**Q1 . 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_1:

Optimal value of alpha in Ridge's regression: 2.0

Optimal value of alpha in Lasso's regression: 0.0003

Double the value of Ridge and Lasso's Alpha values.

Ridge's top features after doubled:

1. 1stFlrSF
2. 2ndFlrSF
3. TotalBsmtSF
4. GarageArea
5. Neighborhood_Crawford

Lasso's top features after doubled:

1. 1stFlrSF
2. 2ndFlrSF
3. TotalBsmtSF
4. GarageArea
5. Neighborhood_Somerst

**Q2: 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_2 :

I will select Lasso's regression:

Reason:

1. Ridge's Regression considers all features of the model which is not the feasible way every time.
2. Ridge's Regression has lesser R-square value compared to Lasso's model(0.894 > 0.887)
3. Ridge's Regression model has relatively higher mean squared error(0.00596 > 0.00545)

**Q3: 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_3:

After dropping 5 top features from Lasso's model . We run it again with the same alpha value .
Hereby we get these features as the top one's:

- TotRmsAbvGrd
- GarageCars
- Fireplaces
- BsmtFinSF1
- Foundation_PConc

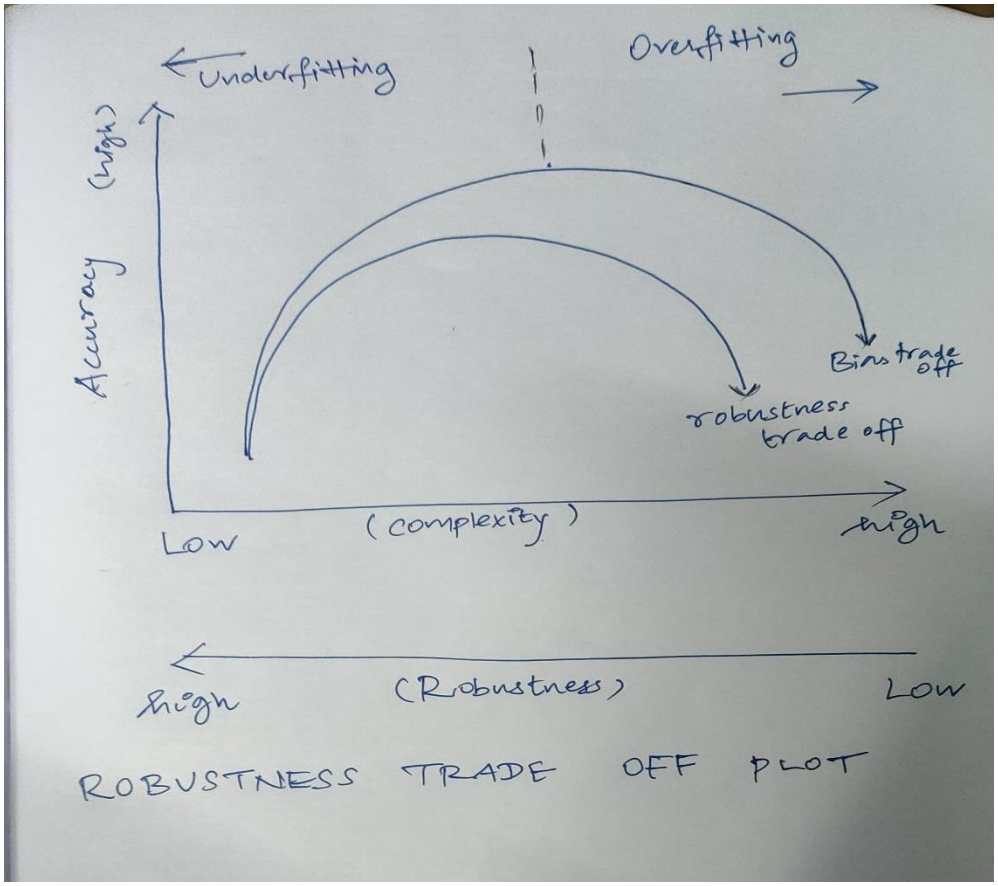
**Q4: How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?**

More Robust and Generalised Model :

1. The model performs well on unseen data
2. This can be achieved by training the model on outliers. So that it could perform data for which it hasn't seen in training or the test field
3. The model which overfits the data will be less generalised , hence by regularization methods we will train the model not to overfit any train data. Which in turn makes the model more generalisable
4. A robust model has high Generalisability

Accuracy of the model:

1. Highly accurate or robust model might overfit the data . More simple model with underfit the data . So there's always a trade-Off between robustness/accuracy and complexity.




2. We have a chance to select an optimal point where the trade off is very less . We use optimal alpha to get an optimal model.