

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 for Ridge Regression is - 100

The optimal value for Ridge Regression is – 0.01

If the alpha value is doubled, not much difference in the accuracy of the algorithm is seen. I think the accuracy will affect if we choose the alpha from different order, but doubling the alpha keeps the order same hence not affecting the accuracy a lot.

The most important predictor variable remains GrLivArea in both cases.

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 –

The optimal value for Ridge Regression is - 100

The optimal value for Ridge Regression is – 0.01

These values are chosen because they lead to better accuracy on the test set.

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 – Five most important predictor variables for Lasso Regression –

1. GrLivArea
2. OverallQual
3. YearBuilt
4. TotalBsmtSF
5. 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 –

To make sure model is generalizable we must make sure that we are not over fitting the model. The lambda values for Ridge and Lasso regression can be helpful in making sure the predictor coefficients are adjusted correctly to reduce the effect of overfitting. This might show some toll on the accuracy on trading data, but model would perform better on the unseen data.