Report

Question-4

Problem 1 : Lasso

Alpha varies: giving lower values of alpha results in better accuracy.

Alpha	Accuracy(%)
1	63.942
0.1	67.777
0.01	70.192
0.0001	71.634
0.00001	71.634

Problem 2 : Ridge

Alpha varies: giving lower values of alpha results in better accuracy.

Alpha	Accuracy(%)
1	63.942
0.1	71.634
0.01	71.634
0.0001	72.115
0.00001	72.596

Problem 3 : Elastic Net

- Since ElasticNet is a combination of Ridge and Lasso, the first observation was related to l1_ratio.
- For L1_ratio <=0.2, maximum number of iterations required were lesser and thus, it behaved as Ridge.
- For L1_ratio >=0.8, maximum number of iterations required were atleast 100 and thus, it behaved as Lasso.
- ElasticNet accuracy was better with lower values of alpha.
- One more observation is that, for greater values of alpha, not only does the accuracy increase, but it also starts becoming less dependent on the hyper-parameter - L1_ratio.

Alpha	l1_ratio	Accuracy(%)
0.01	0.1	69.23
0.01	0.2	69.23
0.01	0.6	70.67
0.01	0.8	69.71
0.01	0.95	70.19
0.001	0.1	70.67
0.001	0.2	70.19
0.001	0.6	70.19
0.001	0.8	70.19
0.001	0.95	70.19

Problem 4 : No regularisation (Logistic Regression)

In regularization, the cost function includes a regularization expression, and keep in mind that the C parameter in sklearn regularization is the inverse of the regularization strength.

C in this case is 1/lambda, subject to the condition that $C > 0$.

Therefore, when C approaches infinity, then lambda approaches 0. When this happens, then the cost function becomes your standard error function, since the regularization expression becomes, for all intensive purposes, 0.

C	Accuracy(%)
100	69.71
100000	70.67
1000000000	71.63

We observe that, as the value of C becomes larger, the accuracy increases for a non regularised Logistic Regression.