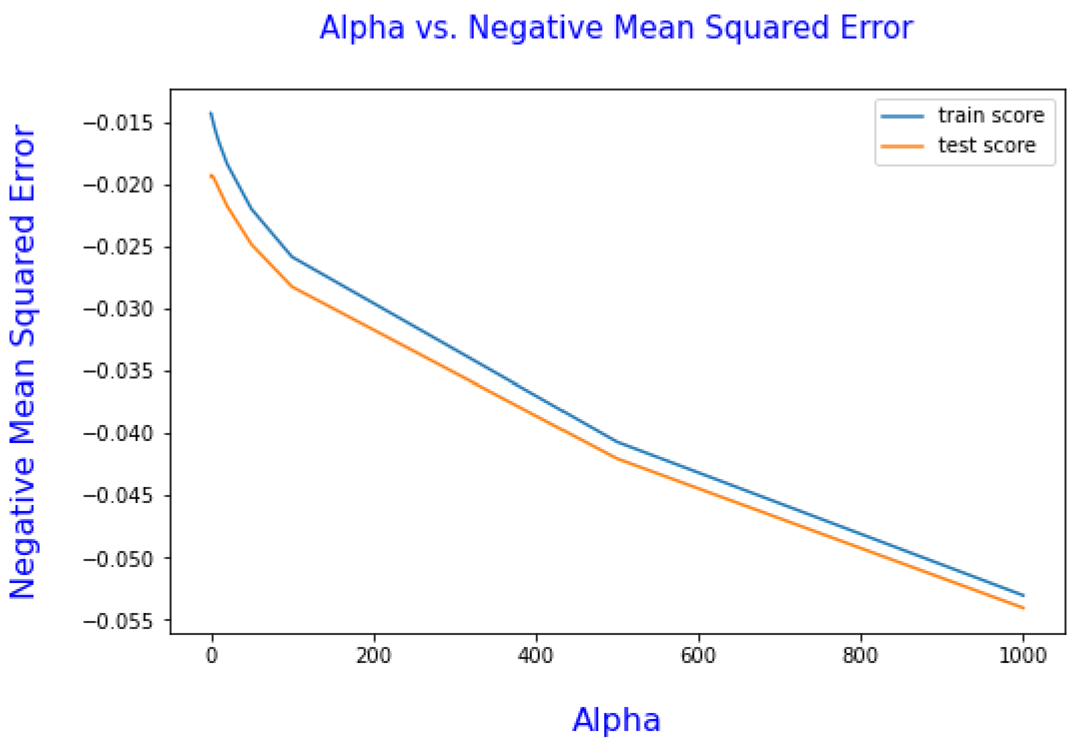
Question 1

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: The best value for Alpha for Ridge Regression is 0.7

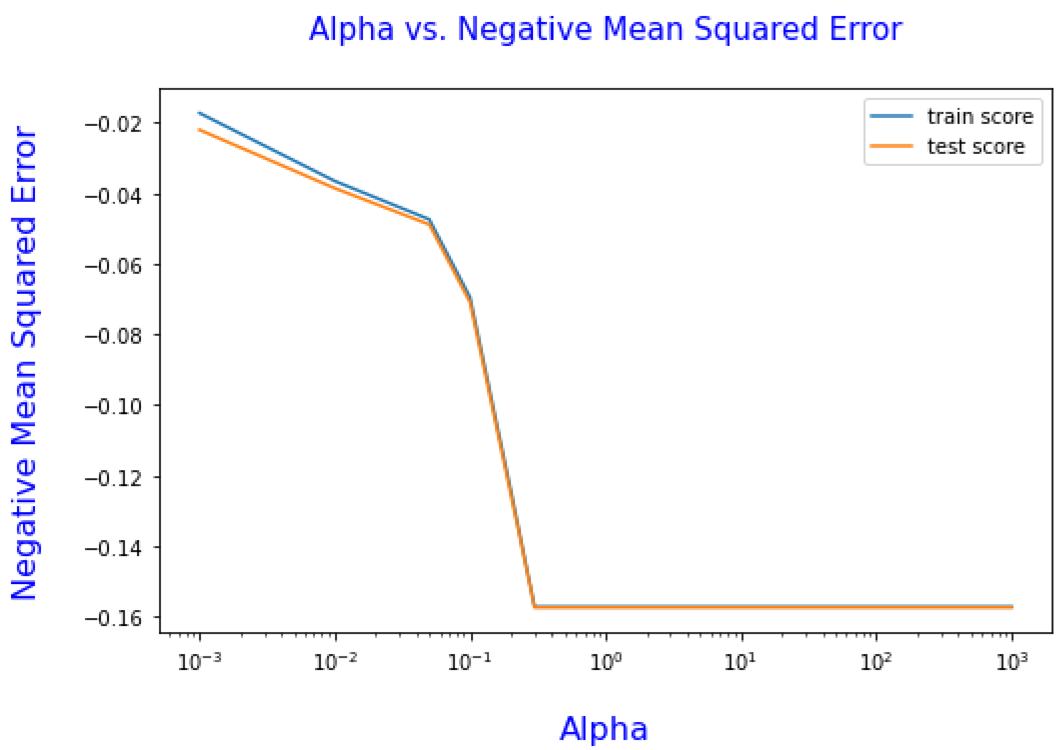


Model performance matrix for Ridge with alpha 0.7 on test and train data:

```
Train Dataset:
------------------------------------
Ridge Regression: Evaluation metrics
------------------------------------
MSE: 0.014837397564007357
RMSE: 0.12180885667309811
R2 Square: 0.9055811035739851
```

```
Test Dataset:
-----------------------------------
Ridge Regression: Evaluation metrics
-----------------------------------
MSE: 0.02093971313917017
RMSE: 0.14470560852700276
R2 Square: 0.8729500249966605
```

The best value for lasso regression is 0.001



Model performance matrix for lasso with alpha 0.001 on test and train data:

```
Train Dataset:
-----------------------------------
Lasso Regression: Evaluation metrics
-----------------------------------
MSE: 0.01786892917643991
RMSE: 0.13367471405033904
R2 Square: 0.8862897239306435
```

```
Test Dataset:
------------------------------------
Lasso Regression: Evaluation metrics
------------------------------------
MSE: 0.022334359208428613
RMSE: 0.14944684408989242
R2 Square: 0.8644881254921093
```

When we double the values of alpha for Ridge and Lasso we get the model parameters as following:

Ridge Regression:

```
Train Dataset:
------------------------------------
Ridge Regression: Evaluation metrics
------------------------------------
MSE: 0.014990790576871362
RMSE: 0.12243688405407646
R2 Square: 0.9046049755884944
```

```
Test Dataset:
------------------------------------
Ridge Regression: Evaluation metrics
------------------------------------
MSE: 0.020782593209280265
RMSE: 0.1441616911987379
R2 Square: 0.8739033371567826
```

Lasso:

```
Train Dataset:
------------------------------------
Lasso Regression: Evaluation metrics
------------------------------------
MSE: 0.020653043852085555
RMSE: 0.1437116691576768
R2 Square: 0.868572800591229
```

```
Test Dataset:
------------------------------------
Lasso Regression: Evaluation metrics
------------------------------------
MSE: 0.024817493823541793
RMSE: 0.15753569063403314
R2 Square: 0.8494219118967611
```

Top 5 with Lasso at Alpha = 0.001

	Features	Coefficient
38	OverallQual_Excellent	0.241190
48	OverallCond_Fair	-0.213345
61	BsmtQual_No Basement	-0.171611
43	OverallQual_Very Good	0.148668
20	Neighborhood_Crawfor	0.139267

Top 5 with Lasso at Alpha = 0.001

	Features	Coefficient
60	BsmtFinType1_No Basement	-0.202942
25	Neighborhood_NridgHt	0.153298
65	KitchenQual_Fa	-0.147356
0	1stFlrSF	0.135606
67	KitchenQual_TA	-0.127217

Top 5 with Ridge at Alpha = 0.7

	Features	Coefficient
41	OverallQual_Poor	-0.394913
38	OverallQual_Excellent	0.227101
48	OverallCond_Fair	-0.221398
49	OverallCond_Very Poor	-0.184210
44	OverallQual_Very Poor	-0.184210

Top 5 with Ridge at Alpha = 0.14

	Features	Coefficient
41	OverallQual_Poor	-0.309161
38	OverallQual_Excellent	0.221562
48	OverallCond_Fair	-0.215513
25	Neighborhood_NoRidge	0.153925
20	Neighborhood_Crawfor	0.153544

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 best value for alpha in case of Ridge nd Lasso are:

- Ridge = 0.7
- Lasso = 0.001

The MSE for Ridge and Lasso on Train and Test are:

- Ridge (MSE Train) = 0.014837397564007357
- Ridge (MSE Test) = 0.02093971313917017

- Lasso (MSE Train) = 0.01786892917643991
- Lasso (MSE Test) = 0.022334359208428613

The R2 Score for Ridge and Lasso on Train and Test are:

- Ridge (R2 Train) = 0.9055811035739851
- Ridge (R2 Test) = 0.8729500249966605

- Lasso (R2 Train) = 0.8862897239306435
- Lasso (R2 Test) = 0.8644881254921093

As we can see that there is no major difference in the outcomes of the performance in Ridge or Lasso.

We will go with Lasso here, because it helps in feature selection/reductions as the coefficient values can become 0. The performance of lasso is good hence we will select variable predicted by Lasso to predict the price of the house

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?

Top 5 with Lasso at Alpha = 0.001 before dropping 5 imp predictors:

	Features	Coefficient
38	OverallQual_Excellent	0.241190
48	OverallCond_Fair	-0.213345
61	BsmtQual_No Basement	-0.171611
43	OverallQual_Very Good	0.148668
20	Neighborhood_Crawfor	0.139267

Top 5 with Lasso at Alpha = 0.001 after dropping 5 imp predictors:

	Features	Coefficient
60	BsmtFinType1_No Basement	-0.202942
25	Neighborhood_NridgHt	0.153298
65	KitchenQual_Fa	-0.147356
0	1stFlrSF	0.135606
67	KitchenQual_TA	-0.127217

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:

Occam's Razor: This is a thumb rule in ML , Advantages:

- Simple model is more generic than a complex model, as they perform better on unseen data sets
- Fewer data training points.
- No significant change to minor changes in the training data points
- Can outperform in complex models while processing new data

Therefore, to make sure the mode is robust and generalizable making the model simple is effective. But here the level of simplicity needs to be taken care of as over simplification of the model can make the model unless.

Making the model leads to Bias – Variance Trade off:

- Complex model: Little change in data will change the model significantly
- Simple Model : Unlikely to have major change even if some points are added or removed

Bias: Difference between the expected and the true value of the parameter being estimated. High bias causes the algorithm to miss relevant relation between features

Variance: Error/ small fluctuations in the training set. High variance can cause noise in the training data.

To have the accuracy of the model, a balance in Bias and Variance is required to be kept.