Machine Learning - HW #2

1) Decision regions induced by the following instances using Euclidean distance:

Let $A = (4,4)$ - pos

$B = (4,6)$ - pos

$C = (6,2)$ - neg

Euclidean distance AB $(4,4)\ (4,6) = \sqrt{(4-4)^2 + (4-6)^2} = \sqrt{0+4} = 2$

$d(A(4,4)) \leq d(B(4,6))$

$\Rightarrow \sqrt{(x_1-4)^2 + (x_2-4)^2} \leq \sqrt{(x_1-4)^2 + (x_2-6)^2}$

$\Rightarrow \cancel{x_1^2} - \cancel{8x_1} + \cancel{16} + \cancel{x_2^2} - 8x_2 + 16 \leq \cancel{x_1^2} - \cancel{8x_1} + \cancel{16} + \cancel{x_2^2} - 12x_2 + 36$

$\Rightarrow -8x_2 + 12x_2 + 16 - 36 \leq 0$

$\Rightarrow 4x_2 - 20 \leq 0 \Rightarrow \boxed{x_2 - 5 \leq 0} \rightarrow (1)$

Euclidean distance BC $(4,6)\ (6,2) = \sqrt{(6-4)^2 + (2-6)^2} = \sqrt{4+16} = 2\sqrt{5}$

$d(B(4,6)) \leq d(C(6,2))$

$\Rightarrow \sqrt{(x_1-4)^2 + (x_2-6)^2} \leq \sqrt{(x_1-6)^2 + (x_2-2)^2}$

$\Rightarrow \cancel{x_1^2} - 8x_1 + 16 + \cancel{x_2^2} - 12x_2 + 36 \leq \cancel{x_1^2} - 12x_1 + 36 + \cancel{x_2^2} - 4x_2 + 4$

$\Rightarrow -8x_1 - 12x_2 + 52 \leq -12x_1 - 4x_2 + 40$

$\Rightarrow 4x_1 - 8x_2 + 12 \leq 0 \Rightarrow \boxed{x_1 - 2x_2 + 3 \leq 0} \rightarrow (2)$

Euclidean distance CA $(6,2)\ (4,4) = \sqrt{(6-4)^2 + (2-4)^2} = \sqrt{4+4} = 2\sqrt{2}$

$d(C(6,2)) \leq d(A(4,4))$

$\Rightarrow \sqrt{(x_1-6)^2 + (x_2-2)^2} \leq \sqrt{(x_1-4)^2 + (x_2-4)^2}$

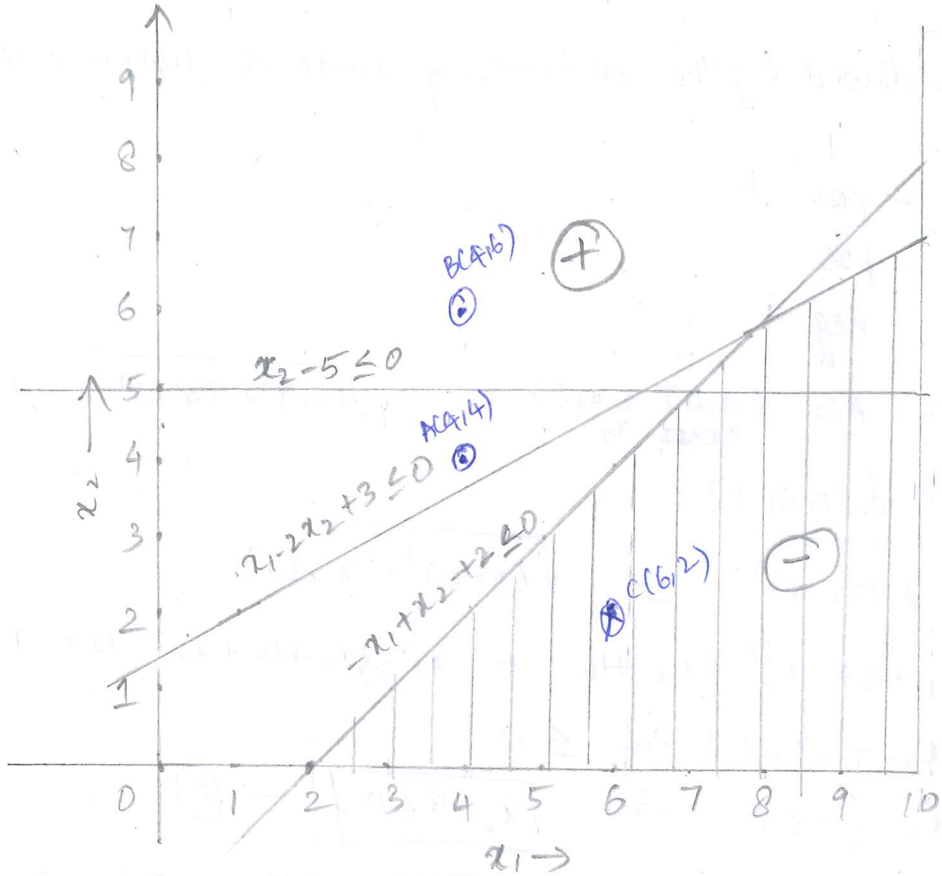
$\Rightarrow \cancel{x_1^2} - 12x_1 + 36 + \cancel{x_2^2} - 4x_2 + 4 \leq \cancel{x_1^2} - 8x_1 + 16 + \cancel{x_2^2} - 8x_2 + 16$

$\Rightarrow -12x_1 - 4x_2 + 40 \leq -8x_1 - 8x_2 + 32$

$\Rightarrow -4x_1 + 4x_2 + 8 \leq 0$

$\Rightarrow \boxed{-x_1 + x_2 + 2 \leq 0} \rightarrow (3)$

Decision boundaries obtained from Eq ①, ② and ③:



2. Nearest Neighbours of $x^{(q)} = (7,10)$

given points : a(2,11) b(3,12) c(5,10) d(2,8) e(2,4) f(6,3) g(9,1) h(12,5) i(10,10) j(13,12)

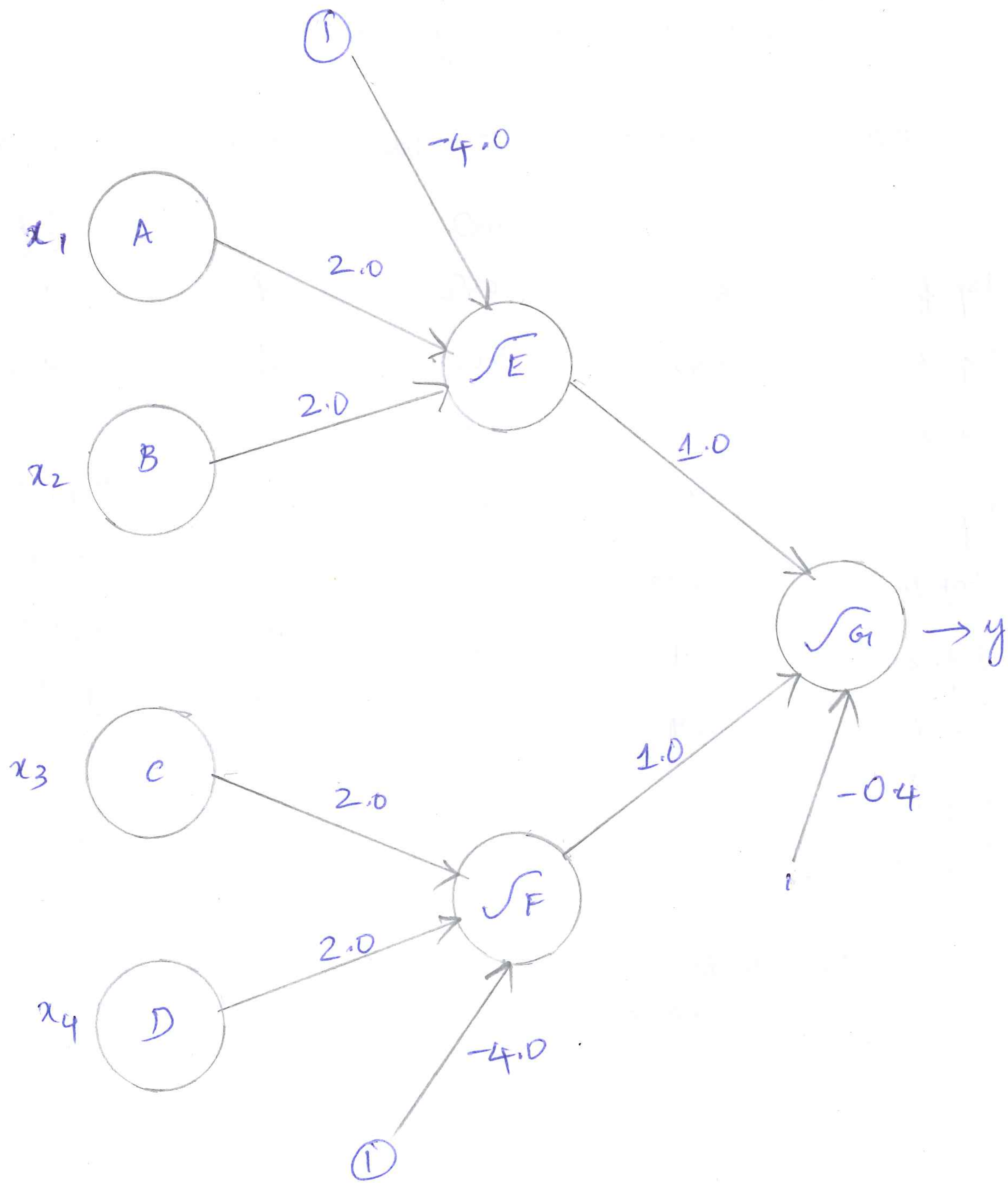
OPERATION	DISTANCE	BEST DISTANCE	BEST NODE	PRIORITY QUEUE
		∞		{f,0}
Pop f	$5\sqrt{2}$	$5\sqrt{2}$	f	(h,0) (c,1)
Pop h	$5\sqrt{2}$	$5\sqrt{2}$	f	(i,0) (c,1) (g,5)
Pop i	3	3	i	(c,1) (j,3) (g,5)
Pop c	2	2	c	(b,0) (e,0) (j,3) (g,5)
Pop b	$2\sqrt{5}$	2	c	(e,0) (j,3) (a,4) (g,5)
Pop e	$\sqrt{61}$	2	c	(d,0) (j,3) (a,4) (g,5)
Pop d	$\sqrt{29}$	2	c	(j,3) (a,4) (g,5)
Pop j				
Return c				

∴ Best Node = c

Best distance = 2

3. Neural network for logical function $y = (x_1 \wedge x_2) \vee (x_3 \wedge x_4)$

- By trial and error,



$w_{EA} = 2$

$w_{EB} = 2$

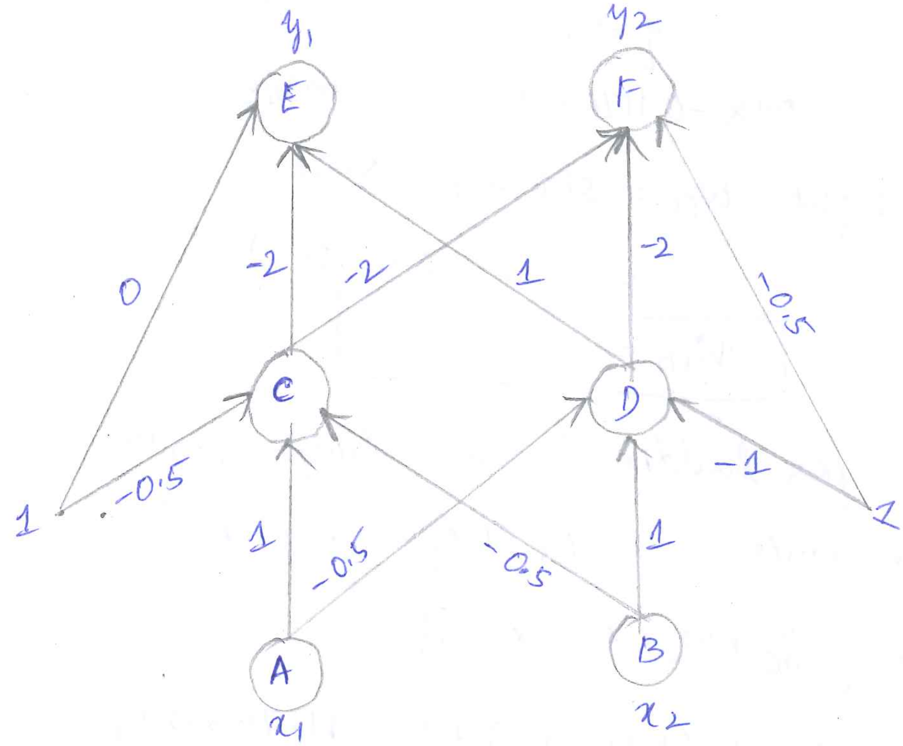
$w_{FC} = 2$

$w_{FD} = 2$

$w_{GE} = 1.0$

$w_{GF} = 1.0$

4.



Training Instance : $x = [0,1]$ $y = [1,0]$

Net I/p to C = $(0 \times 1) + (1 \times -0.5) + (1 \times -0.5) = -1$

$O_C = \dfrac{1}{1+e^{-(-1)}} = \dfrac{1}{1+e} = 0.2689$

Net I/p to D = $(1 \times 1) + (0 \times -0.5) + (1 \times -1) = 0$

$O_D = \dfrac{1}{1+e^0} = 0.5$

Net I/p to E = $(1 \times 0) + (0.2689 \times -2) + (0.5 \times 1) = -0.5379 + 0.5 = -0.0379$

$O_E = \dfrac{1}{1+e^{0.0379}} = 0.4905$

Net I/p to F = $(1 \times -0.5) + (0.5 \times -2) + (0.2689 \times -2) = -0.5 + (-1) + (-0.5379)$
$= -2.0378$

$O_F = \dfrac{1}{1+e^{2.0378}} = 0.11529 \cong 0.1153$

Error in Output Units : $\delta_j = O_j(1-O_j)(y_j - O_j)$

$\delta_E = 0.4905 \times (1-0.4905) \times (1-0.4905) = 0.1273$

$\delta_F = 0.1153 \times (1-0.1153) \times (0-0.1153) = -.01176$

Weight updation for erroneous o/p units: $\Delta w_{ji} = \eta \delta_j O_i$

$\eta = 0.1$

$\therefore \Delta W_{FD} = \eta \delta_F O_D = 0.1 \times -0.1176 \times 0.5 = -0.000588$

$\therefore$ New weight $W_{FD} = \text{Old } W_{FD} + \Delta W_{FD}$

$= -2 + (-0.000588)$

$\boxed{\therefore W_{FD} = -2.000588}$

Weight updation for hidden layers: $\Delta w_{ji} = \eta \delta_j O_i$

Error in Hidden units $\delta_j = O_j (1 - O_j) \sum_k \delta_k w_{kj}$

$\delta_D = O_D \times (1 - O_D) [\delta_E \times W_{ED} + \delta_F \times W_{FD}]$

$= 0.5 \times (1 - 0.5) \times [(0.1273 \times 1) + (-0.01176 \times -2)]$

$= 0.5 \times 0.5 \times (0.1273 + 0.02352) = 0.37705$

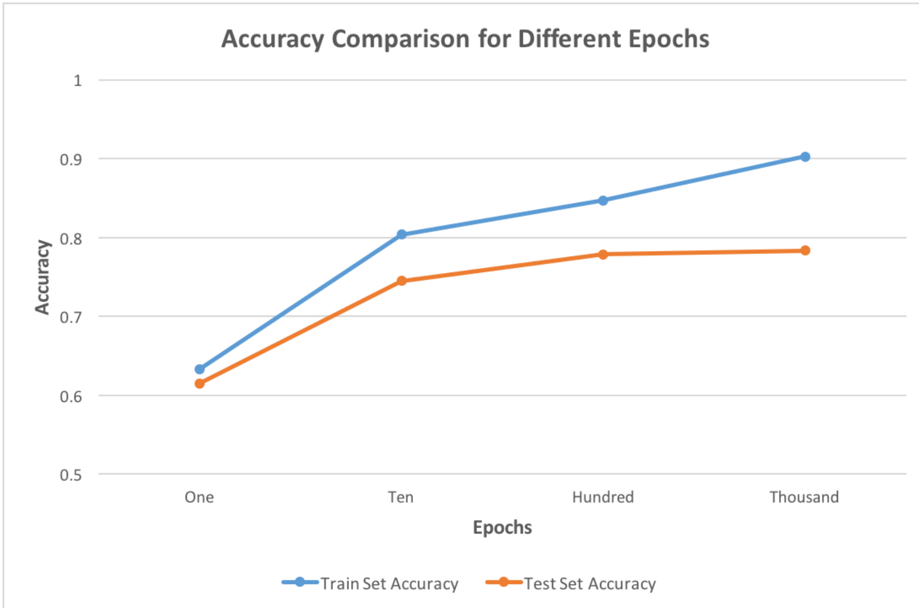
$\therefore \Delta W_{DA} = \eta \delta_D O_A = 0.1 \times 0.37705 \times 0 = 0$

$\therefore$ New weight $W_{DA} = \text{old } W_{DA} + \Delta W_{DA}$

$= -0.5 + 0$

$\therefore \boxed{W_{DA} = -0.5}$

5. Graph showing training and test set accuracy after 1, 10, 100 and 1000 training epochs with 10-fold cross validation:



6. ROC curve for a run of stratified 10-fold cross validation with a learning rate of 0.1 and 100 training epochs: