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 Lambda:
Ridge : 10
Lasso : 0.001

On doubling alpha the metrics is as follows:

Ridge:

Training R2 Score: 0.9388401207889624
Testing R2 Score : 0.9009241328601169
Adj. RSquared Train: 0.9184534943852832
Adj. RSquared Test: 0.7621090442850058
MSE Train: 0.0076336268384604375
MSE Test: 0.012184835885558852
Train_RSS : 7.793933002068107
Test_RSS : 5.336958117874777

Lasso:

Training R2 Score: 0.9197874665599054
Testing R2 Score : 0.8892200181302384
Adj. RSquared Train: 0.8930499554132072
Adj. RSquared Test: 0.7340063072687593
MSE Train: 0.010011670329438881
MSE Test: 0.013624265297445483
Train_RSS : 10.221915406357098
Test_RSS : 5.967428200281121

Top Predictor Variables after doubling Lambda :

Ridge:

('MSSubClass', 0.064),
('LotArea', 0.066),
('OverallQual', 0.093),
('OverallCond', 0.068),
('YearBuilt', 0.046),
('YearRemodAdd', 0.042),
('MasVnrArea', 0.041)

Lasso:

('MSSubClass', 0.081),
 ('LotArea', 0.054),
 ('OverallQual', 0.052),
 ('OverallCond', 0.106),
 ('YearBuilt', 0.053)]

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:

Since, we have a huge number of predictors, hence Lasso model is more appropriate in this case as it helps in eliminating features by pushing model coefficients towards 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:

Top predictors after removing the top predictors from original Lasso model:

[('constant', 11.887),
 ('MSSubClass', 0.079),
 ('LotArea', 0.054),
 ('OverallQual', 0.104),
 ('OverallCond', 0.083),
 ('YearBuilt', 0.067),
 ('YearRemodAdd', 0.063)]

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:

Occam's Razor states that a model should be as simple as possible without compromising on the quality of the model. Simpler models are always preferred for their ease of interpretation and robustness.
Advantages of simpler models:

• Generalisability: Simpler models have a better chances of being applicable in a wider sense and as model complexity increases, it's specificity to handle the particular dataset increases.

• Robustness: Simpler models are robust and indifferent to changes in the data set while complex model can alter their form very quickly and significantly when met with a different set of data.

• Need for fewer datapoints: Simpler models need fewer datapoints to train while complex models need more data. In real life, the amount of data can be limited and therefore, simpler models become the go to choice.

To make a model robust and generalisable, it is important to make simple models as possible without compromising with its prediction quality.

Regularisation is one of the ways to simplify the model, by pushing the coefficients of the model towards zero and eliminating the number of predictor variables by making the coefficients zero.
Since model complexity is depends on the number of predictor variables and the value of coefficients of the predictor variables, Regularisation simplifies the model and makes it robust and generalisable.