Advanced Linear Regression Subjective Questions

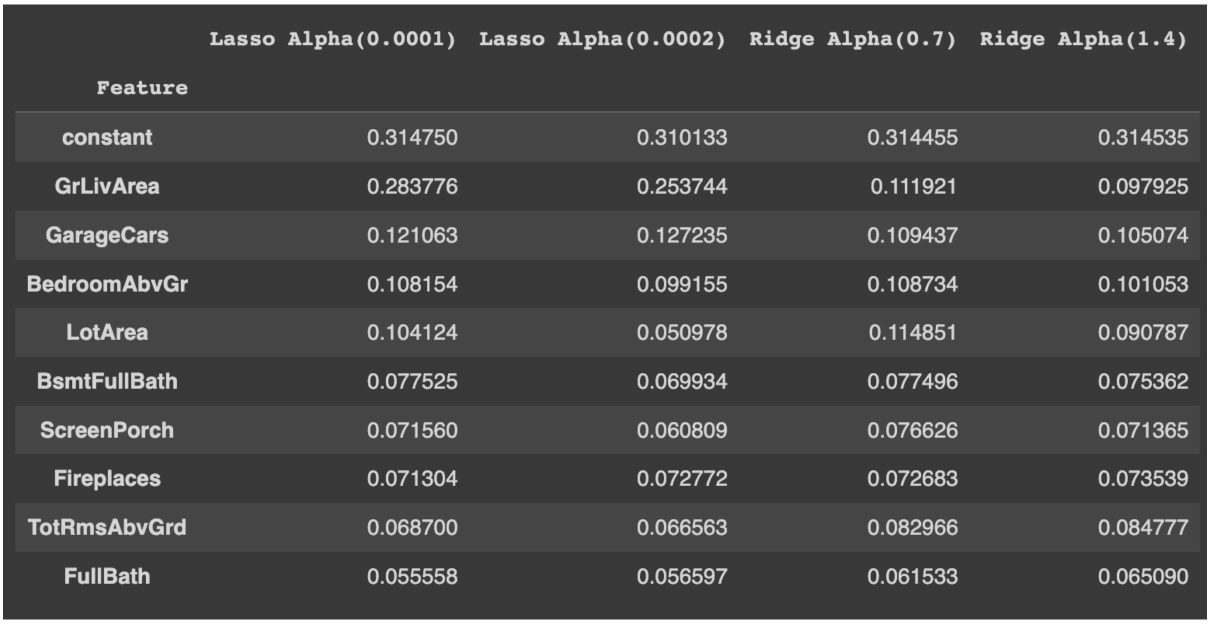
1. What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

A: Optimal Value of Alpha for :

Ridge : 0.7

Lasso : 0.0001

On doubling the Alpha for ridge and Lasso, the model coefficients are penalized more heavily. In case of ridge coefficient will move closer to zero and in case of Lasso coefficient will be set to zero.



Feature	Lasso Alpha(0.0001)	Lasso Alpha(0.0002)	Ridge Alpha(0.7)	Ridge Alpha(1.4)
constant	0.314750	0.310133	0.314455	0.314535
GrLivArea	0.283776	0.253744	0.111921	0.097925
GarageCars	0.121063	0.127235	0.109437	0.105074
BedroomAbvGr	0.108154	0.099155	0.108734	0.101053
LotArea	0.104124	0.050978	0.114851	0.090787
BsmtFullBath	0.077525	0.069934	0.077496	0.075362
ScreenPorch	0.071560	0.060809	0.076626	0.071365
Fireplaces	0.071304	0.072772	0.072683	0.073539
TotRmsAbvGrd	0.068700	0.066563	0.082966	0.084777
FullBath	0.055558	0.056597	0.061533	0.065090

Optimal Value for Lasso and Ridge

Lasso Regression with 0.0001
=====================================
R2 score (train) : 0.8401700603462888
R2 score (test) : 0.8457108625506681
RMSE (train) : 0.06184073772425473
RMSE (test) : 0.06347499664001059

Ridge Regression with 0.7
=====================================
R2 score (train) : 0.841350367643952
R2 score (test) : 0.8420300317600945
RMSE (train) : 0.06161197479775505
RMSE (test) : 0.06422768613393358

On doubling alpha for lasso and ridge

```
Lasso  Regression with  0.0002
=====================================
R2 score (train) :  0.8360877559124604
R2 score (test) :  0.8481937209212929
RMSE (train) :  0.06262551238422062
RMSE (test) :  0.06296219770020679

Ridge  Regression with  1.4
=====================================
R2 score (train) :  0.8401941252091704
R2 score (test) :  0.8445529281043156
RMSE (train) :  0.06183608201050448
RMSE (test) :  0.06371274019691095
```

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?**

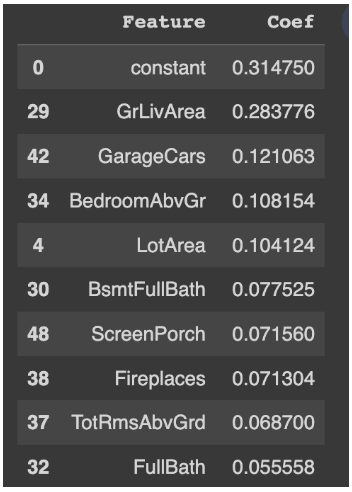
A: Idea is reduce to the number of features or data is showing multicollinearity then lasso regression should be preferred. As it suggests simple model by eliminating coefficients. Usually Ridge is considered when there is a high correlation between the model parameters. If collinearity is high it can create some bias also.
for our housing dataset , I would consider Lasso Regression, as R2 is slightly higher than ridge plus the model is similar because of the feature eliminations.

```
Lasso -
R2 score (train) :  0.8401700603462888
R2 score (test) :  0.8457108625506681
RMSE (train) :  0.06184073772425473
RMSE (test) :  0.06347499664001059

Ridge –
R2 score (train) :  0.841350367643952
R2 score (test) :  0.8420300317600945
RMSE (train) :  0.06161197479775505
RMSE (test) :  0.06422768613393358
```

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?**

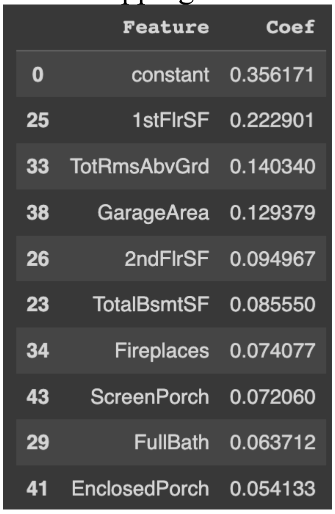
A:



	Feature	Coef
0	constant	0.314750
29	GrLivArea	0.283776
42	GarageCars	0.121063
34	BedroomAbvGr	0.108154
4	LotArea	0.104124
30	BsmtFullBath	0.077525
48	ScreenPorch	0.071560
38	Fireplaces	0.071304
37	TotRmsAbvGrd	0.068700
32	FullBath	0.055558

These were the top 5 feature from the lasso regression.

After dropping the these top5 feature we got :



	Feature	Coef
0	constant	0.356171
25	1stFlrSF	0.222901
33	TotRmsAbvGrd	0.140340
38	GarageArea	0.129379
26	2ndFlrSF	0.094967
23	TotalBsmtSF	0.085550
34	Fireplaces	0.074077
43	ScreenPorch	0.072060
29	FullBath	0.063712
41	EnclosedPorch	0.054133

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?**

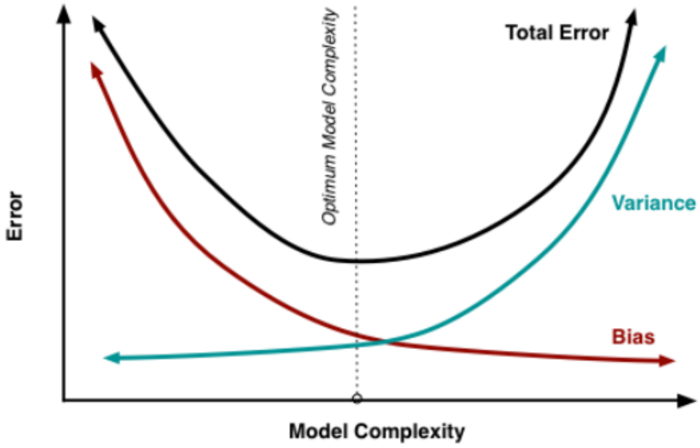
A: Model is considered to be robust if its output and predictions are consistently accurate even if one or more of the predictors are changed.
A generalizable model is able to adapt to new or unseen data taken from the same distribution , from which model is created.

To make the model more robust and generalizable, make the model simple but not simpler which will not be of any use.
there are a few typical ways of looking the complexity of a model :

1. Number of parameters required to specify the model completely
2. degree of the function. Higher means more complex
3. Size of the best-possible representation of the model. More complex the coefficient of model more complex the model.

Usually Simpler model are more generic and easier to train.
Simpler model makes more error while training.
Simpler models are more robust .

simple models have low variance, high bias and complex models have low bias, high variance. (e bias-variance trade off)



Using Regularization Techniques , we can make model simpler but not naive. As the Regularization technique add more cost as the number of model parameter increases.

Implication of the same for the accuracy of the model

1. Making model robust and generalizable , so they are not impacted by outliers in training set.
2. Making generalizable makes sure test accuracy doesn’t drop.
3. Outliers Analysis needs to be done only those required should be retained. Outliers that doesn’t make sense should be removed.