

### 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 are:-

a). Ridge – 4.0 b). Lasso – 0.01

If we double the alpha for Ridge and Lasso Regression, the model under fits the data and the test metric reduces.

The most prominent features after the change is implemented remain the same.

### 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 :-

I will chose to apply the lasso regression because it is able to generalize better as compared to ridge regression primarily because the model does feature selection by making the coefficients of some features 0.

### 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 3

The five most important predictor variables after the change is implemented are:- 1. LowQualFinSF 2. BsmtCond 3. GarageArea 4. Functional 5. TotRmsAbvGrd

#### 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 4

We can ensure that the model is robust and generalisable by ensuring:-

1. The training data is representative of the population
2. Having enough training data
3. Making the model less complex by doing feature selection
4. Applying regularization so that the model does not learn patterns specific to training data If the model is robust and generalisable, It will have maximum possible performance on the data it has never seen before and the goal of machine learning is not to perform well on training data but to perform well on testing data or the data the model has never seen before.