Advanced Regression Assignment

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: Optimal Value of lambda for Ridge: 10 and for Lasso: 0.001.

If we choose to double the value of alpha for both ridge and lasso then for ridge that will lower the coefficients and for Lasso there would be more less important features coefficients turning 0.

The most important predictor variable after the change is implemented are:

Lasso:

```
('GarageFinish_RFn', 0.059),
 ('GarageFinish_Unf', 0.066),
 ('SaleCondition_Normal', 0.07),
 ('SaleCondition_Others', 0.104),
 ('SaleCondition_Partial', 0.156)
```

Ridge:

```
('GarageFinish_RFn', 0.075),
 ('GarageFinish_Unf', 0.077),
 ('SaleCondition_Normal', 0.079),
 ('SaleCondition_Others', 0.089),
 ('SaleCondition_Partial', 0.108)
```

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: Optimal Value of lambda for ridge: 10 and for Lasso: 0.001

As the result we can go with Lasso regression as it results in model parameters lesser important features coefficients become zero.

Ridge: Train: 90.9 Test: 87.4

Lasso: Train: 89.8 Test: 86.5.

Question 3: After building the model, you realized 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 : I have executed the same notebook and removed the top 5 significant variables and I found below variables as next 5 significant.

-Ridge:

```
('GarageType_BuiltIn', 0.09),
('GarageType_Detchd', 0.094),
('GarageType_No Garage', 0.097),
('GarageType_Others', 0.103),
('GarageFinish_No Garage', 0.141)]
```

-Lasso

```
('GarageType_BuiltIn', 0.084),
('GarageType_Detchd', 0.086),
('GarageType_No Garage', 0.093),
('GarageType_Others', 0.118),
('GarageFinish_No Garage', 0.193)]
```

Question 4: How can you make sure that a model is robust and generalizable? What are the implications of the same for the accuracy of the model and why?

Answer: We can ensure model robustness and generalizable:

1. Split data for train-test.
2. Cross-validation.
3. Quality data, handle missing values.
4. Feature engineering.
5. Optimize hyperparameters.
6. Apply regularization.
7. Adversarial testing.
8. Evaluate on real-world data.

Implications for accuracy:

1. Balance training accuracy and generalization.
2. Manage bias-variance tradeoff.
3. Control model complexity with regularization.
4. Test for robustness to distribution shift.
5. Understand and acknowledge model limitations.