**Subjective Questions**

**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:

**Ridge Regression:**

The optimal value of alpha for ridge regression is: **4**

The R2 score on test set with optimal value of alpha come out to be **0.897**

And the best predictor variables for the optimal alpha are –

|  |  |
| --- | --- |
| Features | Coefficients |
| Neighborhood\_Crawfor | 0.140 |
| Neighborhood\_StoneBr | 0.130 |
| Functional\_Typ | 0.128 |
| CentralAir\_Y | 0.101 |
| SaleCondition\_Normal | 0.098 |

If we double the value of alpha, there isn’t much change in R2 score.

The best predictors are -

|  |  |
| --- | --- |
| Feature | Coefficient |
| Neighborhood\_Crawfor | 0.121 |
| Functional\_Typ | 0.114 |
| Neighborhood\_StoneBr | 0.102 |
| SaleCondition\_Normal | 0.090 |
| CentralAir\_Y | 0.090 |

The coefficients value decreases for each feature as we doubled the alpha.

**Lasso Regression:**

The optimal value of alpha for lasso regression is **0.001**

The R2 Scores on test set with above optimal values of alpha is: **0.90**

And the best predictor variables for the optimal alpha are –

|  |  |
| --- | --- |
| Feature | Coefficient |
| Neighborhood\_Crawfor | 0.140 |
| Functional\_Typ | 0.124 |
| CentralAir\_Y | 0.107 |
| Neighborhood\_StoneBr | 0.095 |
| SaleType\_New | 0.082 |

On doubling the alpha,

R2 score on test set turns to be: 0.89

The best predictor variables –

|  |  |
| --- | --- |
| Feature | Coefficient |
| Neighborhood\_Crawfor | 0.113 |
| Functional\_Typ | 0.106 |
| SaleType\_New | 0.080 |
| CentralAir\_Y | 0.077 |
| SaleCondition\_Normal | 0.055 |

Again, there is significant decrease in the coefficient value on doubling the alpha.

More number of coefficients have value zero. Hence, feature elimination tendency increased on doubling the alpha.

**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:

Since, during model evaluation both Ridge & Lasso gives similar results on test set but lasso has an advantage that it penalizes more when there are high number of features.

This characteristic of lasso helps in feature elimination. Thus, we will choose lasso as our final model in this problem.

**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:

Originally below are the top 5 features

|  |  |
| --- | --- |
| Feature | Coefficient |
| Neighborhood\_Crawfor | 0.140 |
| Functional\_Typ | 0.124 |
| CentralAir\_Y | 0.107 |
| Neighborhood\_StoneBr | 0.095 |
| SaleType\_New | 0.082 |

After excluding the above 5 features, we get the model with best 5 features as -

|  |  |
| --- | --- |
| Feature | Coefficient |
| SaleType\_WD | 0.084 |
| MSZoning\_FV | 0.081 |
| SaleType\_Oth | 0.078 |
| HouseStyle\_SLvl | 0.078 |
| MSZoning\_RL | 0.063 |

**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:

A model is said be robust if it is stable, i.e. it does not change drastically when changing the training set. Whereas a model is said to be generalisable if it does not overfits the training data and performs well on the unseen data.

We can use regularization to make sure our model is robust and generalized.

Regularization ensures the model is as simple as possible and the model doesn’t learn unnecessary and redundant features.

We should take care of following parameters while training the model:

* R squared – Should be maximized (near to 1)
* Adj R square – Should be maximized (near to 1)
* AIC – Should be minimized
* BIC – Should be minimized

This will ensure our model is simple, robust and generalizable.

Robustness and Generalizability has a huge implication on model accuracy.

A robust and generalizable model will perform good on both training and testing set, i.e. the accuracy on training data and on testing data is close.