

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 for Ridge is 500 and Lasso is 1000

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

Will apply Lasso, since it makes variables which not useful to almost 0 and makes the equation simple

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

GrLivArea, YearBuilt, YearRemodAdd, GarageArea, Fireplaces

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

Identify the variables which are not suitable to predict the target variable.

Clean the data, remove the invalid values.

Apply the Lasso and Ridge regression models.

Compare the resultant co-efficients.

Compare the values of RSS, R^2 & MSE. Chose the model which has similar values for RSS for test and training data. Also R^2 and MSE should be low.