

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 –

1. Optimal value of alpha for ridge and lasso regression
 - a. Optimal Value of Alpha in Ridge Regression – 10
 - b. Optimal Value of Alpha in Lasso Regression – 0.001
2. The most important predictor variables are -
 - a. OverallQual_9
 - b. GrLivArea
 - c. OverallQual_8
 - d. Neighborhood_Crawfor
 - e. Functional_Typ
 - f. Exterior1st_BrkFace
 - g. CentralAir_Y
 - h. Neighborhood_Somerst
 - i. TotalBsmtSF
 - j. Condition1_Norm
3. Changes in the model if you choose double the value of alpha for both ridge and lasso

Value of R2

Regression	Lasso	Ridge
R2 Score (Train)	0.92	0.94
R2 Score (Test)	0.93	0.93

With Double Value of R2

Regression	Lasso	Ridge
R2 Score (Train)	0.91	0.93
R2 Score (Test)	0.91	0.93

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 –

Here in the given case study many variables has been provided used hence I will prefer to use Lasso Regression, as Lasso also helps in the feature selection. However in case of few feature we can use Ridge Regression as just reduction of coefficients will be the objective

Question 3

After building the model, you realized 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 –

After dropping 5 top variables observed following 5 most important predictors -

1. 2ndFlrSF (Second floor square feet)
2. Exterior1st_BrkFace (Exterior covering on house is Brick Face)
3. 1stFlrSF (First Floor square feet)
4. MSSubClass_70 (2-STORY 1945 & OLDER)
5. Neighborhood_Somerst (Physical locations within Ames city - Somerset)

Question 4

How can you make sure that a model is robust and generalizable? What are the implications of the same for the accuracy of the model and why?

Answer –

1. Models needs to be so perfect and robust it should work with the unseen various type of data.
2. Model should not be too complex so that should not overfit
3. Model needs to work perfectly with train and test data.