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?

10 and 100 are the values

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?

The lasso regression model had a slightly lower MSE, which indicates better performance on the test data. Additionally, lasso's feature selection can lead to a simpler model, which is often desirable.

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?

Sale Condition, Garage Type, House Style, Roof style are new variables

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?

Strategies for Robust and Generalizable Models

1. Cross-Validation:
2. Regularization:
3. Hyperparameter Tuning:
4. Feature Selection:
5. Data Augmentation:
6. Validation on Separate Test Set:
7. Model Ensemble:

Implications for Model Accuracy

1. Bias-Variance Trade-off:

2. Overfitting vs. Underfitting:
3. Model Complexity:

Summary

By employing these strategies, you can build models that are robust and generalizable. The implications for model accuracy are significant: a well-generalized model is likely to perform consistently well on new, unseen data, ensuring that its predictions are reliable and useful in real-world applications. This balance between bias and variance, achieved through regularization, cross-validation, and other techniques, is key to developing high-performing, dependable models.