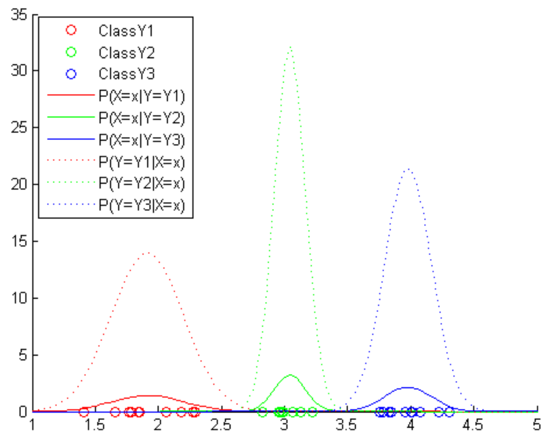
CSL407 Machine Learning

HW2

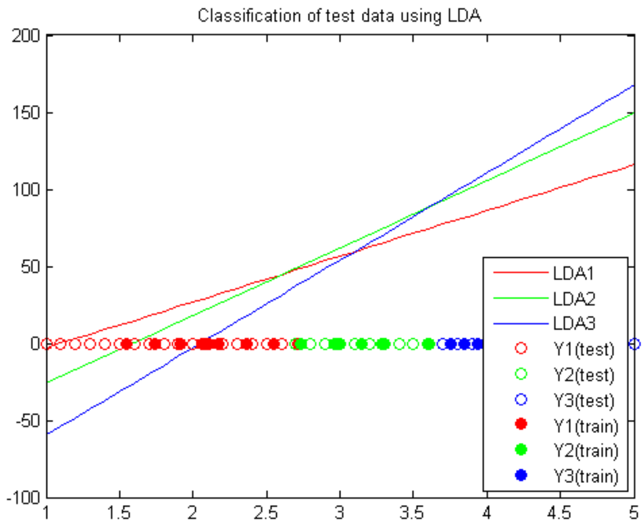
Jaskaran Singh(2011cs1012)

Ans1

Plot1



Plot2



Ans2

Regularized Logistic Regression Cost Function

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m} y^{(i)} log h_\theta(x^{(i)}) + (1 - y^{(i)})\log(1 - h_\theta(x^{(i)})) + \frac{\lambda}{2m}\sum_{j=1}^{n}\theta_j^{\,2}$$

Derivation of weight update equation for the Logistic regression function without regularization

$$\frac{\delta J}{\delta \theta_j} = -\frac{1}{m}\sum_{i=1}^{m}(y^i\left(\frac{1}{h_\theta(x^i)}\right) + \left(1-y^i\right)(\frac{-1}{1-h_\theta(x^i)}))\frac{\delta h_\theta(x^i)}{\delta \theta_j}$$

Solving which gives

$$1)\frac{\delta J}{\delta \theta_j} = -\frac{1}{m}\sum_{i=1}^{m}\frac{y^i - h_\theta(x^i)}{h_\theta(x^i)(1-h_\theta(x^i))}\frac{\delta h_\theta(x^i)}{\delta \theta_j}$$

Since,

$$h_\theta\left(x^i\right) = \frac{1}{1+e^{-\theta^T x^i}}$$

Solving for $\frac{\delta h_\theta(x^i)}{\delta \theta_j}$

$$2)\frac{\delta h_\theta(x^i)}{\delta \theta_j} = \frac{-1}{\left(1+e^{-\theta^T x^i}\right)^2}\left(-x_j^i\right)\left(e^{-\theta^T x^i}\right) = (h_\theta(x^i))^2\left(x_j^i\right)\left(\frac{1}{h_\theta(x^i)}-1\right)$$
$$= h_\theta\left(x^i\right)(1-h_\theta\left(x^i\right))x_j^i$$

Substituting 2) in 1)

We get

$$3)\frac{\delta J}{\delta \theta_j} = -\frac{1}{m}\sum_{i=1}^{m}\frac{y^i - h_\theta(x^i)}{h_\theta(x^i)(1-h_\theta(x^i))}h_\theta\left(x^i\right)(1-h_\theta\left(x^i\right))x_j^i = \frac{1}{m}\sum_{i=1}^{m}(h_\theta\left(x^i\right)-y^i)x_j^i$$

For regularized logistic regression,

To 3) we add the derivative of the regularized term

So,

For j=0

$$\frac{\delta J}{\delta \theta_0} = \frac{1}{m}\sum_{i=1}^{m}(h_\theta\left(x^i\right)-y^i)x_0^i$$

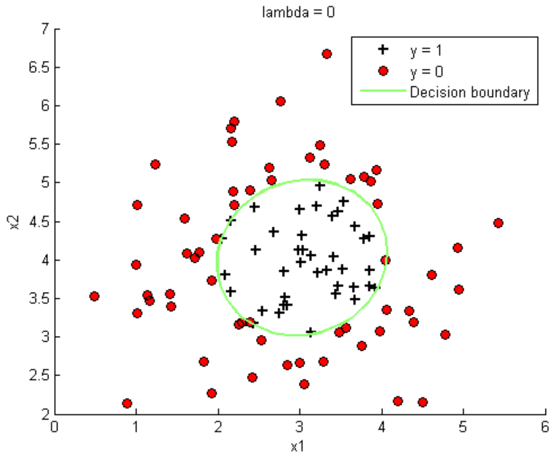
And for j=1,2,3,….,n

$$\frac{\delta J}{\delta \theta_j} = \frac{1}{m}\sum_{i=1}^{m}(h_\theta\left(x^i\right)-y^i)x_j^i + \frac{\lambda}{m}\theta_j$$

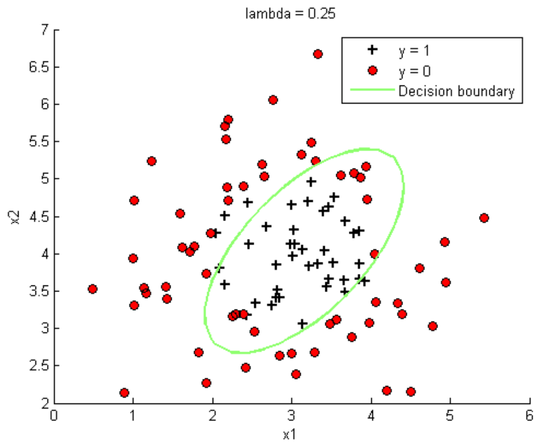
Ans3

Varying the lamda values in the set {0,0.25,0.5,0.75,1}

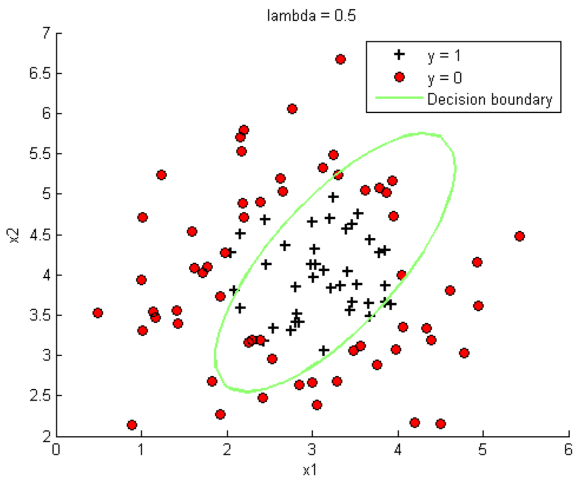
Lambda=0,Training Accuracy=98%(Overfitting)



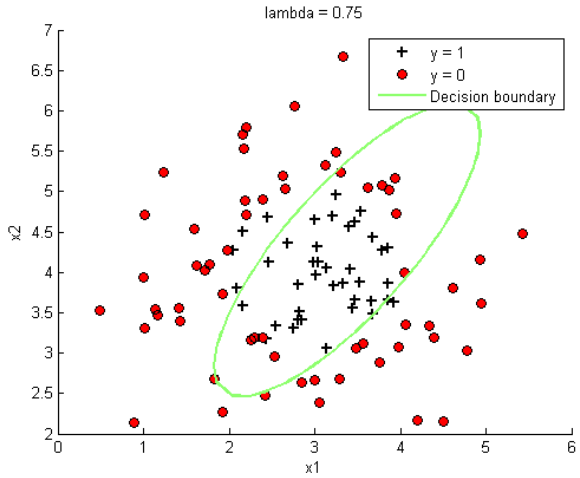
Lambda=0.25,Training Accuracy-86%



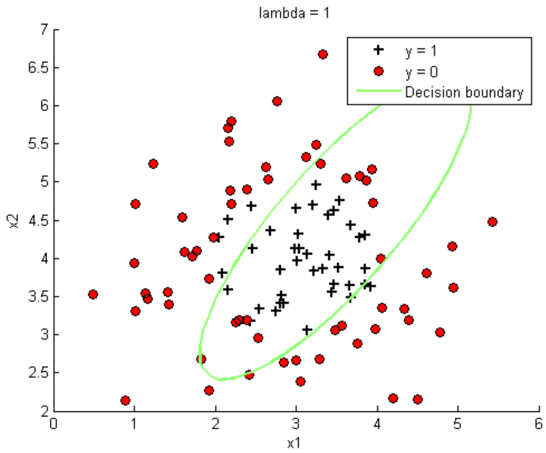
Lambda=0.50,Training Accuracy=83%



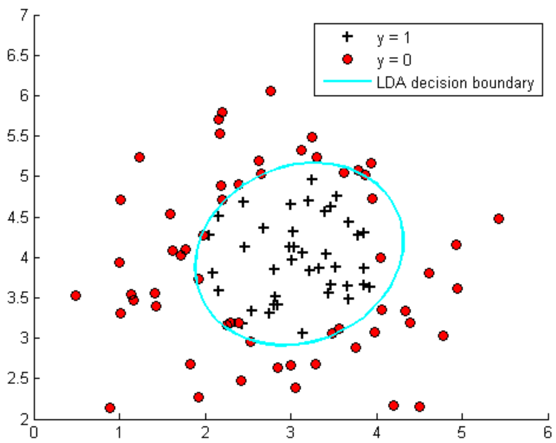
Lambda=0.75,Training Accuracy=82%



Lambda=1,Training Accuracy=81%(Underfitting)



The linear decision boundary(very similar to plotdecisionboundary where lambda=0)



Ans4

a)

x1=5hrs,x2=7.5gpa

So,

$$\theta^T X = -8 + .05(5) + 7.5 = -0.25$$

Thus,

$$h_\theta(x) = \frac{1}{1 + e^{0.25}} = 0.4378$$

b)

Since,

$$h_\theta(x) = \frac{1}{1 + e^{-\theta^T X}} = 0.6$$

$$\theta^T X = -ln(\frac{1}{.6} - 1) = \ln(\frac{3}{2}) = -8 + 0.05(x1) + 7.5$$

Solving for x1

We get x1=(0.5+ln(3/2))/0.05=18.11hrs