REPORT LAB 2

Linear and Logistic Regression

Submitted By: Abhinav Jindal

Entry Number: 2016csb1026

21 September 2018

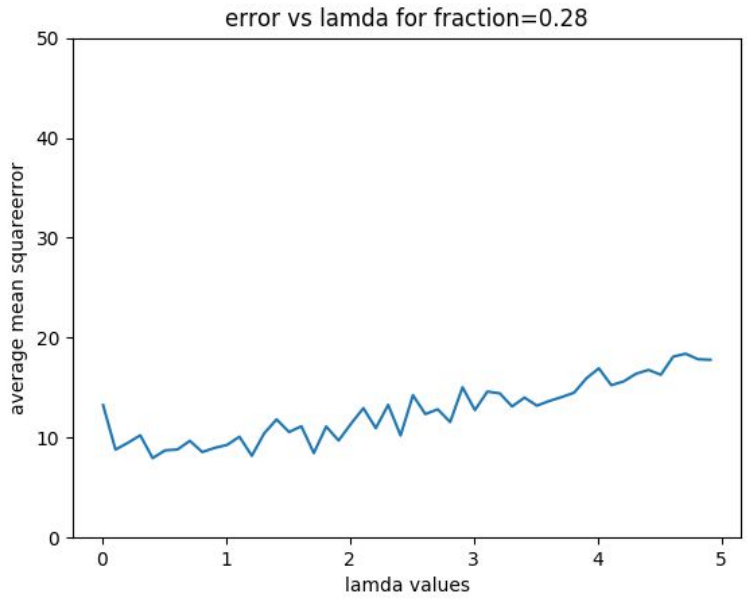
PART 1

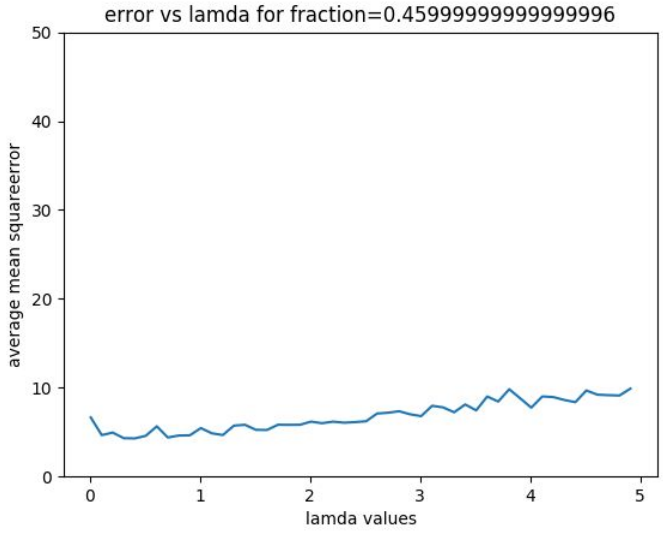
A partition is created after standardizing the data for the test and training data where test data is 20% of the total data and training data is 80% of the total data.

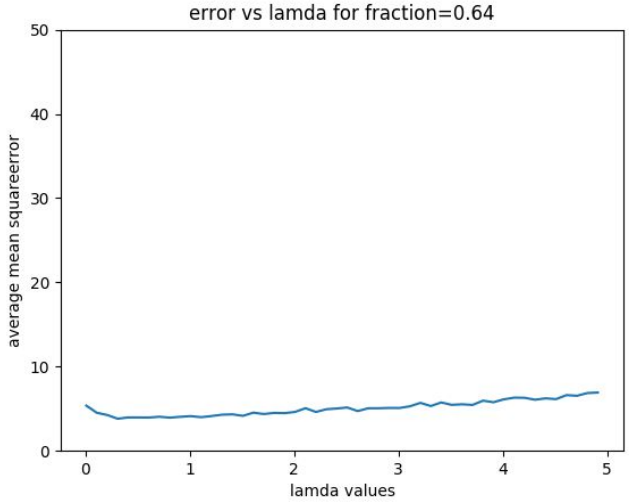
Effect of λ on error change:

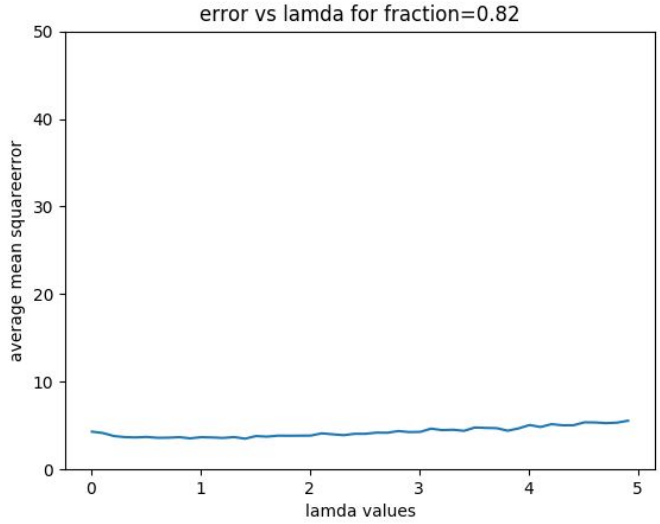
Many instances of training data and validation data are created for each different value of fractions from the training data created in previous step are used to calculate the mean squared error for different values of lambda.

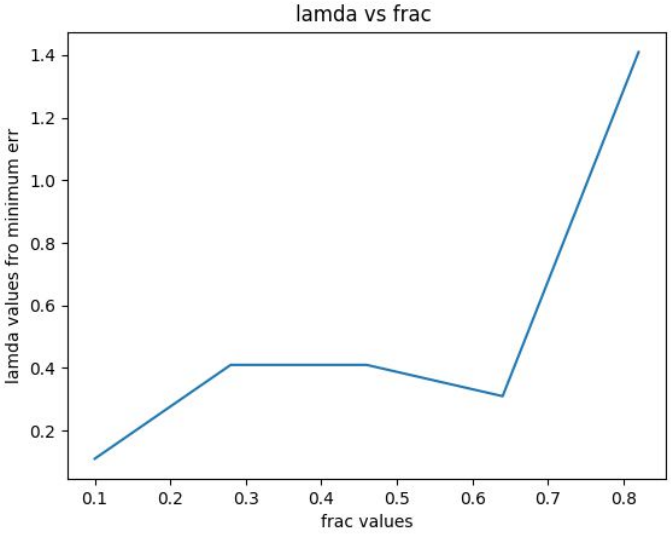

error vs lamda for fraction=0.1
average mean squareerror
lamda values
0
10
20
30
40
50
0
1
2
3
4
5

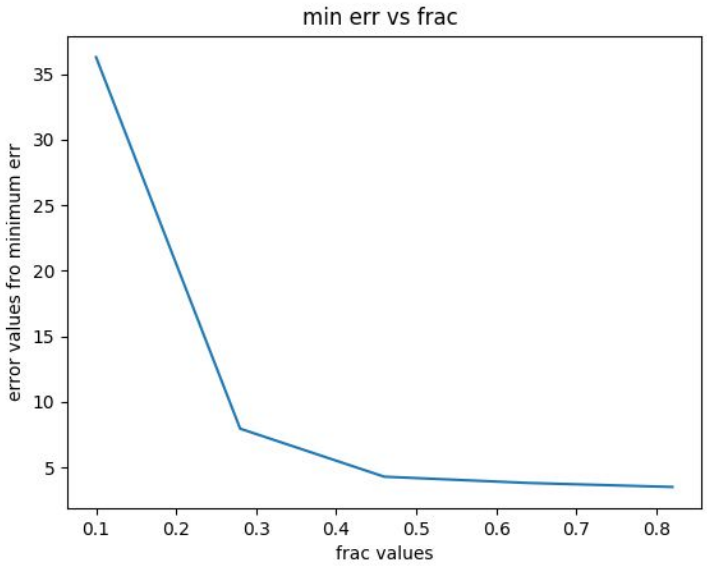

error vs lamda for fraction=0.28
average mean squareerror
lamda values
0
10
20
30
40
50
0
1
2
3
4
5


error vs lamda for fraction=0.45999999999999996
average mean squareerror
lamda values
0
10
20
30
40
50
0
1
2
3
4
5


error vs lamda for fraction=0.64
average mean squareerror
lamda values
0
10
20
30
40
50
0
1
2
3
4
5


error vs lamda for fraction=0.82
average mean squareerror
lamda values
0
10
20
30
40
50
0
1
2
3
4
5


lamda vs frac
lamda values fro minimum err
frac values
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8

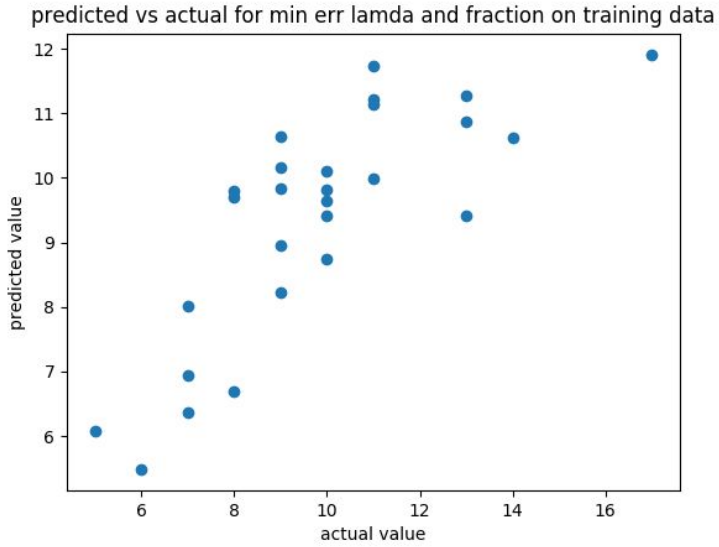


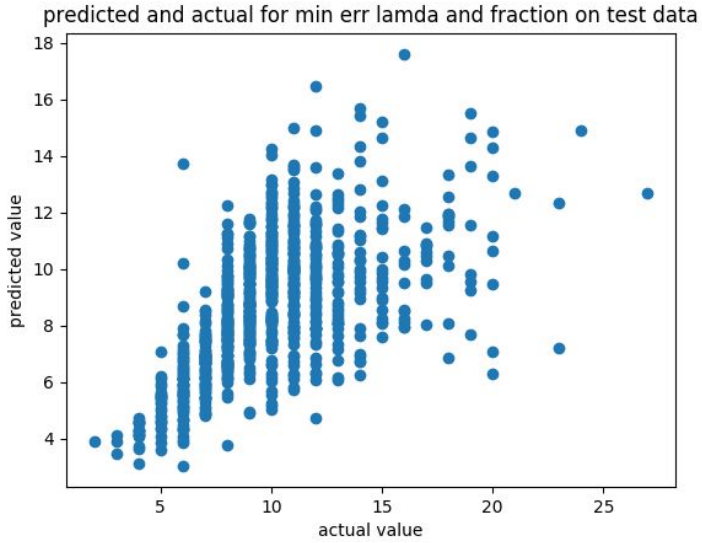
Observations: The value of lambda does not follow a fixed pattern. This indicates that the value of lambda is chosen according to the dataset and cannot be fixed for a model. However is is best to keep its value between 0.1 and 0.9. As we can see that increasing the fraction increases the accuracy and hence decreasing the average mean squared error. This is because we are able to train our model using more no of instances.

How to predict the accuracy of the model:

The model with the least average mean squared error over all the instances on test data is considered to be more accurate according to our model.

In our example the model is most accurate for **λ=1.41** and fraction=0.82 . The graph for these values for the predicted and actual value of rings is shown below:

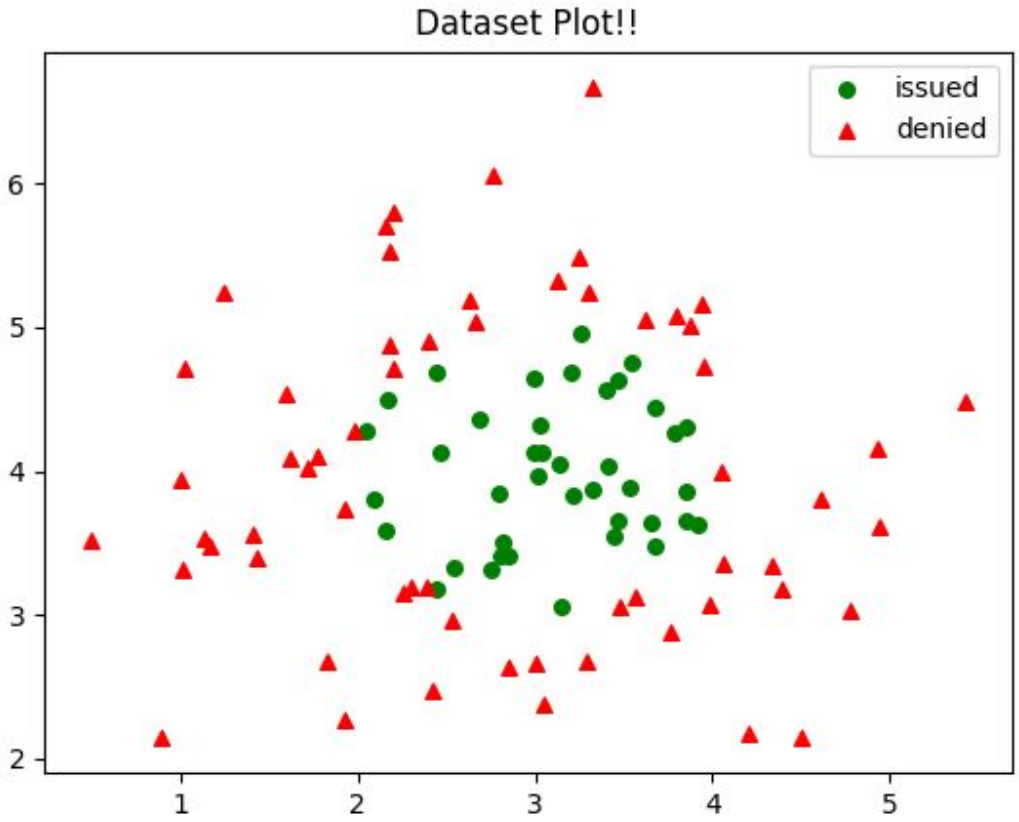




Observations: As we can see that the points lie somewhat around the line y=x this shows that our model is pretty good. For some instances the error is a little greater which tends to increase the average mean squared error.

PART 2

Plot for dataset for 2 classes :



Performance of gradient descent vs newton raphson:

(the number of iterations are fixed to not exceed 25000)

Accuracy of gradient descent = 0.55

Accuracy of newton raphson = 0.54

Observations: We see that the accuracy is close to half for both. Therefore, both the models predict with almost the same accuracy. But this is almost half which is same as just guessing the solution as there are only 2 classes. Therefore this shows that the data is not linearly separable.

Higher Degree Logistic Regression:

Degree	**Accuracy**
2	60
3	78
4	75
5	70
6	60

Observations: Accuracy increases for higher degree as compared to one degree logistic regression. However increasing the degree too much can lead to poor results as it results in overfitting.