Problem Statement - Part II
(subjective Assignment)

By- Rishabh Tripathi
email id -rishabh.tripathi.19041@iitgoa.ac.in

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 ; So initially we use alpha as 10^{-15} ,10^{-10},10^{-8},10^{-4},10^{-3},10^{-2},1 ,5 ,10 ,20 ,50 we get optimal value of alpha for ridge regression is 50 . similarly optimal value for lasso regression is 50. If we choose double value of alpha for both ridge and lasso regression than our respected alpha value become as $2x10^{-15}$,$2x10^{-10}$,$2x10^{-8}$,$2x10^{-4}$,$2x10^{-3}$,$2x10^{-2}$,2 ,10 ,20 ,40 ,100.Than we get that our most optimal value of alpha is 40 for both ridge and lasso regression. Now our most important predictor variable is states as -
'MSZoning' , 'LandContour' , 'LotConfig' , 'Neighborhood' ,'Condition1' , 'Condition2 ,
'BldgType' , 'HouseStyle' ,'RoofStyle' , 'RoofMatl' , 'Exterior1st' , 'Exterior2nd' ,'MasVnrType' ,
'ExterQual' , 'ExterCond' , 'Foundation' ,'BsmtQual' , 'BsmtCond' , 'BsmtExposure' ,
'HeatingQC' ,'Electrical' , 'KitchenQual' , 'Functional' , 'GarageType' ,'GarageFinish' ,
'GarageQual' , 'GarageCond' , 'SaleCondition' ,'MSSubClass' , 'LotFrontage' , 'LotArea' ,
'OverallQual' , 'OverallCond' ,'YearBuilt' , 'YearRemodAdd' , 'MasVnrArea' , 'TotalBsmtSF' ,
'2ndFlrSF' ,'LowQualFinSF' , 'GrLivArea' , 'BsmtFullBath' ,'BsmtHalfBath' ,'FullBath',
'HalfBath' , 'BedroomAbvGr' , 'KitchenAbvGr' , '1stFlrSF' , 'TotRmsAbvGrd' ,'Fireplaces' ,
'GarageYrBlt' , 'GarageCars' , 'GarageArea' , 'WoodDeckSF' ,'OpenPorchSF' , 'EnclosedPorch'
'3SsnPorch' , 'ScreenPorch' , 'PoolArea' ,'MiscVal' , 'MoSold' , 'YrSold'.

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 ;We determined the optimal value of lambda is 50 for both ridge and lasso regression .so now i will use lasso regression to predict the desired outcome because Lasso regression have many advantage over ridge regression as it not only overcome Overfitting problem along with in many case it give more r_2score value in compare to ridge regression r_2score values. So as we know that in regression problem the model which r_2score more they are more accurate in predicting the desired result. And Lasso regression use most important predictor variable .For giving result. Lasso

has an advantage of completely reducing unnecessary parameters(feature) in the given model .So that why i will use Lasso regression in my model.

Question 3;

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 ; when we get low variance and low bias than we can say our model is generalisable along with this when we get almost r_2 score approximately 0.9 or above this. And our root mean square error nearer to zero. Than we can say that our model is robust. So when these satisfy we can say our model is robust and generalisable .So if our mode is robust and generalisable than it indicates accuracy should be more than 90 %.unless its not use full for any work completely for company because company want benefit which can be achived by using correct model with high accuracy.