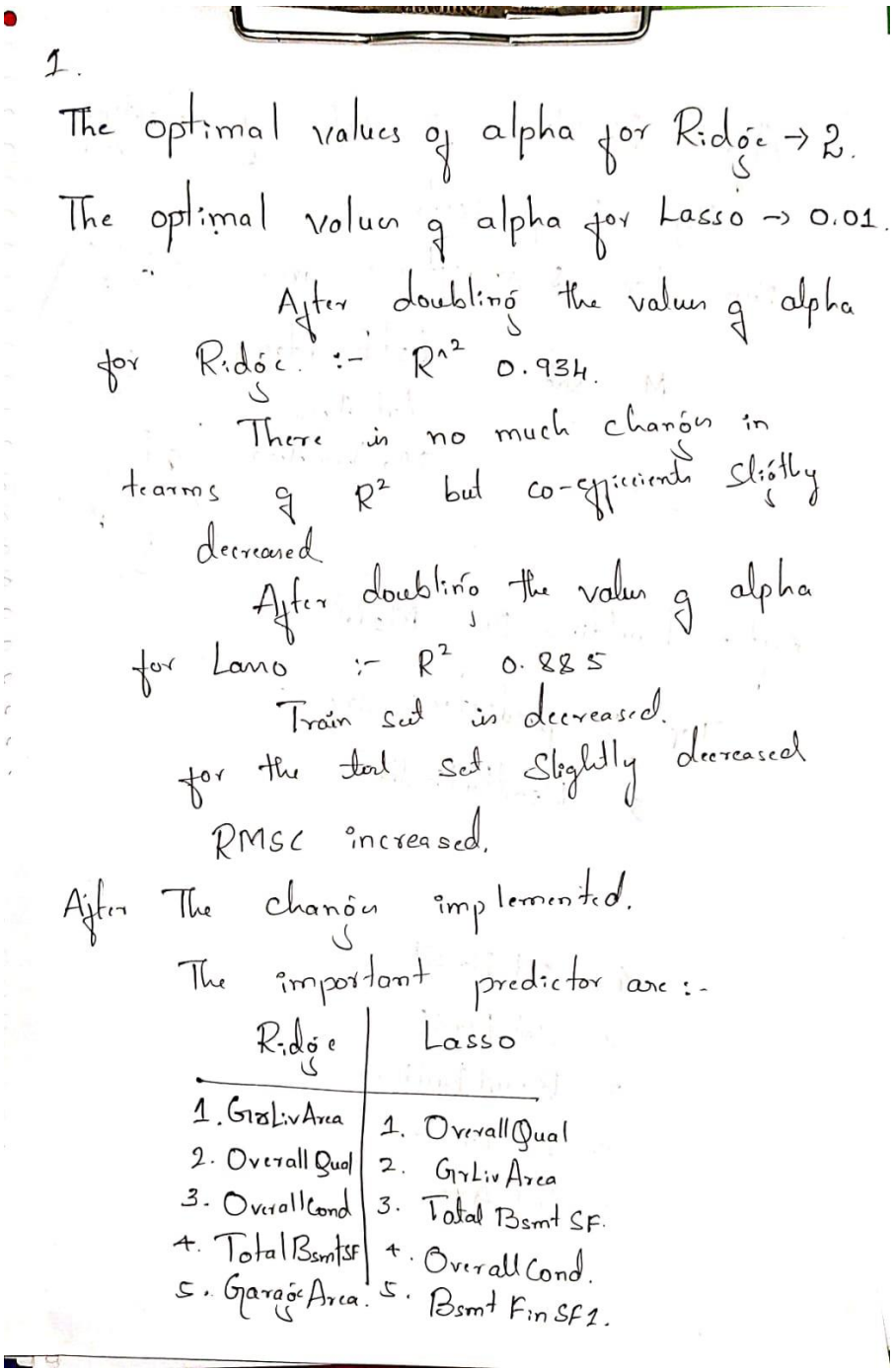
Assignment: Subjective Questions

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?

--



1.

The optimal values of alpha for Ridge → 2.

The optimal value of alpha for Lasso → 0.01.

After doubling the value of alpha for Ridge :- R^2 0.934.

There is no much change in terms of R^2 but co-efficients slightly decreased

After doubling the value of alpha for Lasso :- R^2 0.885

Train set is decreased.

for the test set. Slightly decreased RMSE increased.

After The change implemented.

The important predictor are :-

Ridge	Lasso
1. GrLivArea	1. OverallQual
2. OverallQual	2. GrLivArea
3. OverallCond	3. TotalBsmtSF
4. TotalBsmtSF	4. OverallCond.
5. GarageArea.	5. BsmtFinSF1.

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?

According to this data set the optimal values of lambda for lasso and ridge
Better role played by lasso regression so choosing the lasso Regression is the good and optimal values of lambda

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?

- ***1stflrSF***
- ***2ndFlrSF***
- ***BsmtFullBath***
- ***LotFrontsage***
- ***LotArea***

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?

when the model is said to be robust in any new changes in the set also its performance she not be affecting
model should be fitting in unseen data also so it's a generalisable
while making in the predications of the model it should be accuracy
for the good outcome of the business and and making business decision
implicating by visualization ,outliers treatment and data cleaning and applying good model fitting performance