Question 1

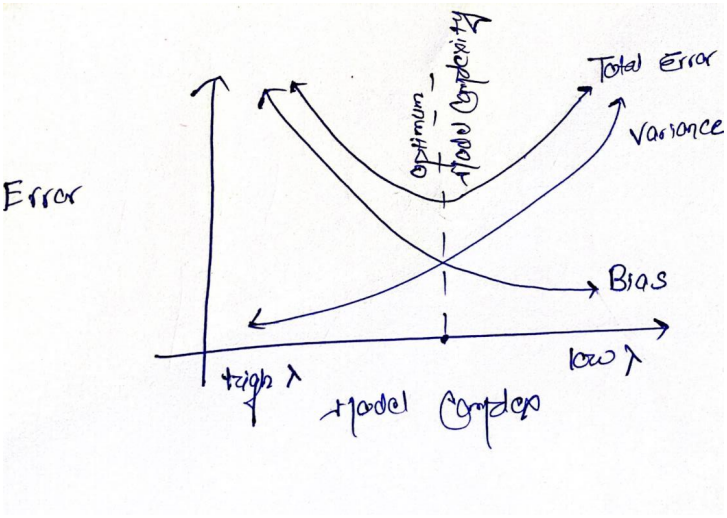
What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Solution : The optimal lambda value in case of Ridge and Lasso is as below:

Ridge - 0.4

Lasso - 0.0001

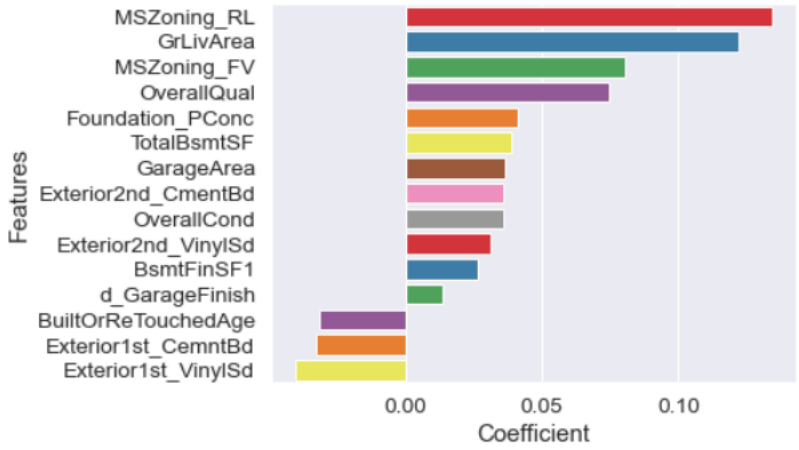
When we double the value of alpha , we do see the penalty will increase for the model . This result is a model being more simpler. A higher value of alpha would means we move towards left of the optimum model complexity on the graph below . This leads to decrease in model complexity , more bias and with lesser variance.



Important predictor after making the change:

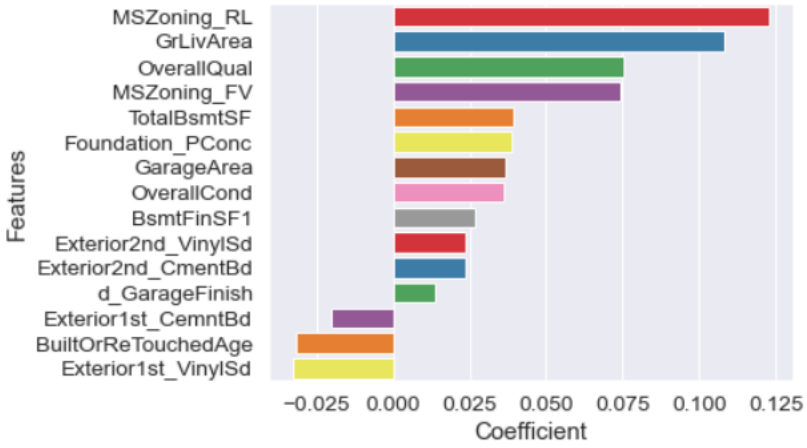
For Ridge :

MSZoning_RL
GrLivArea
MSZoning_FV
OverallQual
Foundation_PConc
TotalBsmtSF
GarageArea
Exterior2nd_CmentBd
OverallCond
Exterior2nd_VinylSd



For lasso :

MSZoning_RL
GrLivArea
OverallQual
MSZoning_FV
TotalBsmtSF
Foundation_PConc
GarageArea
OverallCond
BsmtFinSF1
Exterior2nd_VinylSd
Exterior2nd_CmentBd



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?

Solution : We choose Lasso . This is because :

- The Mean Squared error in case of Ridge and Lasso are:

  Ridge -

  - MSE: 0.015888159710745184
  - RMSE : 0.1260482435845307

  Lasso -

  - MSE : 0.01578521041943166
  - RMSE : 0.125639207333665

  The Mean Squared Error of Lasso is slightly lower than that of Ridge .

- Also, since Lasso helps in feature reduction (as the coefficient value of one of the feature became 0), Lasso has a better edge over Ridge.

- Ridge includes all variables in final model . In contrast, the LASSO does both parameter shrinkage and variable selection automatically.
- Our business scenario is such that we need to have a few variables to be presented with a huge impact, also in such a scenario we go with lasso.

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?

Solution :

Top 5 predictors are now as follow .

From Lasso :

2ndFlrSF
1stFlrSF
TotalBsmtSF
OverallCond

BsmtFinSF1

From Ridge :

2ndFlrSF
1stFlrSF
TotalBsmtSF
OverallCond
Exterior2nd_VinylSd

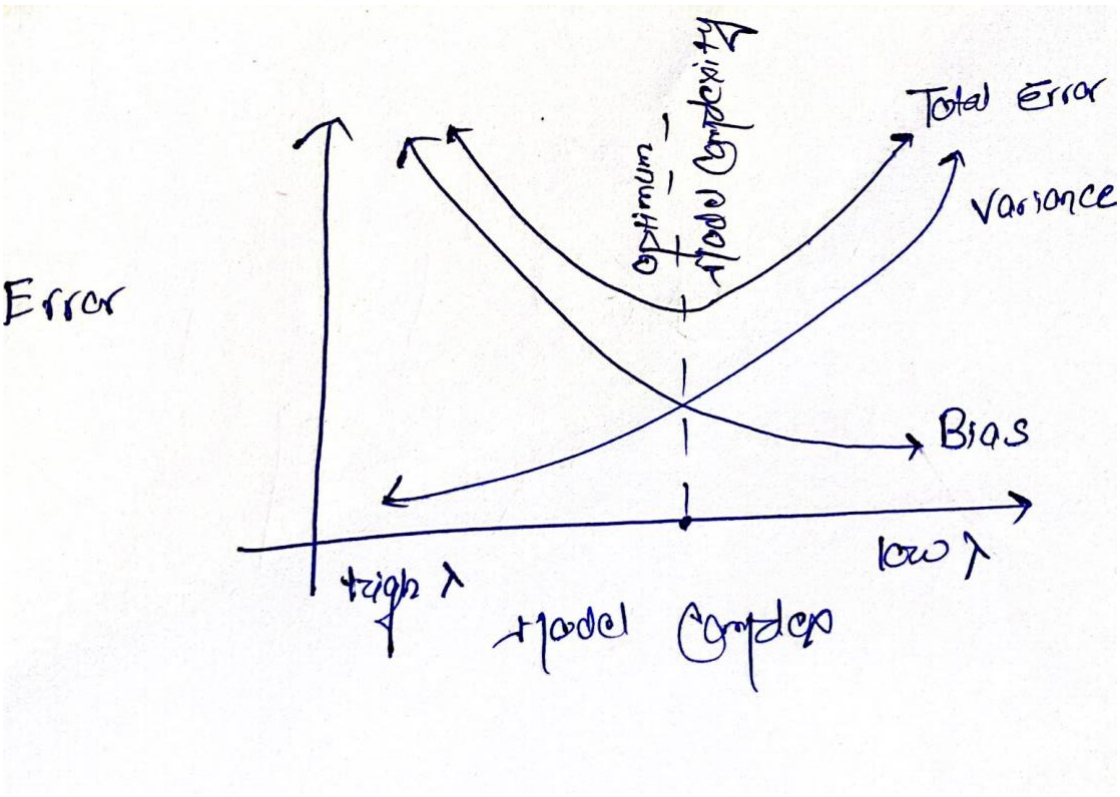
We will choose Lasso ones as the final one .

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?

Solution :

- We can make sure by trying to maintain a sweet spot between Bias and Variance .
- As per bias variance trade off we know, the simpler the model , more general and robust it is but with higher bias and low variance . Although it does take a hit on accuracy on training data but it is not tightly coupled on train data points , so its more robust. Hence , the implication is such that a simpler model does not tend to memorize the whole set data points and performs generally well on both training and test data .
- On the contrary , complex models are less general but with low bias and high variance .
- Making a model too simple will lead to underfitting , more complex to overfitting . So , We use Regularization technique to make a model optimum simpler. It allows to maintain that optimum sweet spot as shown below . The dotted line shows the sweet spot , which Regularisation allows to achieve . It allows to get the optimum value of alpha and attain the sweet spot where model is simple and robust and balance between bias and variance as shown below .


Error
Total Error
Variance
Bias
optimum
Model Complexity
high λ
low λ
Model Complex