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?

Ans :- Ridge regression :- While building model using ridge regression it was noticed that the train error is showing increase trend when value of alpha increases. In case of test error is minimum so taken alpha as 2 for building the model.

Lasso regression :- Taking alpha 0.01 gives an optimal model. It add penalty to the feature bringing it down to 0.

When we double the value of alpha for our ridge regression no we will take the value of alpha equal to 10 the model will apply more penalty on the curve and try to make the model more generalized that is making model more simpler and no thinking to fit every data of the data set .from the graph we can see that when alpha is 10 we get more error for both test and train.
Similarly when we increase the value of alpha for lasso we try to penalize more our model and more coefficient of the variable will reduced to zero, when we increase the value of our r2 square also decreases.
The most important variable after the changes has been implemented for ridge regression are as follows:-

MSZoning_FV
MSZoning_RL
Neighborhood_Crawfor
MSZoning_RH
MSZoning_RM
SaleCondition_Partial
Neighborhood_StoneBr
GrLivArea
SaleCondition_Normal
Exterior1st_BrkFace

The most important variable after the changes has been implemented for lasso regression are as follows:-

GrLivArea
OverallQual

OverallCond

TotalBsmtSF

BsmtFinSF1

GarageArea

Fireplaces

LotArea

LotFrontage

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?

Ans :- It is best to use Lasso regression for building the model as it will bring the down feature data point to 0 and act as a feature selection attribute.

The mean square error achieved by Lasso regression is better so priority is given to Lasso regression over Ridge regression.

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?

Ans :- The 5 most important predictor variables that will be excluded are :-

GrLivArea
OverallQual
OverallCond
TotalBsmtSF
GarageArea

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?

Ans :-

Per, Occam’s Razor— given two models that show similar ’performance’ in the finite training or test data, we should pick the one that makes fewer on the test data due to following reasons:-

- Simpler models are usually more ’generic’ and are more widely applicable
- Simpler models require fewer training samples for effective training than the more complex ones and hence are easier to train.
- Simpler models are more robust.

  Complex models tend to change wildly with changes in the training data set

  Simple models have low variance, high bias and complex models have low bias, high variance

- Simpler models make more errors in the training set. Complex models lead to overfitting — they work very well for the training samples, fail miserably when applied to other test samples

Therefore, to make the model more robust and generalizable, make the model simple but not simpler which will not be of any use.

Regularization can be used to make the model simpler. Regularization helps to strike the delicate balance between keeping the model simple and not making it too naive to be of any use. For regression, regularization involves adding a regularization term to the cost that adds up the absolute values or the squares of the parameters of the model.

Also, Making a model simple leads to Bias-Variance Trade-off:

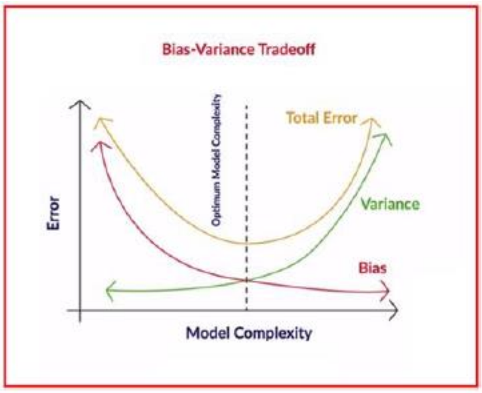
- A complex model will need to change for every little change in the dataset and hence is very unstable and extremely sensitive to any changes in the training data.
- A simpler model that abstracts out some pattern followed by the data points given is unlikely to change wildly even if more points are added or removed.

Bias quantifies how accurate is the model likely to be on test data. A complex model can do an accurate job prediction provided there is enough training data. Models that are too naïve, for e.g.,

one that gives same answer to all test inputs and makes no discrimination whatsoever has a very large bias as its expected error across all test inputs are very high.

Variance refers to the degree of changes in the model itself with respect to changes in the training data.

Thus, accuracy of the model can be maintained by keeping the balance between Bias and Variance as it minimizes the total error as shown in the below graph


Bias-Variance Tradeoff
Optimum Model Complexity
Total Error
Variance
Bias
Error
Model Complexity