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 :

Optimal value of alpha is 0.3,100 respectively for the ridge and lasso regression.

If we double the alpha value in Ridge regression the R2 score on train and test is 0.8566020762625707 , 0.8315642268655783 this is a good for a lasso regression the R2 score on train and test is 0.8240162674141438,0.8435397218703375 this is very rare situation that R2 score of test set is greater than train set. The important predictor of both Ridge regression and Lasso regression is **1stFlrSF.**

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 :

Value of lambda is 2 after the hit and try. From 5 to 2. We have a good fit model on the lambda 2 because in lasso regression we have a lowest different in R2 Score of train and test data.

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 : TotalBsmtSF, LotArea **,** OverallQual **,** 2ndFlrSF**,** BsmtFinSF1

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 : There is less difference in R2 score of train and test data such in lasso regression model such that we can say that model is robust and generalisable. If model is underfit or ove rfit we can get huge difference in in R2 score of train and test data