Report on Assignment #4

IRM2015003

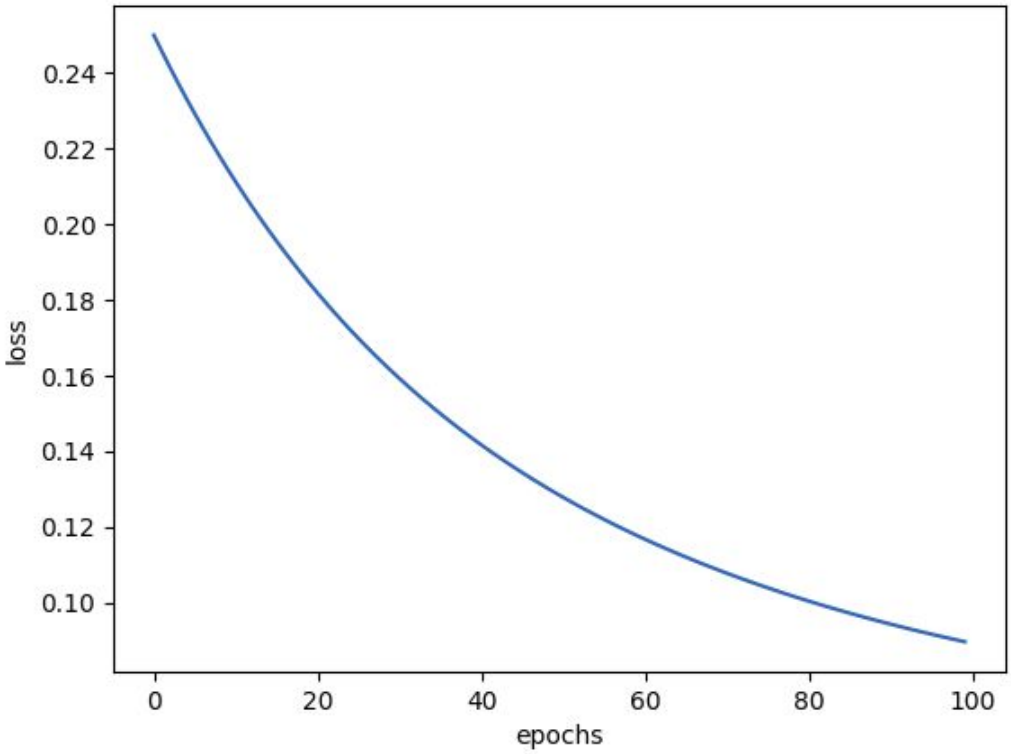
Some observations after implementing the logistic regression algorithm on both the datasets.(Using newton's method)

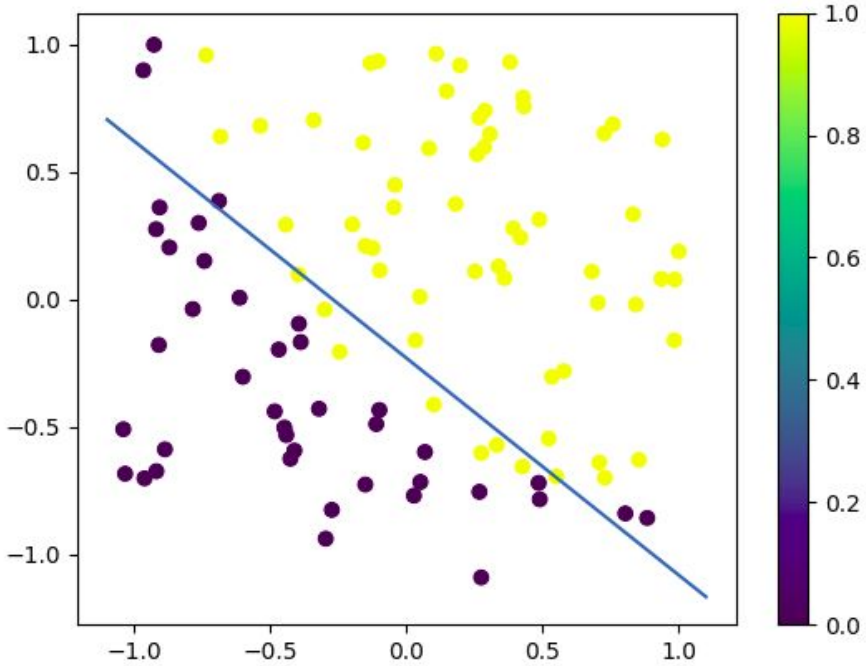
The newton's method although faster is computationally more expensive as it has $O(N^3)$ overhead of inversing a hessian matrix. The algorithm converges very fast as compared to the alpha learning owing to adaptive nature of learning rate in form of hessian matrix.

The calculation of hessian matrix is itself more computationally expensive than normal alpha learning. The hessian matrix is the second derivative of the cost function.

Examination dataset:

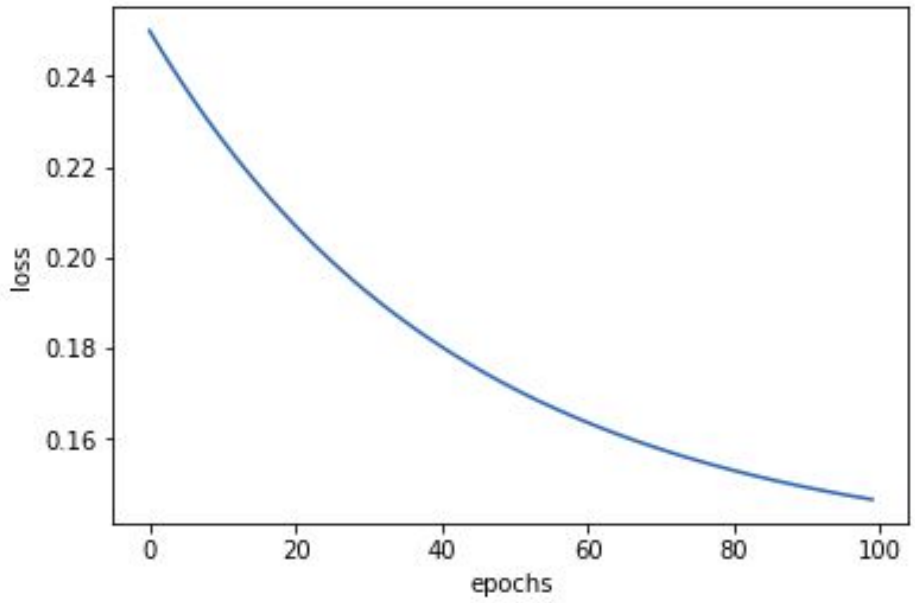
1. For this dataset the input were in order of 10s (around 70-80), so the data needed to be normalized in order to make the hypothesis less sensitive to the data.
2. The data was separable by a straight line.
3. The convergence was very fast fast as compared to alpha learning method.
4. The accuracy achieved was very good around 90%-95%.

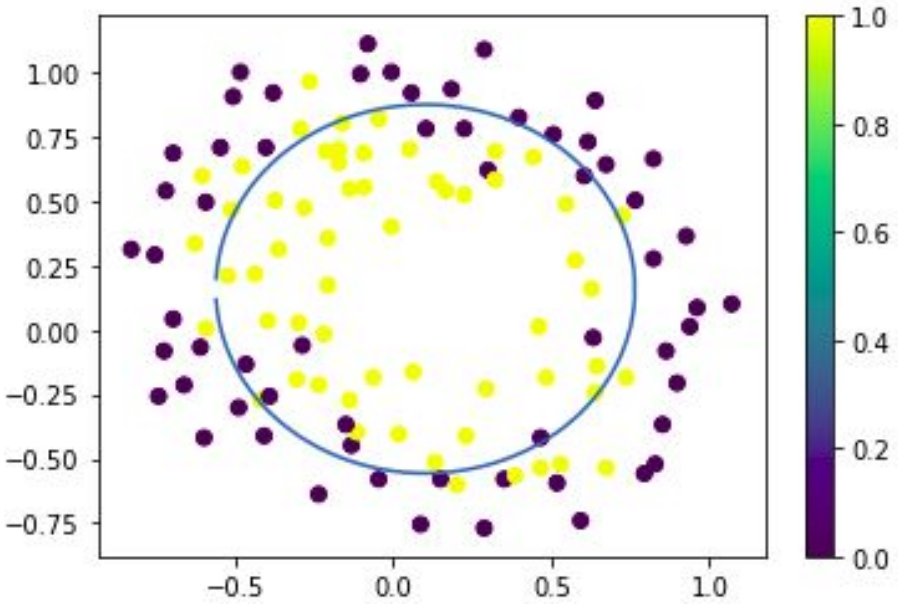

0.24
0.22
0.20
0.18
0.16
0.14
0.12
0.10
loss
0
20
40
60
80
100
epochs


1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
1.0
0.8
0.6
0.4
0.2
0.0

Microchip testing dataset:

1. For this data it was observed from visually plotting the data that the data was circular it was concluded that the data cannot be linearly separated. When normal($w_0 + w_1x_1 + w_2x_2$) data was used to apply logistic regression it showed poor results owing to its circular nature. Shown in figure below.
2. The problem was that we were trying to define a decision boundary that was a straight line for circular data. This was solved fitting a ellipse to the data so that better classification can be done.the equation used was $w_0 + w_1X_1^2 + w_2x_2^2$.
3. The accuracy was found to around 80%-85%.




1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−0.5
0.0
0.5
1.0
1.0
0.8
0.6
0.4
0.2
0.0