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 of Regression model is as follows:-
Ridge Regression - 12
Lasso Regression - 0.001

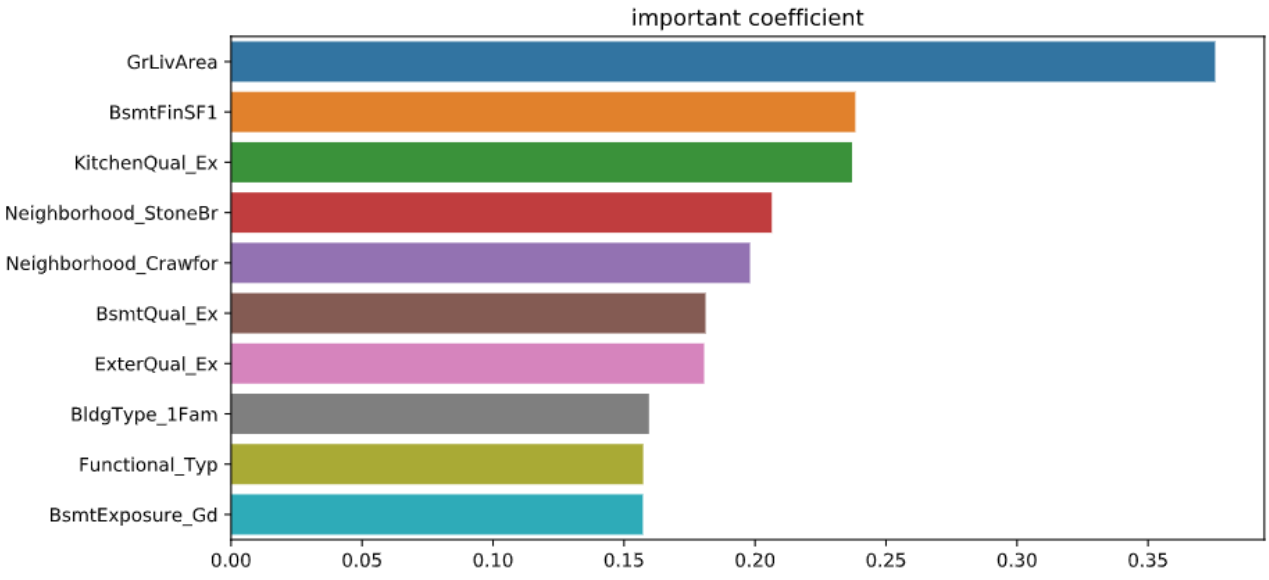
When we double the value of alpha ,the model is more penalised for increase in coefficient value, so the value of coefficients is less than the original values predicted for optimal hyper parameter.

In Ridge more number of value will be nearing zero and in lasso more value would have been eliminated by getting a coefficient of exact zero.

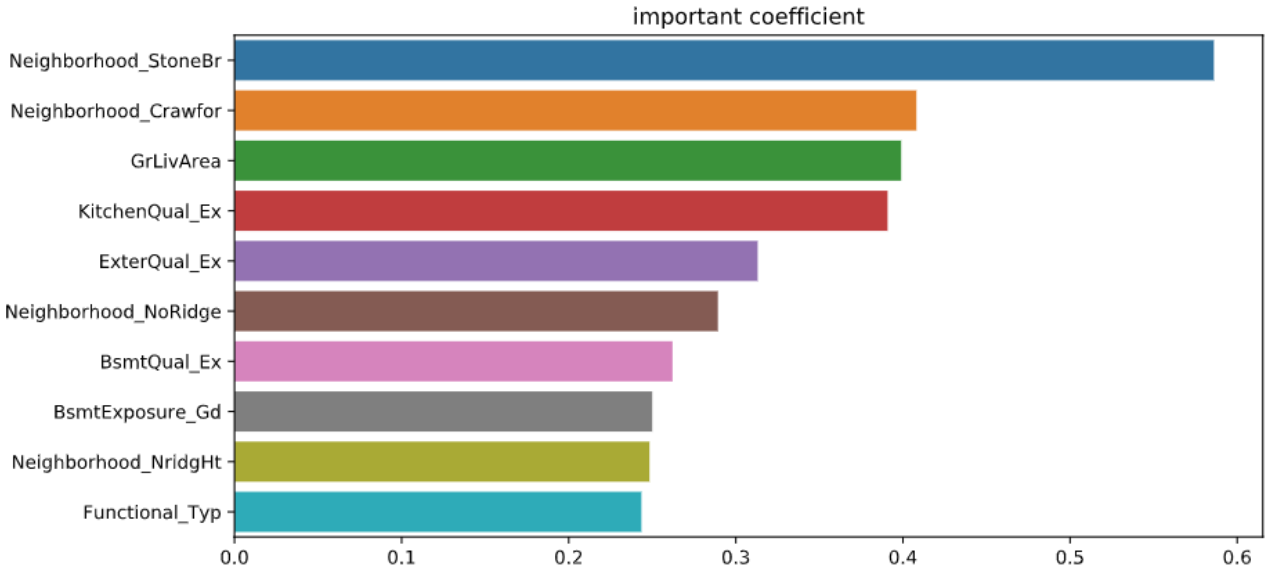
The training and test accuracy decreases for both ridge and lasso regression.

It is observed that there is are little changes in order of importance of the variable.

Important variable calculated after change in optimal value in ridge regression: - (ordered in descending order)



Important variable calculated after change in optimal value in Lasso regression: - (ordered in descending order)



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:
Each regularization method has their pros and cons,

Ridge regression can be used to treat multicollinearity and it is fast since it is implemented into the formula itself and no iterative process needed.

But it doesn't perform feature elimination as lasso does. Mostly all variables are retained since features coefficient may near 0 but not exactly 0.

Lasso on the other hand can perform feature elimination by making the coefficient of features exactly 0 .

But it is computationally intensive. Accuracy of lasso will comparatively be lesser to the Ridge since features are eliminated.

for the given problem I would choose Lasso regression over Ridge regression.

This is because both the upon implementing two models , they both have almost same accuracy in both training and testing data set with negligible differences.

While in lasso this accuracy is achieved by eliminating over 154 less relevant variable , In ridge regression most of the variables are retained.

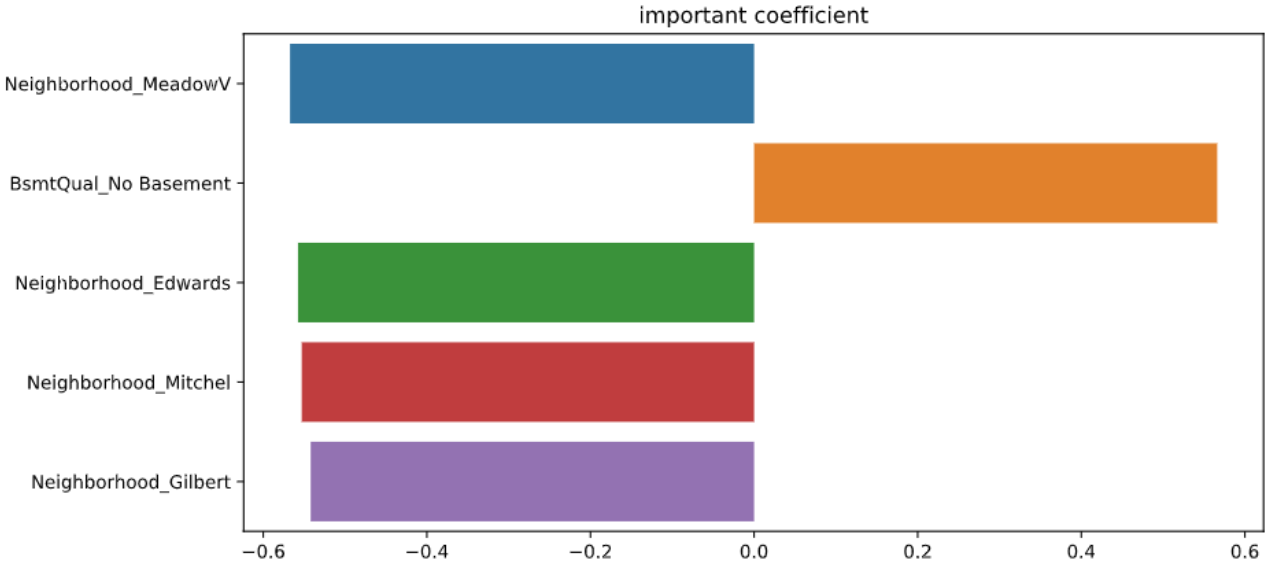
while Ridge has a computational advantage over lasso , the processing time difference is too small to be noticeable in this assignment.

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?

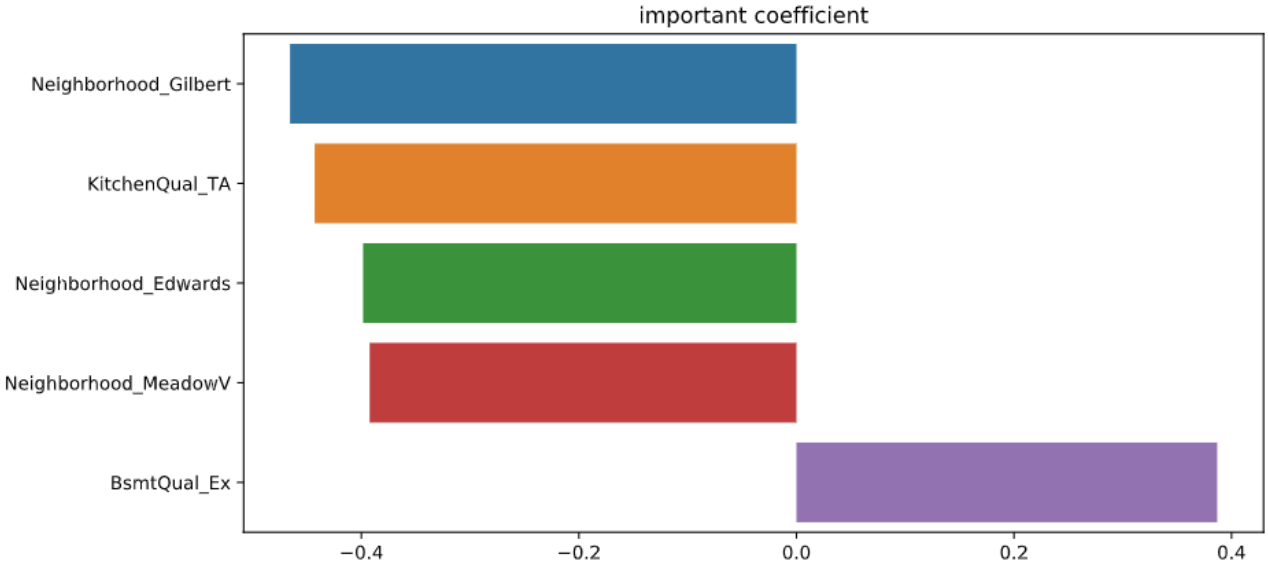
Answer:
while removing a top 5 Features , it is not necessary that next immediate top 5 feature will be available as present top 5 feature.

When a predictor variable is removed , the new model will try to predict the output value with less than optimal feature set.Here the importance given to new feature sets may change in any direction to predict the output value as close as possible.

The next important five features in the list are as follows Lasso Regression,



The next important five features in the list are as follows Ridge Regression,



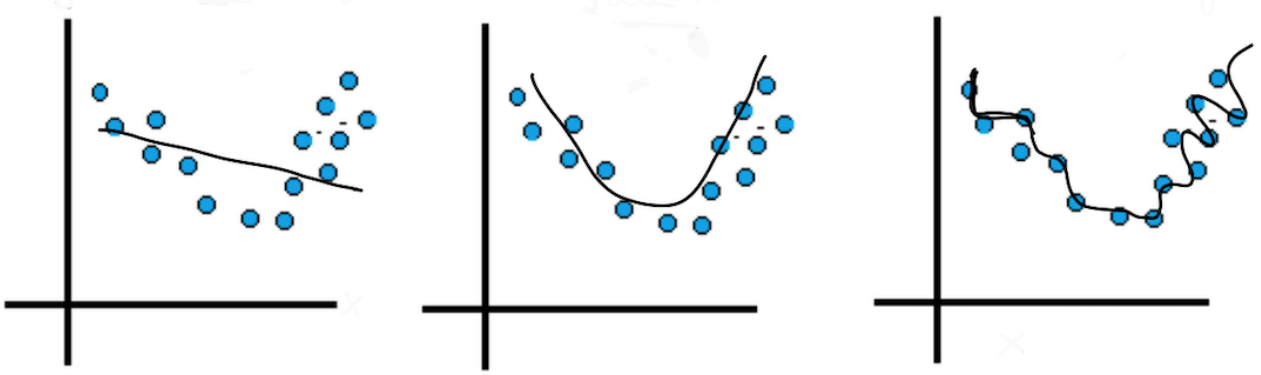
Question 4

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:

We can make sure that the model is robust and generalisable by using Regularisation.

Regularisation is a technique that is literally used to tune power of the model so that it doesn't overfit and instead learn the pattern underlying the dataset.



generalisation of a model is usually not achieved due to overfitting of the model. The model instead of learning the pattern it learns the data in it (as shown in 3rd diagram).

Regularisation helps in avoiding overfitting by penalising the model when the value of coefficient of the features increases.

This helps in keeping coefficient values small and controls variance of the model(regularized model doesn't zigzag widely as shown in 2nd diagram).

Precaution must be taken that the model is not regularised too much that it is unable learn the inherent pattern in the dataset (Example shown in 1st diagram).

regularisation helps the model in generalising as a cause-effect It leads decrease in accuracy of the model.

Since regularization reduces the variance of the model , it doesn't fit the output exactly . This leads to reduction in accuracy.

how exactly the model fits / how accurate the model is on the trained dataset depends on how less it is regularized.

A good model strike a correct balance between generalisability and accuracy by tuning the hyper parameter of the regularisation algorithm to optimal value.

In Linear Regression , Ridge and Lasso are common techniques used in Regularisation.