Advanced Regression Assignment

Subjective Questions

Question 1

What is the optimal value of alpha for ridge and lasso regression?

Optimal value of Ridge Regression is – 6.0

Optimal value of Lasso Regression is – 100

What will be the changes in the model if you choose double the value of alpha for both ridge and lasso?

After Doubling the alpha vaues –

What will be the most important predictor variables after the change is implemented?

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?

Below are the Optimal Values, R2, RMSE values for

Ridge Regression

Optimal Value of Lambda – 6

R2 Score (Train) – 0.88

R2 Score (Test) – 0.85

Lasso Regression – 100

R2 Score (Train) – 0.88

R2 Score (Test) – 0.84

	Metric	Linear Regression	Ridge Regression	Lasso Regression
0	R2 Score (Train)	9.069813e-01	8.804186e-01	8.893337e-01
1	R2 Score (Test)	-2.400766e+22	8.505200e-01	8.487007e-01
2	RSS (Train)	5.936317e+11	7.631514e+11	7.062565e+11
3	RSS (Test)	6.783905e+34	4.223894e+11	4.275303e+11
4	MSE (Train)	2.411269e+04	2.733962e+04	2.630076e+04
5	MSE (Test)	1.243104e+16	3.101875e+04	3.120694e+04

Since both the regressions, have the R2 score approximately the same, but since Lasso shrinks some of the variable coefficients to 0 and helps in variable selection, those features are automatically

excluded and hence the model becomes simple with less number of features. Hence Lasso Regression is preferred.

Question 3

After building the model, you realized that the 5 most important predictor variables in lasso model are not available in the incoming data. You will now have to create another model excluding the 5 most important predictor variables. Which are the 5 most important predictor variables now?

The five most important predictor variables as per Lasso Regression model are –

- TotalBsmtSF
- 2ndFlrSF
- OverallQual
- Neighborhood_NoRidge
- GarageCars

Now after removing the above 5 predictor variables below are the other 5 important predictor variables –

Question 4

How can you make sure that a model is robust and generalisable?

As per the "Occam's Razor" – the fundamental principle of all the machine learning algorithms are the predictive model has to be as simple as possible, but no simpler.

Some of the parameters to look for complexity of a model is

- Number of parameters required to specify the model completely.
- Degree of the function, if it is a polynomial.
- Size of the best-possible representation of the model.
- Depth or size of a decision tree.

One simple thumb rule as per Occam's Razor is given 2 models that show similar performance, in the finite training or test data, we should pick the one that makes fewer assumptions about the data that is yet to be seen. That means we need to pick the "simpler" of the 2 models. The relationship between the complexity of a model & its usefulness in a learning context is

- Simpler models are usually more "generic".
- Simpler models require fewer training samples.
- Simpler models are more robust.
- Simpler models make more errors in the training set – Complex models lead to overfitting. Reducing the Overfitting makes the model more robust & generalized.

What are the implications of the same for the accuracy of the model and why?