20'

Lasso-Alpha=' 0.001'

When we double the values of alpha for Ridge and lasso that is 40 & 0.002 , some of the coefficients values of the features are changed and new features are introduced.For Ridge OverallQual feature is added.For Lasso RoofStyle_Shed is added.

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:

As per our analysis from the model , we have chosen the the Lasso regression as our final model , because when we see the values of the coefficient , it gets clamp down to zero which are irrelevant which makes the model more sparse. As we increase the value of alpha we see more coefficients becomes zero which helps in selecting the best features for prediction

Question 3

After building the model, you realized 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:

After removing the five most important feature from the data set the alpha value gets changed to 0.0001 and the best score is also changed from 0.9143561099481419 to 0.8875739717820892.Now the five most important Predictors varaibles are :

1. BsmtFinSF1 0.349783
2. GarageCars 0.237725
3. GarageType_BuiltIn 0.186798
4. BsmtFinSF2 0.133765
5. GarageType_No Garage 0.124928

Question 4

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

Answer:

The Lasso model is more robust and generalizable because it balances the bias variance trade off. The model drops the irrelevant coefficient thus making the model not becoming too complex and more generalizable and robust. As per the graphs it is clearly visible that the accuracy model gets dropped for both the test and train data as the alpha value is increased and the error term is increased and regularization term and as we decrease the value of alpha opposite will happen.