SMAI Assignment 2

Chaitanya Patel : 201501071

Question 4

1. Lasso

Multiple runs with random shuffle and the results are as follows.

This is classification problem being solved with regression. Labels are 0 and 1. Linear model is trained. While testing, the model will predict the number between 0 and 1. I am putting threshold as 0.5 i.e. greater than 0.5 will be taken as 1 and less than 0.5 will be taken as 0.

alpha accuracy

1e-05 -> 0.75

0.0001 -> 0.75

0.001 -> 0.760869565217

0.01 -> 0.717391304348

0.1 -> 0.576086956522

1 -> 0.54347826087

10 -> 0.54347826087

100 -> 0.54347826087

1000 -> 0.54347826087

1e-05 -> 0.771739130435

0.0001 -> 0.782608695652

0.001 -> 0.782608695652

0.01 -> 0.75

0.1 -> 0.695652173913

1 -> 0.663043478261

10 -> 0.663043478261

100 -> 0.663043478261

1000 -> 0.663043478261

1e-05 -> 0.782608695652

0.0001 -> 0.79347826087

0.001 -> 0.79347826087

0.01 -> 0.70652173913

0.1 -> 0.684782608696

1 -> 0.619565217391

10 -> 0.619565217391

100 -> 0.619565217391

1000 -> 0.619565217391

1e-06 -> 0.75

1e-05 -> 0.739130434783

0.0001 -> 0.739130434783

0.001 -> 0.739130434783

0.01 -> 0.70652173913

0.1 -> 0.695652173913

1 -> 0.684782608696

10 -> 0.684782608696

100 -> 0.684782608696

1000 -> 0.684782608696

Thus overall on multiple runs, alpha = 0.0001 looks promising.

Hence alpha = 0.0001

2. Ridge

alpha accuracy

1e-06 -> 0.695652173913

1e-05 -> 0.695652173913

0.0001 -> 0.70652173913

0.001 -> 0.70652173913

0.01 -> 0.70652173913

0.1 -> 0.70652173913

1 -> 0.717391304348

10 -> 0.728260869565

100 -> 0.684782608696

1000 -> 0.630434782609

1e-06 -> 0.739130434783

1e-05 -> 0.739130434783

0.0001 -> 0.75

0.001 -> 0.75

0.01 -> 0.739130434783

0.1 -> 0.739130434783

1 -> 0.739130434783

10 -> 0.728260869565

100 -> 0.728260869565

1000 -> 0.663043478261

1e-06 -> 0.739130434783

1e-05 -> 0.739130434783

0.0001 -> 0.728260869565

0.001 -> 0.75

0.01 -> 0.771739130435

0.1 -> 0.782608695652

1 -> 0.771739130435

10 -> 0.70652173913

100 -> 0.673913043478

1000 -> 0.652173913043

1e-06 -> 0.739130434783

1e-05 -> 0.739130434783

0.0001 -> 0.75

0.001 -> 0.728260869565

0.01 -> 0.728260869565

0.1 -> 0.728260869565

1 -> 0.75

10 -> 0.684782608696

100 -> 0.641304347826

1000 -> 0.663043478261

1e-06 -> 0.771739130435

1e-05 -> 0.771739130435

0.0001 -> 0.760869565217

0.001 -> 0.75

0.01 -> 0.75

0.1 -> 0.739130434783

1 -> 0.739130434783

10 -> 0.739130434783

100 -> 0.728260869565

1000 -> 0.619565217391

1e-06 -> 0.782608695652

1e-05 -> 0.79347826087

0.0001 -> 0.79347826087

0.001 -> 0.771739130435

0.01 -> 0.760869565217

0.1 -> 0.760869565217

1 -> 0.760869565217

10 -> 0.75

100 -> 0.760869565217

1000 -> 0.586956521739

Here, alpha = 0.0001 looks promising.

Hence, alpha = 0.0001

3. Elastic Net

Over multiple random splits of train and validate, I observed that alpha values from 0.01 to 0.00001 works better. So I tested on that range with different l1_ratio’s.

alpha   l1_ratio   accuracy

1e-06 , 1e-06 -> 0.684782608696

1e-06 , 1e-05 -> 0.684782608696

1e-06 , 0.0001 -> 0.684782608696

1e-06 , 0.001 -> 0.684782608696

1e-06 , 0.01 -> 0.684782608696

1e-06 , 0.1 -> 0.684782608696

1e-06 , 0.5 -> 0.684782608696

1e-06 , 0.9 -> 0.717391304348

1e-05 , 1e-06 -> 0.684782608696

1e-05 , 1e-05 -> 0.684782608696

1e-05 , 0.0001 -> 0.684782608696

1e-05 , 0.001 -> 0.684782608696

1e-05 , 0.01 -> 0.684782608696

1e-05 , 0.1 -> 0.684782608696

1e-05 , 0.5 -> 0.684782608696

1e-05 , 0.9 -> 0.684782608696

0.0001 , 1e-06 -> 0.684782608696

0.0001 , 1e-05 -> 0.684782608696

0.0001 , 0.0001 -> 0.684782608696

0.0001 , 0.001 -> 0.684782608696

0.0001 , 0.01 -> 0.684782608696

0.0001 , 0.1 -> 0.684782608696

0.0001 , 0.5 -> 0.684782608696

0.0001 , 0.9 -> 0.684782608696

0.001 , 1e-06 -> 0.684782608696

0.001 , 1e-05 -> 0.684782608696

0.001 , 0.0001 -> 0.684782608696

0.001 , 0.001 -> 0.684782608696

0.001 , 0.01 -> 0.684782608696

0.001 , 0.1 -> 0.684782608696

0.001 , 0.5 -> 0.695652173913

0.001 , 0.9 -> 0.695652173913

0.01 , 1e-06 -> 0.684782608696

0.01 , 1e-05 -> 0.684782608696

0.01 , 0.0001 -> 0.684782608696

0.01 , 0.001 -> 0.684782608696

0.01 , 0.01 -> 0.684782608696

0.01 , 0.1 -> 0.684782608696

0.01 , 0.5 -> 0.695652173913

0.01 , 0.9 -> 0.695652173913

1e-07 , 1e-06 -> 0.695652173913

1e-07 , 1e-05 -> 0.695652173913

1e-07 , 0.0001 -> 0.695652173913

1e-07 , 0.001 -> 0.695652173913

1e-07 , 0.01 -> 0.695652173913

1e-07 , 0.1 -> 0.695652173913

1e-07 , 0.5 -> 0.695652173913

1e-07 , 0.9 -> 0.695652173913

1e-06 , 1e-06 -> 0.70652173913

1e-06 , 1e-05 -> 0.70652173913

1e-06 , 0.0001 -> 0.70652173913

1e-06 , 0.001 -> 0.70652173913

1e-06 , 0.01 -> 0.70652173913

1e-06 , 0.1 -> 0.70652173913

1e-06 , 0.5 -> 0.70652173913

1e-06 , 0.9 -> 0.70652173913

1e-05 , 1e-06 -> 0.70652173913

1e-05 , 1e-05 -> 0.70652173913

1e-05 , 0.0001 -> 0.70652173913

1e-05 , 0.001 -> 0.70652173913

1e-05 , 0.01 -> 0.70652173913

1e-05 , 0.1 -> 0.70652173913

1e-05 , 0.5 -> 0.70652173913

1e-05 , 0.9 -> 0.70652173913

0.0001 , 1e-06 -> 0.70652173913

0.0001 , 1e-05 -> 0.70652173913

0.0001 , 0.0001 -> 0.70652173913

0.0001 , 0.001 -> 0.70652173913

0.0001 , 0.01 -> 0.70652173913

0.0001 , 0.1 -> 0.70652173913

0.0001 , 0.5 -> 0.70652173913

0.0001 , 0.9 -> 0.70652173913

0.001 , 1e-06 -> 0.717391304348

0.001 , 1e-05 -> 0.717391304348

0.001 , 0.0001 -> 0.717391304348

0.001 , 0.001 -> 0.717391304348

0.001 , 0.01 -> 0.717391304348

0.001 , 0.1 -> 0.717391304348

0.001 , 0.5 -> 0.70652173913

0.001 , 0.9 -> 0.70652173913

0.01 , 1e-06 -> 0.684782608696

0.01 , 1e-05 -> 0.684782608696

0.01 , 0.0001 -> 0.684782608696

0.01 , 0.001 -> 0.684782608696

0.01 , 0.01 -> 0.684782608696

0.01 , 0.1 -> 0.673913043478

0.01 , 0.5 -> 0.673913043478

0.01 , 0.9 -> 0.663043478261

1e-07 , 0.001 -> 0.739130434783

1e-07 , 0.01 -> 0.739130434783

1e-07 , 0.1 -> 0.739130434783

1e-07 , 0.3 -> 0.739130434783

1e-07 , 0.5 -> 0.739130434783

1e-07 , 0.7 -> 0.739130434783

1e-07 , 0.9 -> 0.739130434783

1e-06 , 0.001 -> 0.739130434783

1e-06 , 0.01 -> 0.739130434783

1e-06 , 0.1 -> 0.739130434783

1e-06 , 0.3 -> 0.739130434783

1e-06 , 0.5 -> 0.739130434783

1e-06 , 0.7 -> 0.739130434783

1e-06 , 0.9 -> 0.739130434783

1e-05 , 0.001 -> 0.760869565217

1e-05 , 0.01 -> 0.760869565217

1e-05 , 0.1 -> 0.760869565217

1e-05 , 0.3 -> 0.760869565217

1e-05 , 0.5 -> 0.760869565217

1e-05 , 0.7 -> 0.760869565217

1e-05 , 0.9 -> 0.75

0.0001 , 0.001 -> 0.760869565217

0.0001 , 0.01 -> 0.760869565217

0.0001 , 0.1 -> 0.760869565217

0.0001 , 0.3 -> 0.760869565217

0.0001 , 0.5 -> 0.760869565217

0.0001 , 0.7 -> 0.760869565217

0.0001 , 0.9 -> 0.760869565217

0.001 , 0.001 -> 0.75

0.001 , 0.01 -> 0.75

0.001 , 0.1 -> 0.760869565217

0.001 , 0.3 -> 0.760869565217

0.001 , 0.5 -> 0.760869565217

0.001 , 0.7 -> 0.760869565217

0.001 , 0.9 -> 0.760869565217

0.01 , 0.001 -> 0.728260869565

0.01 , 0.01 -> 0.728260869565

0.01 , 0.1 -> 0.717391304348

0.01 , 0.3 -> 0.695652173913

0.01 , 0.5 -> 0.684782608696

0.01 , 0.7 -> 0.684782608696

0.01 , 0.9 -> 0.695652173913

As observed here, l1_ratio is not impacting much on the results when alpha is less than 0.001. But on a bigger overview l1_ratio works better if it is approx 0.1.

Hence alpha is 0.0001 and l1_ratio is 0.1

4. No Regularization

No hyper parameter.......