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 alpha is
Ridge Regression : 5.0
Lasso Regression : 0.0007

Double the value of alpha

Lasso: 5.0X2= 10

Ridge: .0007X2=.0014

MSE before:
Mean Square Error for Lasso Regression: 0.11833529463591849
Mean Square Error for Ridge Regression: 0.11795124311878702
MSE after doubling:

Ridge: MSE score: 0.10651634693528603

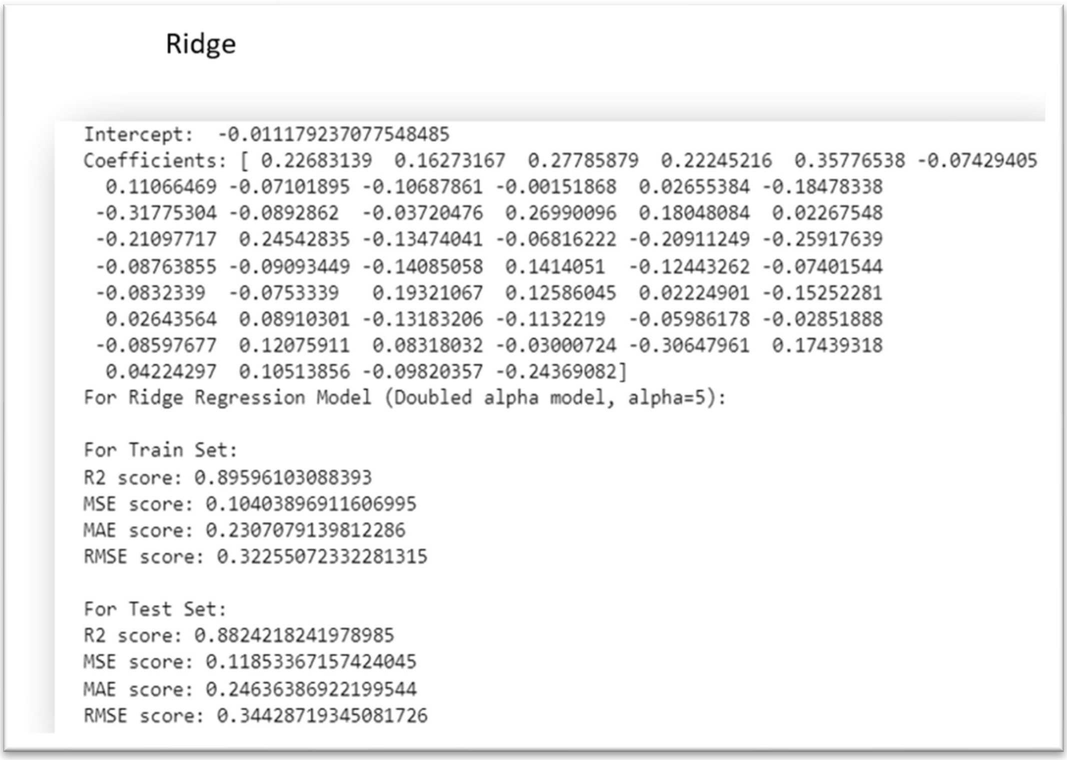
Lasso: MSE score: 0.10669838522935521
# More details in notebook.

Top 10 features after change implemented:

Lasso		Ridge	
0	Exterior1st_BrkComm	0	GrLivArea
1	GrLivArea	1	YearBuilt
2	MSSubClass_160	2	MSSubClass_160
3	MSZoning_FV	3	OverallQual
4	YearBuilt	4	MSZoning_FV
5	Neighborhood_Crawfor	5	GarageType_No Garage
6	GarageType_No Garage	6	Neighborhood_Crawfor
7	OverallQual	7	TotalBsmtSF
8	TotalBsmtSF	8	MSZoning_RL
9	Neighborhood_StoneBr	9	Neighborhood_MeadowV

Q2: 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: To choose one of it, let’s see the comparison in scores between the two:

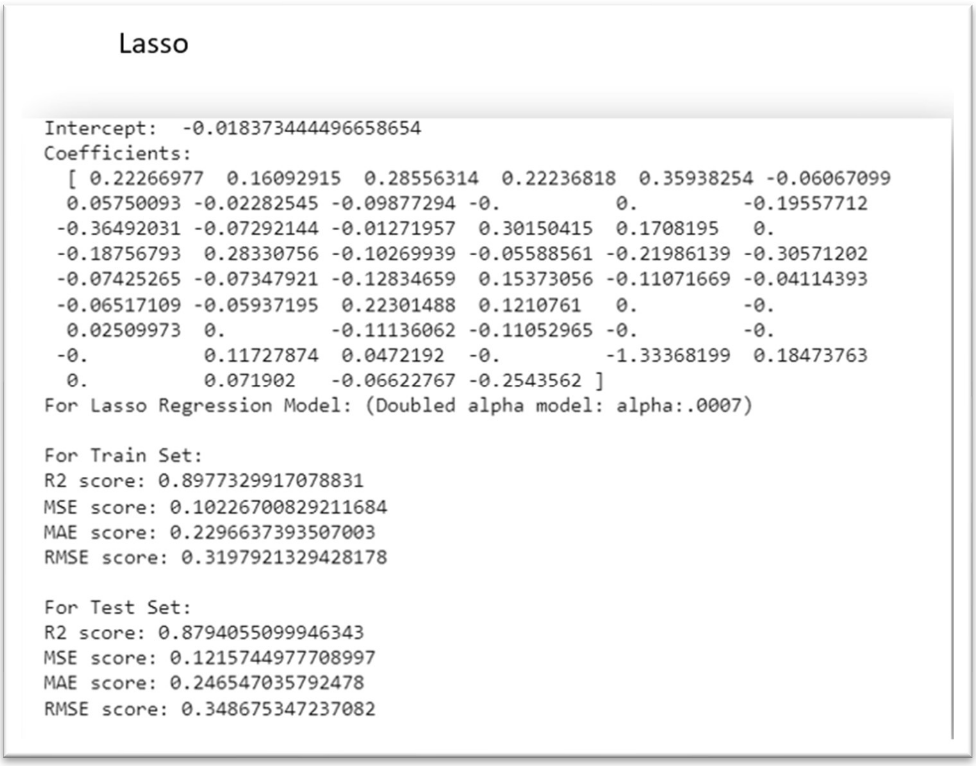


Ridge

```
Intercept:  -0.011179237077548485
Coefficients: [ 0.22683139  0.16273167  0.27785879  0.22245216  0.35776538 -0.07429405
  0.11066469 -0.07101895 -0.10687861 -0.00151868  0.02655384 -0.18478338
 -0.31775304 -0.0892862  -0.03720476  0.26990096  0.18048084  0.02267548
 -0.21097717  0.24542835 -0.13474041 -0.06816222 -0.20911249 -0.25917639
 -0.08763855 -0.09093449 -0.14085058  0.1414051  -0.12443262 -0.07401544
 -0.0832339  -0.0753339   0.19321067  0.12586045  0.02224901 -0.15252281
  0.02643564  0.08910301 -0.13183206 -0.1132219  -0.05986178 -0.02851888
 -0.08597677  0.12075911  0.08318032 -0.03000724 -0.30647961  0.17439318
  0.04224297  0.10513856 -0.09820357 -0.24369082]
For Ridge Regression Model (Doubled alpha model, alpha=5):

For Train Set:
R2 score: 0.89596103088393
MSE score: 0.10403896911606995
MAE score: 0.2307079139812286
RMSE score: 0.32255072332281315

For Test Set:
R2 score: 0.8824218241978985
MSE score: 0.11853367157424045
MAE score: 0.24636386922199544
RMSE score: 0.34428719345081726
```



Lasso

```
Intercept:  -0.018373444496658654
Coefficients:
  [ 0.22266977  0.16092915  0.28556314  0.22236818  0.35938254 -0.06067099
  0.05750093 -0.02282545 -0.09877294 -0.          0.         -0.19557712
 -0.36492031 -0.07292144 -0.01271957  0.30150415  0.1708195   0.
 -0.18756793  0.28330756 -0.10269939 -0.05588561 -0.21986139 -0.30571202
 -0.07425265 -0.07347921 -0.12834659  0.15373056 -0.11071669 -0.04114393
 -0.06517109 -0.05937195  0.22301488  0.1210761   0.         -0.
  0.02509973  0.         -0.11136062 -0.11052965 -0.         -0.
 -0.          0.11727874  0.0472192  -0.         -1.33368199  0.18473763
  0.          0.071902   -0.06622767 -0.2543562 ]
For Lasso Regression Model: (Doubled alpha model: alpha:.0007)

For Train Set:
R2 score: 0.8977329917078831
MSE score: 0.10226700829211684
MAE score: 0.2296637393507003
RMSE score: 0.3197921329428178

For Test Set:
R2 score: 0.8794055099946343
MSE score: 0.1215744977708997
MAE score: 0.246547035792478
RMSE score: 0.348675347237082
```

Observation and Conclusion:

We see that R2 score and MSE score are slightly better in ridge regression when seeing the training set, whereas Lasso model is performing better in score with test data. Also Lasso provides simpler model with feature selection, Hence I would choose to apply Lasso Model in this case.

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:

The top5 important predictor variables after dropping the previous top 5:
['Neighborhood_BrDale', 'Neighborhood_MeadowV', 'Exterior2nd_Brk Cmn', 'Neighborhood_IDOTRR', 'MSSubClass_45']

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:

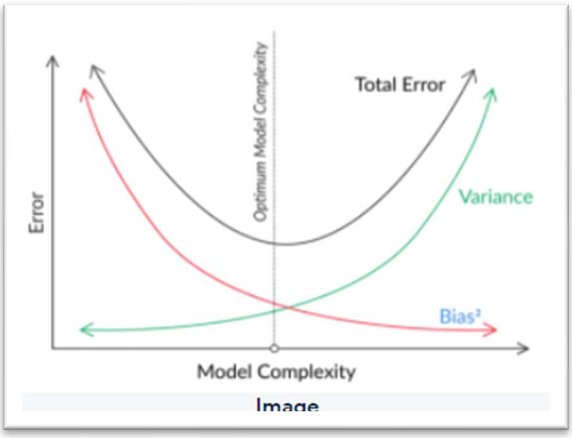
Robust and generalisable

- A robust model should have low variance and bias across different folds of cross-validation, similar or better performance on validation and test sets than on the training set. When we plot it a smooth converging learning curve should come. Also it is robust if its highly consistent values for performance metrics across different datasets.
- A model is generalizable when we train a model on a dataset, and the model is provided with new data absent from the trained set, it may perform well.
- To make sure a model is robust and generalizable, we have to take care it doesn't overfit. Because there's a very convenient way to measure an algorithm's generalization performance: we measure its performance on a test set, consisting of examples it hasn't seen before. If an algorithm works well on the training set but fails to generalize, we say it is overfitting.

Accuracy

- A very complex models has high accuracy but with high variance. To make it more robust and generalisable one should decrease variance which will cause the increase in bias.

- Hence we need to make a balanced model with variance and bias to get consistent accuracy. In below graph we see a mid point where variance and bias cross, that is the point of optimal model complexity.


Optimum Model Complexity
Total Error
Variance
Error
Bias²
Model Complexity
Image