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 to double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Ans:

The optimal value for lambda is:

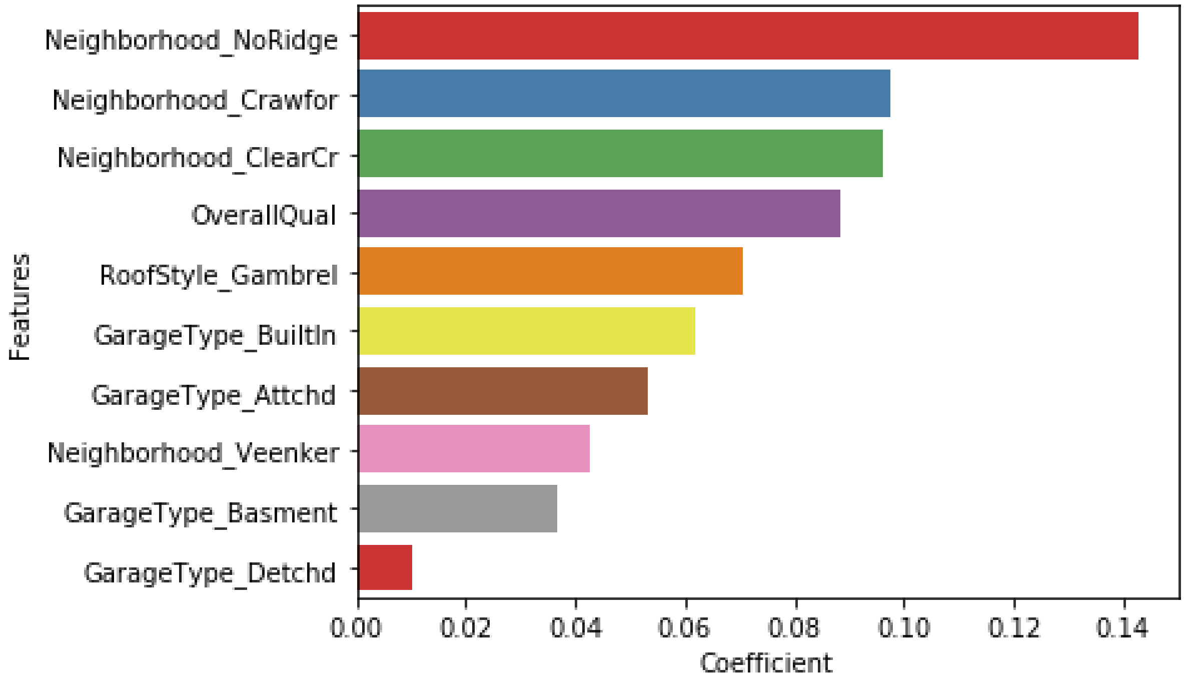
- Ridge: 1
- Lasso: 0.001

When we double the value for alpha

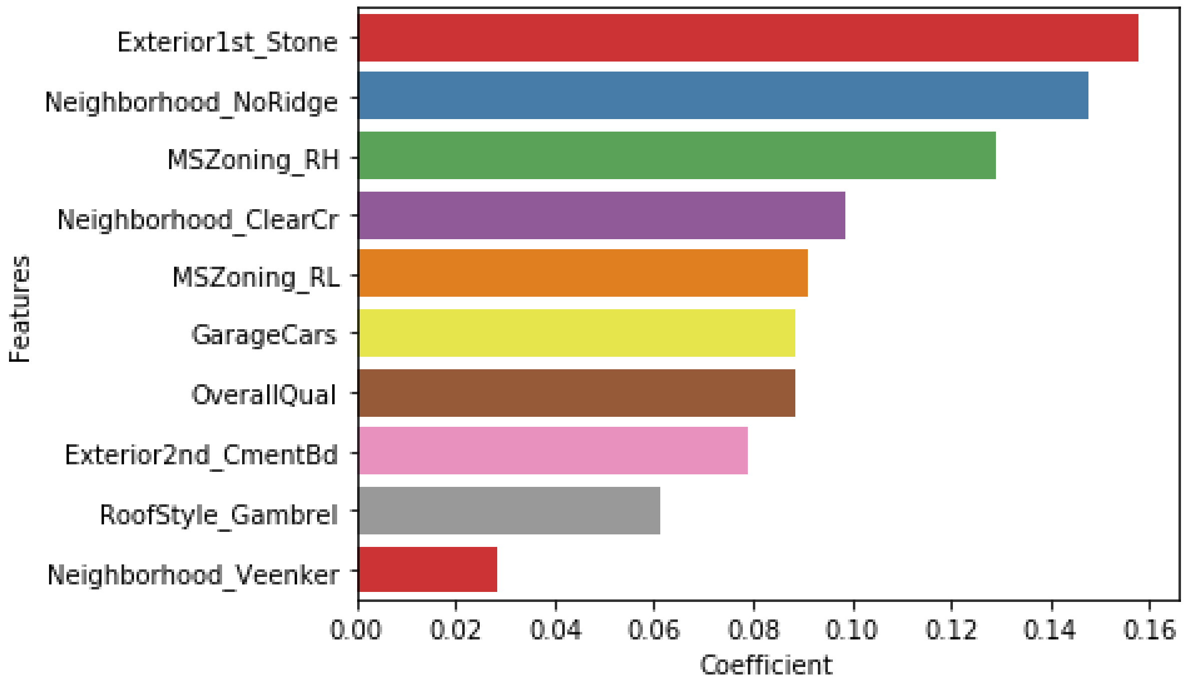
There is no significant change in coefficients for predictor variables, R squared score for train and test and Mean Squared error.

The most important predictor variables after the change is implemented are as follows:

Ridge:



Lasso:



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:

- The optimal value for lambda for Ridge and Lasso is as follow:
    - Ridge: 1
    - Lasso: 0.0001

- Mean Squared Error for Ridge and Lasso is as follow:
    - Ridge: 0.02473
    - Lasso: 0.02490

- R-squared value for Ridge and Lasso is as follow:
    - Ridge
        - Train – 0.8871
        - Test – 0.8269
    - Lasso
        - Train – 0.8871
        - Test – 0.8257

Based on the Mean Squared error, Ridge has lower Mean Squared Error than lasso, also R Squared for test is higher for Ridge than Lasso.

Based on above observation, we will choose and apply Ridge regression to the model.

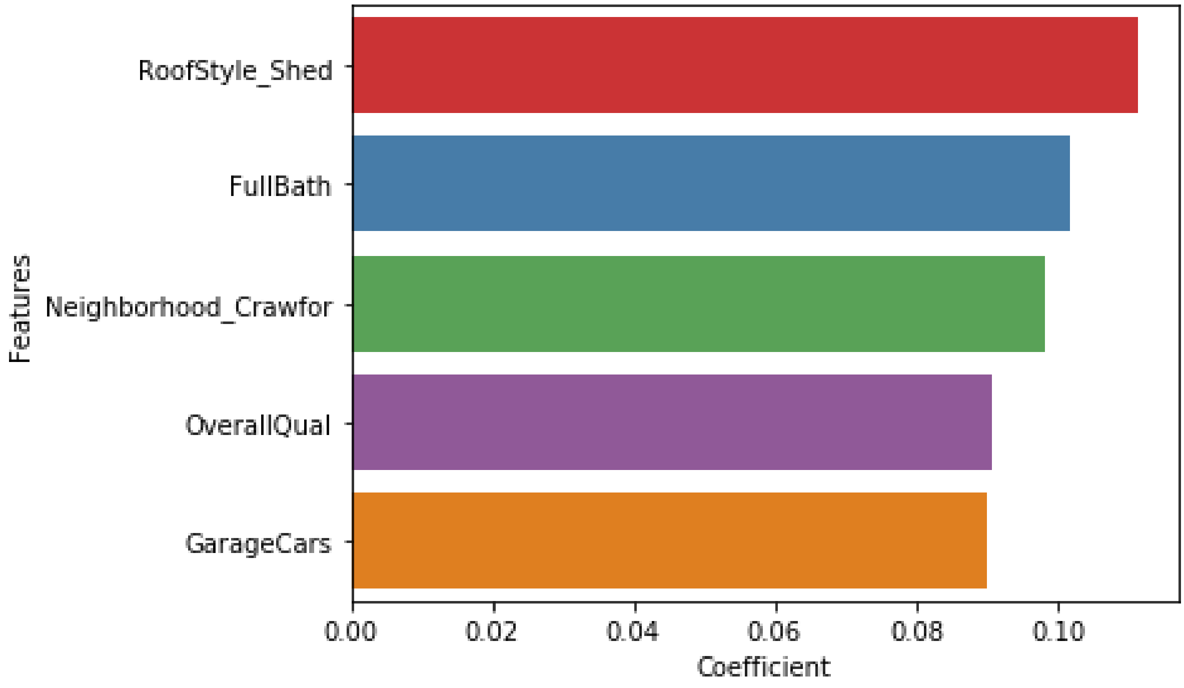
Question 3

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:

After dropping the five most important predictor variables in the lasso model, the new predictor variable given by revised model are as follows

Features	Coefficient
RoofStyle_Shed	0.111
FullBath	0.101
Neighborhood_Crawfor	0.983
OverallQual	0.090
GarageCars	0.090



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:

There are few ways to keep model simple and generalizable

Occam's Razor

It is the problem-solving principle that "entities should not be multiplied without necessity" or, more simply, the simplest explanation is usually the right one.

As per Occam's Razor, if we have two models that show similar performances in the finite training or test data, we should pick the simpler one, because:

- Simplified models are more generalised and can be applied on a wider scale
- They require lesser training samples for effective training in comparison to complex models, making them easier to train
- Simple models are more robust
  - Complex models are more probable to change with changes in the training data set
  - Simple models have low variance, high bias whereas complex models have low bias, high variance
- Simple models are likely to make more errors in the training set whereas complex models lead to over fitting – They work perfectly for training sets, but are catastrophic when applied to other generalised test samples

Thus, to make the model more robust and generalisable, the model should be made simple, but not too simple such that it will deem useless.

Regularization

Regularization is a process used to deliberately simply the model. One can adapt to regularization to make the model simpler. Regularization can help draw the line between keeping the model simple and making it too naïve. A regularization term needs to be added in regression equation, to the cost that adds up the absolute values (Lasso) or the squares of the parameters (Ridge) of the model.

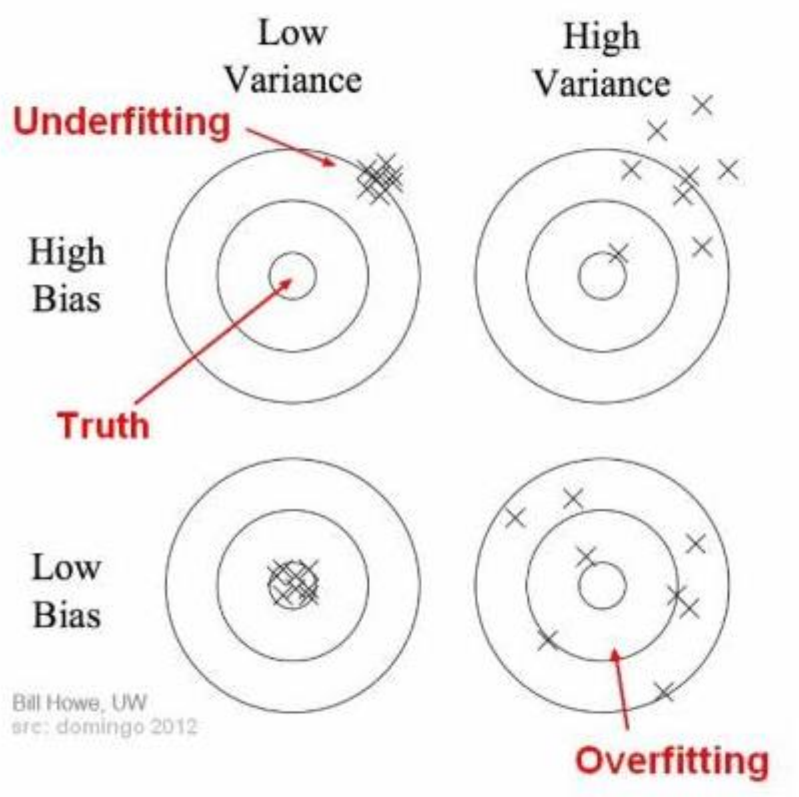
Bias Variance Trade Off

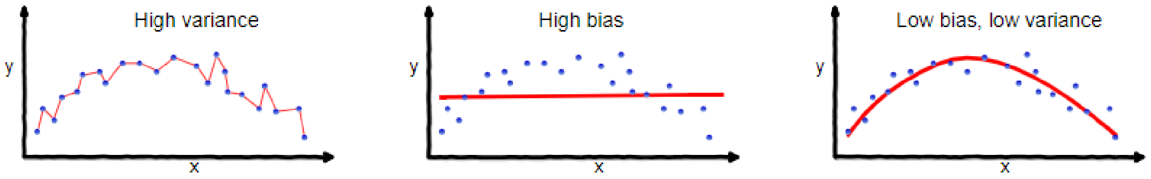
Making a model simple leads to bias Variance trade off:

- Complex models need to adapt to every minor change in the dataset and thus is highly sensitive to small changes in the training data
- A simpler model is unlikely to make major changes even if there are addition or subtraction of points.

Bias quantifies how accurate the model will be on the test data. Complex models can do an excellent job provided the training data has sufficient points. Models that are too naive, which give the same answer to all inputs, with no change in results whatsoever has a very large bias as the expected error across all inputs are very high.

Variance is the amount of change that the model goes through with respect to the changes on the training data.



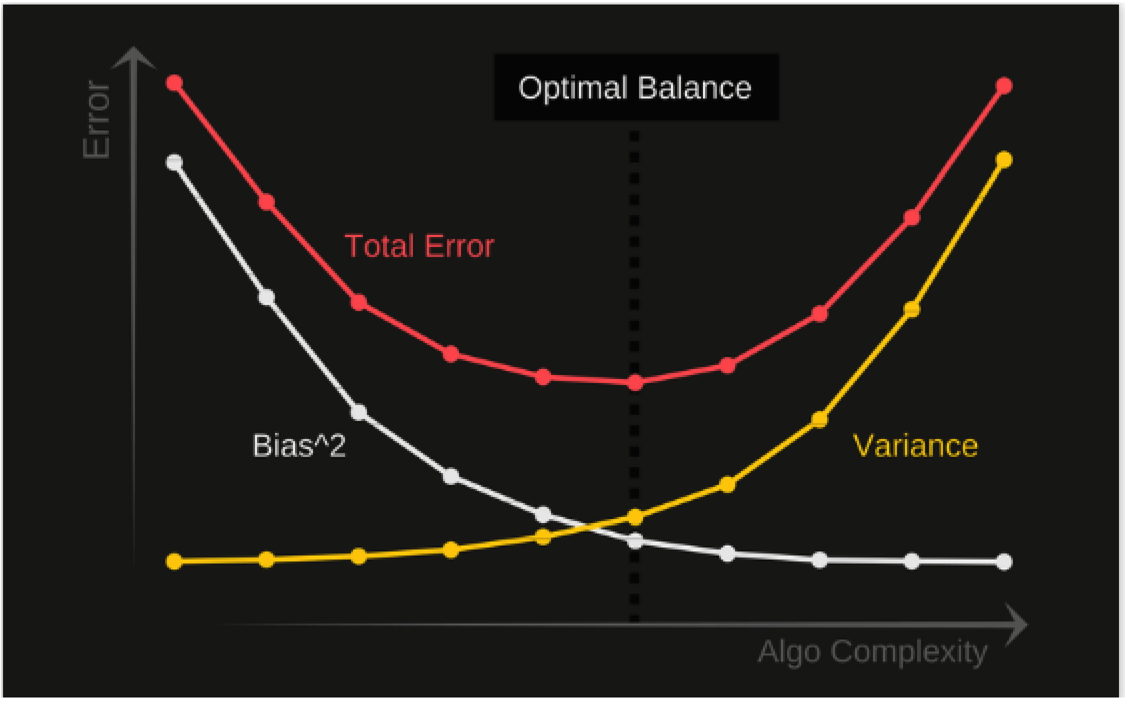


overfitting

underfitting

Good balance

Thus, to maintain the accuracy of the model, there must be a balance between the Bias and Variance as it reduces the total error.


Optimal Balance
Error
Total Error
Bias^2
Variance
Algo Complexity