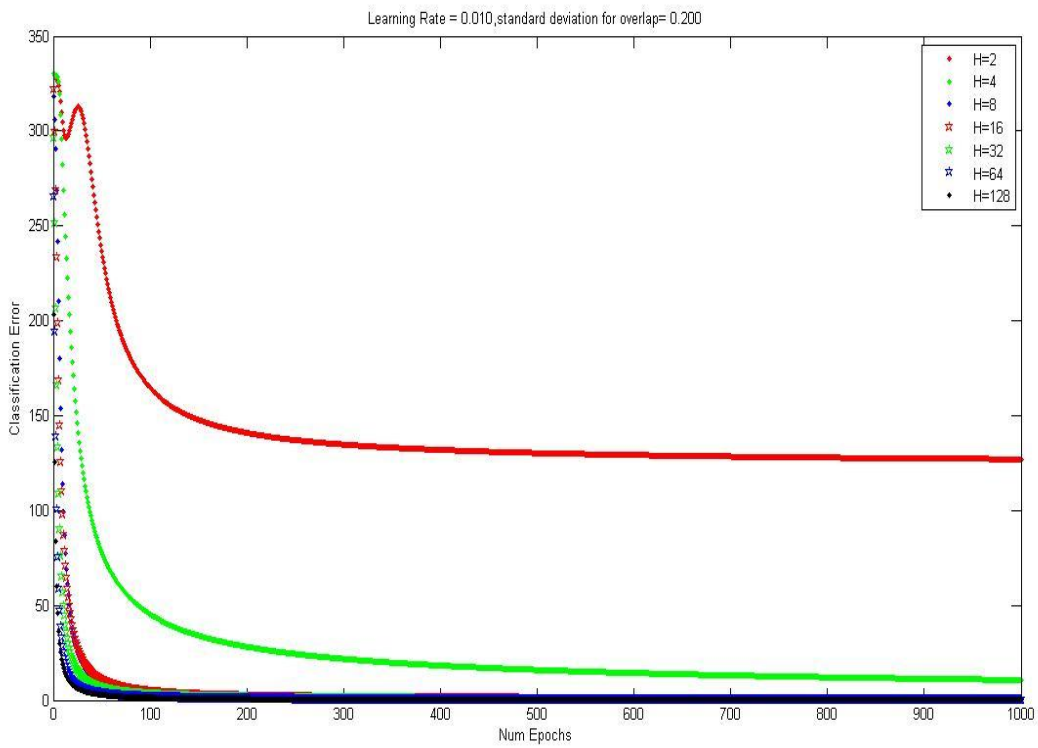
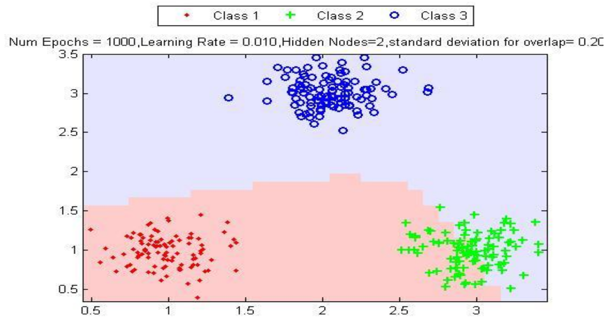
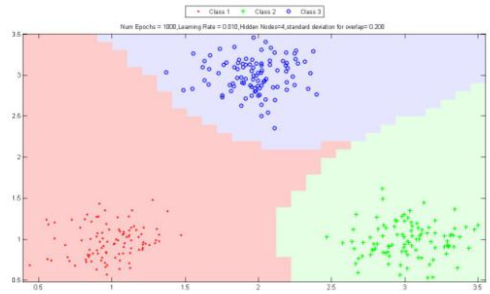
Q1 (d) Figure below shows training error vs epoch for varying number of hidden nodes

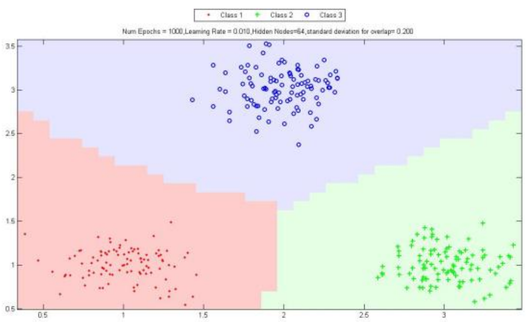


From the plot above we can observe that training accuracy increases with increasing number of hidden layer nodes. Initially this increase in accuracy is quite observable, but the increase in training accuracy is not observed much after H=16.

Following 3 figures show decision boundaries for H=2, 4 and 64 respectively

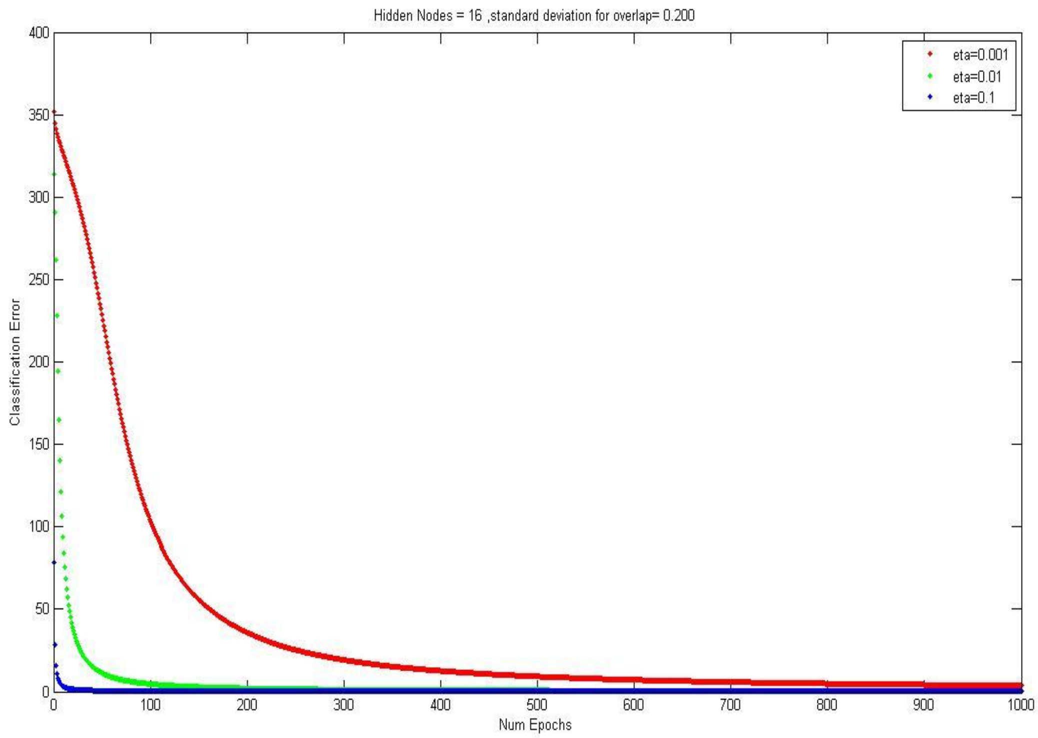






From above 3 figures we can observe that for H=2 and H=4 , there is underfitting. Accuracy is very less in case of H=2. But it gets better for H=64

Q1(e) Figure below shows training error vs. Epoch for varying learning rate



From above figure, we can see that, training error drops quickly as eta increases from 0.001 to 0.1. This is due to fact that weights change very slowly for eta=0.001. Thus weight change per epoch is very less for smaller eta.

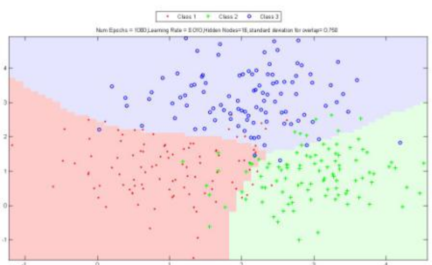
Q1 (f) Training errors were observed as under:

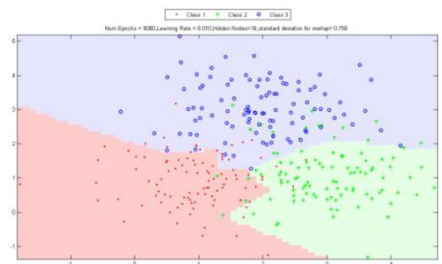
1000 epoch: 115.85

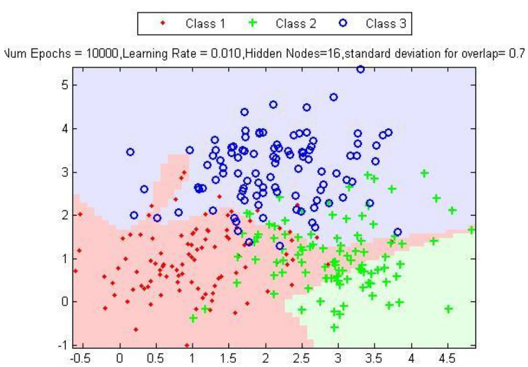
5000 epoch: 92.16

10000 epoch: 85.95

Classification boundary diagrams for epoch =1000, 5000 and 10000 are as under respectively







From the above figures we can see that increasing number of epochs result in over fitting of training data. This over fitting is easily observable for epoch=10000

Q2

First Iteration (Input1)

- C value=0.5498
- D value=0.5387
- W_{C0} =0.1034, W_{CA} =0.1034, W_{CB} =0.1
- V_{D0} =0.2384, V_{DC} =0.1761

First Iteration (Input2)

- C value=0.550
- D value=0.583
- W_{C0} =0.098, W_{CA} =0.106, W_{CB} =0.092
- V_{D0} =0.188, V_{DC} =0.148

Second Iteration (Input1)

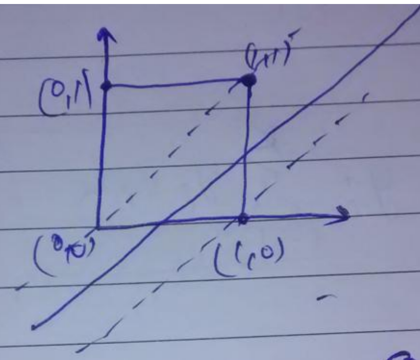
- C value=0.551
- D value=0.567
- W_{C0} =0.099, W_{CA} =0.114, W_{CB} =0.085
- V_{D0} =0.272, V_{DC} =0.194

Second Iteration (Input2)

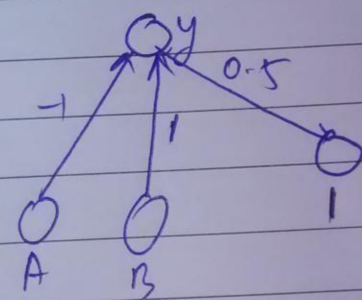
- C value=0.546
- D value=0.593
- W_{C0} =0.091, W_{CA} =0.120, W_{CB} =0.085
- V_{D0} =0.170, V_{DC} =0.139

Q3

A	B	$A \wedge \neg B$
0	0	0
0	1	0
1	0	1
1	1	0



Perceptron →



Q4 (a)

$C(14) = \text{Min} \int$

amazon.com Date:

2 4 8 16 32 64 128

amazon.com Date: ....

Q(a) $E = -\sum_{i=1}^{k} t_i \log y_i$

$$\therefore \frac{\partial E}{\partial v_{Ih}} = \frac{-\partial}{\partial v_{Ih}} \left(t_1 \log y_1 + t_2 \log y_2 - \cdots - t_k \log y_k \right)$$

$$\frac{\partial}{\partial v_{Ih}} = -\left(\frac{t_1}{y_1} \frac{\partial y_1}{\partial v_{Ih}} + \frac{t_2}{y_2} \frac{\partial y_2}{\partial v_{Ih}} \cdots \frac{t_k}{y_k} \frac{\partial y_k}{\partial v_{Ih}} \right)$$

$$= -\left(\sum_{i=1}^{k} \frac{t_i}{y_i} \left(\frac{\partial y_i}{\partial v_{Ih}} \right) \right)$$

Now,

$$\frac{\partial y_I}{\partial v_{Ih}} = \partial \left(\frac{e^{\sum_{h=1}^{H+1} z_h v_{Ih}}}{\sum_{i=1}^{K} e^{\sum_{h=1}^{H+1} z_h v_{Ih}}} \right) \Big/ \partial v_{Ih}$$

$$= \frac{\left(e^{\sum_{h=1}^{H+1} z_h v_{Ih}} \cdot z_h \right) - \left(e^{\sum_{h=1}^{H+1} z_h v_{Ih}} \right)^2 \cdot z_h}{\sum_{i=1}^{K} e^{\sum_{h=1}^{H+1} z_h v_{ih}}}$$

$$\boxed{\frac{\partial y_I}{\partial v_{Ih}} = y_I z_h (1 - y_I)}$$

amazon.com Date: ....

Similarly

$$\left(\frac{\partial y_i}{\partial v_{Ih}}\right)_{i \neq I} = -y_i \, y_I \, z_h$$

$$\therefore \frac{\partial E}{\partial v_{Ih}} = -\Big(-t_1 y_I z_h - t_2 y_I z_h \;--\; -- + t_i(1-y_I) z_h \;---\; -- - t_k y_I z_h\Big)$$

$$\cancel{= t_I z_h}$$

$$= -t_I z_h + \left(\sum_{i=1}^{k} t_i\right) y_I z_h$$

since $\sum_{i=1}^{k} t_i = 1$

$$\therefore \boxed{\frac{\partial E}{\partial v_{Ih}} = (y_I - t_I) z_h = \Delta v_{ih}}$$

amazon.com Date: ....

$$\frac{\partial E}{\partial w_{hj}} = \frac{-\partial}{\partial w_{hj}}\left(t_1 \log(y_1) - \cdots + t_k \log(y_k)\right)$$

$$= -\left(\frac{t_1}{y_1}\frac{\partial y_1}{\partial z_h}\frac{\partial z_h}{\partial w_{hj}} + \cdots + \frac{t_k}{y_k}\frac{\partial y_k}{\partial z_h}\frac{\partial z_h}{\partial w_{hj}}\right)$$

$$\frac{\partial y_i}{\partial z_h} = \partial\left(\frac{e^{\sum_{h=1}^{H+1} z_h v_{ih}}}{\sum_{i=1}^{K} e^{\sum_{h=1}^{H+1} z_h v_{ih}}}\right) \Big/ \partial z_h$$

$$= \frac{e^{\sum_{h=1}^{H+1} z_h v_{ih}} \cdot v_{ih}}{\sum_{k=1}^{K} e^{\sum_{h=1}^{H+1} z_h v_{ih}}} - \frac{e^{\sum_{h=1}^{H+1} z_h v_{ih}}}{\left(\sum_{k=1}^{K} e^{\sum_{h=1}^{H+1} z_h v_{ih}}\right)^2}\left(e^{\sum_{h=1}^{H+1} z_h v_{ih}} v_{ih} - \cdots + e^{\sum_{h=1}^{H+1} z_h v_{kh}} v_{kh}\right)$$

$$\boxed{\frac{\partial y_i}{\partial z_h} = y_i v_{ih} - y_i \sum_{i=1}^{K} y_i v_{ih}}$$

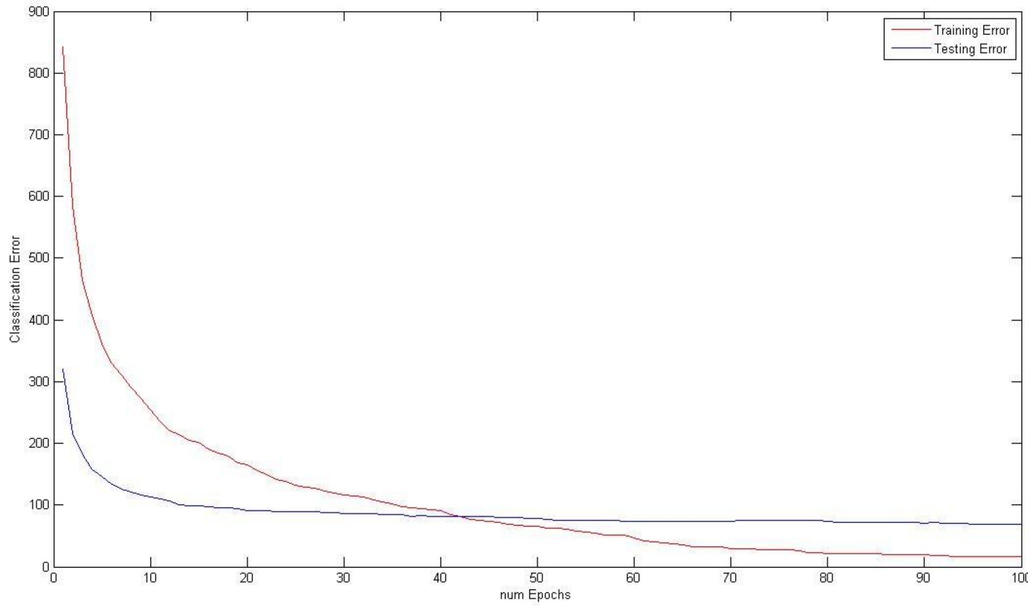
Date: ....

$$\frac{\partial z_h}{\partial w_{hj}} = \frac{\partial \left(\frac{a-b}{a+b} \right)}{\partial w_{hj}} \quad \text{where} \quad a = e^{\sum_{j=1}^{H+1} w_{hj} x_j}, \; b = \frac{1}{a}$$

$$= x_j (1 - z^2)$$

$$\therefore \boxed{\frac{\partial E}{\partial w_{hj}} = -x_j (1 - z^2) \sum_{i=1}^{k} (t_i - y_i) v_{ih} = \Delta w_{hj}}$$

Q4 (c)


Training Error
Testing Error
Classification Error
num Epochs
0
100
200
300
400
500
600
700
800
900
10
20
30
40
50
60
70
80
90