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:

The optimal value of alpha for ridge and lasso regression.

Ridge Alpha 1

Lasso Alpha 10

Ridge Regression: When alpha value is doubled, R2score on training data has decreased but it has increased on testing data.

Lasso Regression: R2score of training data has decrease and it has increase on testing data

After change is Ridge and Lasso alpha, The predictor variables are same, but coefficients of predictor variables are changed.

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:

Lasso regression will be best fit to solve this problem as r2_score is higher than ridge.

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:

The new 5 most predicting variables are,

2ndFlrSF

KitchenAbvGr

GarageArea

Condition2_PosN

RoofMatl_CompShg

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:

To make sure model is robust and generalisable and robust, below things should be considered,

1. Test accuracy should be less than training score.
2. The model should show accurate result on data other that training data.