Simulation of Offline 3

Calculations shown here are not exactly accurate but should be close enough for you to understand.

1. Steps 1 & 2 are trivial
2. Let's say I have the following training set.

Glucose	Blood Pressure	Diabetes?
148	72	1
8	-6	0
183	64	0

   As there are two features, parameter vector will consist of three values.

3. Let's say we initialize it to = [1 2 4]; It's completely random and dependent on your choice.
4. Then, the computation of cost function will be following:

   Input vector for row 1 = [1 148 72]

   Hypotheses value for input row 1 = g(1*1 + 148 * 2 + 72 * 4)

   = g(585)

   = 1/ (1 + e^(-585))

   = 0.99999

   Similarly, hypotheses value for input row 2 = g(1*1 + 8*2 -6*4)

   = 0.00091

   And, hypotheses value for input row 3 = 0.999999

   So, J(Θ) = -1/3 * [{1*log(0.9999) + 0 * log (0.00001)} + {0 * log(0.00091) + 1 * log(0.99909)}

   + {0 * log(1) + 1 * log(0.000001)}]

   = -1/3 * (0 + 0 – 6)

   = - 1/3 * (-6)

   = 2

5. Let's say the learning rate, alpha = 0.05

6. While J(Θ) is not close to 0:
   a. Inner loop iteration 1 (Updating 1st parameter through gradient descent):
      $Θ_0$ = 1 – 0.05 / 3 * [(0.9999-1) *1 + (0.00091-0) * 1 + (0.999999 – 0) * 1]
      = 0.0167

Inner loop iteration two (Updating 2nd parameter through gradient descent):
$\Theta_1 = 2 – 0.05/3 * [(0.9999-1) *148 + (0.00091-0) * 8 + (0.999999 – 0) * 183]$
$= 3.05$
Inner loop iteration three (Updating 3rd parameter through gradient descent):
$\Theta_2 = 4 – 0.05/3 * [(0.9999-1) *72 + (0.00091-0) * (-6) + (0.999999 – 0) * 64]$
$= 1.08$

So, after step 5(a), the new parameter vector becomes = [0.0167 3.05 1.08]

b. Compute J(Θ) again as shown in step 4

Check whether the value of J(Θ) is close to 0 or difference between previous J(Θ) and current J(Θ) is close to 0.

If it is, abort the loop.

7. Done with the training set :D
8. Let's say I have the following test set.

Glucose	Blood Pressure	Diabetes?
8	5	0
83	123	1
183	64	1

9. Let's say at the end your parameter vector becomes:

[0 2 1]

10. So, hypothesis value for test row 1 = g(1*0 + 8*2 + 5*1)

= g(21)

= 0.99

So, your prediction for test row 1 = 1 but actual output = 0

Similarly, your prediction for test row 2 = 1 and actual output = 1

Similarly, your prediction for test row 3 = 1 and actual output = 1

11. So, accuracy is = 2/3 * 100% = 66.67%