ASSIGNMENT 4 - IIT2016039

Problem Statement 1:

We are given a microchip product dataset. Using various parameters in the datasets we need to predict the acceptance or rejection using Logistic regression with delta learning rule using Newton's method and compare the results with LOG(using gradient descent)

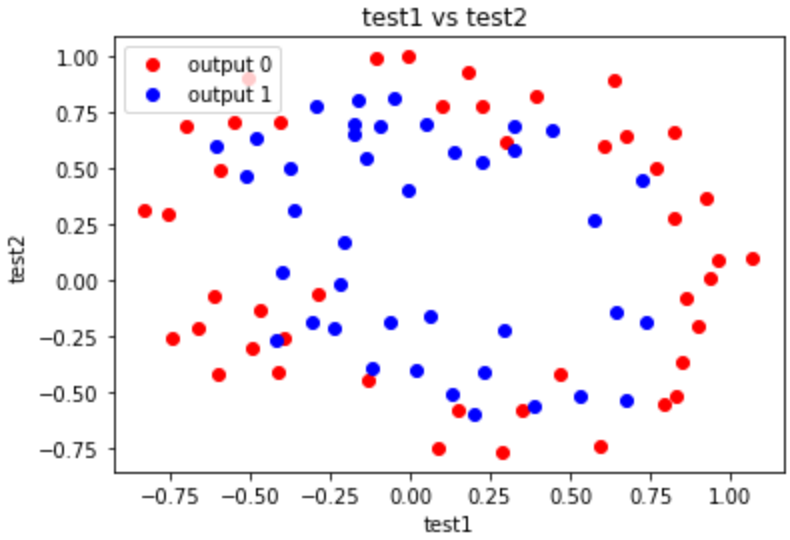
Analysis :

The Output for this can be seen in IIT2016039_Assignment4.html

In the given data we have two features test1 and test2 and we have 118 samples.

We have separated the data into 82 training samples and 36 testing samples.

First we will see the distribution of the dataset with x-axis test1 and y-axis test2 and red are rejected chips and blue are accepted chips. The graph is as follows



So we can say that we need an elliptical structure to fit the data so we are taking the following features X = [1, test1, test2, test1^2, test2^2, test1*test2]

Our weights will be W = [W0, W1, W2, W3, W4, W5]'

Logistic regression predictor function H(X) = 1 / (1 + e^(- W.X))

For H(x) <= 0.5 we take it as 0 i.e W.X >= 0

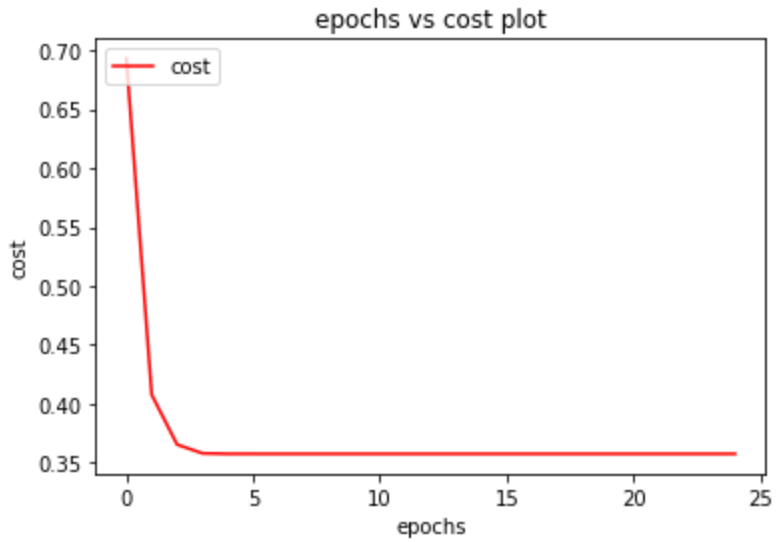
For H(x) > 0.5 we take it as 1 i.e W.X >= 0

Cost function for logistic regression is - (1 / len(Y)) * (sum(Y * log(H) + (1 - Y)*log (1 - H)))

Now while updating the weights in Newton's method we will multiply the Jacobian which is first derivative of cost function with the inverse of the Hessian matrix which is the second derivative of the cost function. By using the Newton the weights are converging within 6 or 7 epochs.

Jacobian J = (H (X) - Y) * X) Hessian = (H (X) * (1 - H (X))) * X * X

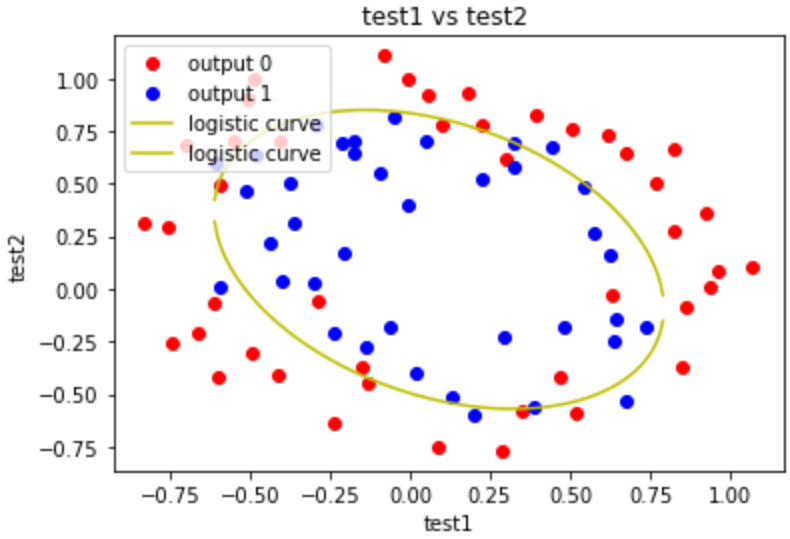
Where H (X) is the prediction we made.
After applying gradient descent with 25 epochs we got Final cost of 0.35739965755444253
graph is as follows :



We got weights W = [4.77123022,
3.19540742,
3.9091817,
-11.77419592,
-11.52007239,
-7.62224018].

When we predicted with the 36 test samples we got 32 of the predictions correctly So accuracy is 88.88888888888889.

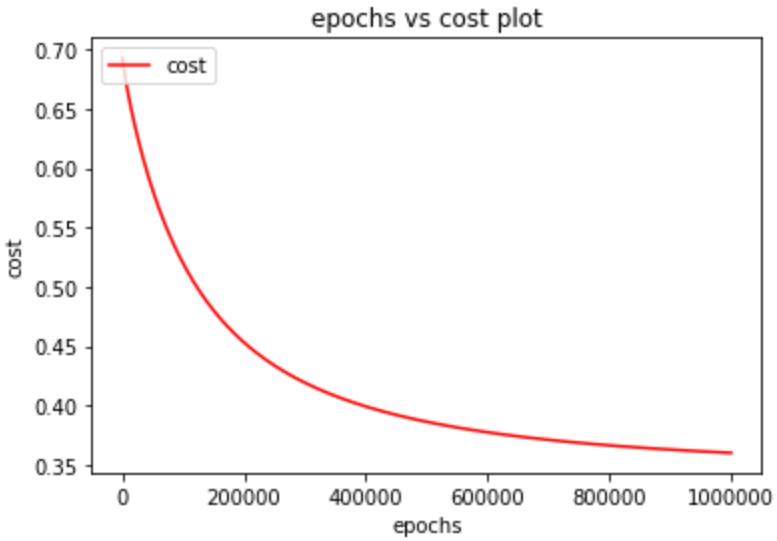
When we draw the graph between two features and then plotting the curve y = - ((W2 + W5 * x) (+/-) (W2 + (W5 * x)) * ((W2 +(W5 * x))) - (4 * W4 * (W0 + (W3 * x * x) + (W1 * x)))) / (2 * W4) predicted is



Comparison

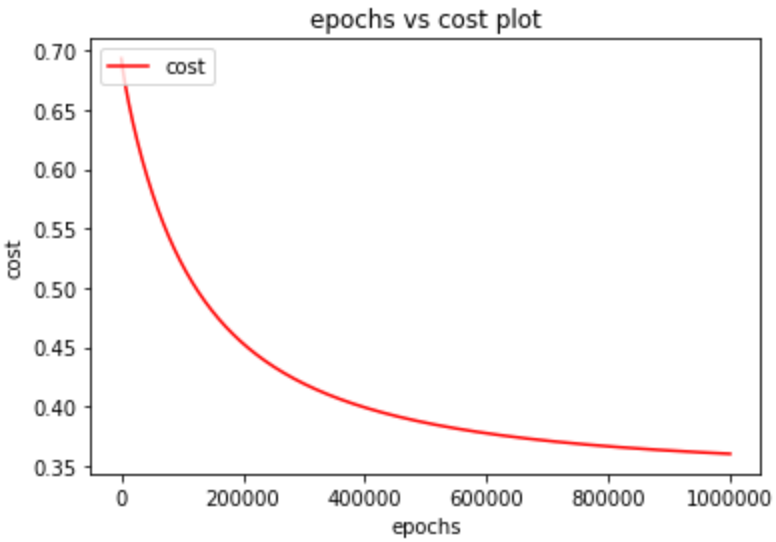
Normal Logistic Regression with Gradient Descent

- ➔ It is taking more than 10,00,000 epochs to converge .
- ➔ After 10,00,000 epochs the cost was 0.3602501194734718
- ➔ Weights we got are W = [3.86038376 2.22604687 2.94760951 -8.95932878 -8.59366206 -5.02147124]
- ➔ When we predicted with whole data we go 19 wrong predictions i.e accuray = 83.898305084



Logistic Regression with Delta Learning Rule:

- ➔ It took just 6 epochs to converge.
- ➔ After 6 epochs the cost was 0.3481053127977324 which is less compared to that of normal method.
- ➔ Weights we got are W = [5.16939737 3.2475016 4.16623174 -12.02686332 -11.8224845 -7.53125756].
- ➔ When we predicted with whole data we go 17 wrong predictions i.e accuray = 85.593220338.


epochs vs cost plot
cost
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
cost
0
200000
400000
600000
800000
1000000
epochs