Q2)

a	b	d
1	0	1
0	1	0

$\eta = 0.3$

Initial weights

$w = [0.1, 0.1, 0.1]$

$v = [0.1, 0.1]$

momentum, $\alpha = 0.9$

Iteration 1

a) Input 1 $(a=1, b=0)$

$c = 0.5498$

$d = 0.5387$

Updated $w = [0.1034, 0.1034, 0.1]$

updated $v = [0.2384, 0.1761]$

b) Input 2 $(a=0, b=1)$

$c = 0.5507$

$d = 0.5831$

Updated $w = [0.0989, 0.1065, 0.0924]$

Updated $v = [0.1880, 0.1483]$

Iteration 2

a) Input 1 ($a=1$, $b=0$)

$c = 0.5512$

$d = 0.5670$

updated $w =$ [0.996, 0.1140, 0.085]

updated $V =$ [0.2726, 0.1948]

b) Input 2 ($a=0$, $b=1$)

$c = \cancel{0.55}$: 0.5461

$d = 0.5936$

updated $w =$ [0.0916, 0.1208, 0.0707]

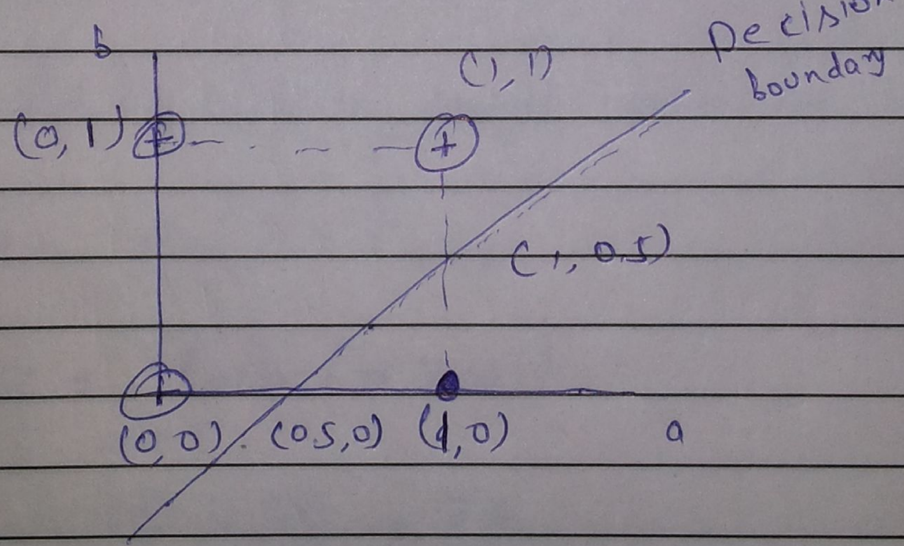
updated $v =$ [0.1706, 0.1394]

Q3 A ∧ ¬B

Truth table for A ∧ ¬B

A	B	A∧¬B
0	0	0
0	1	0
1	0	1
1	1	0

On a 2-d graph, data can be represented as



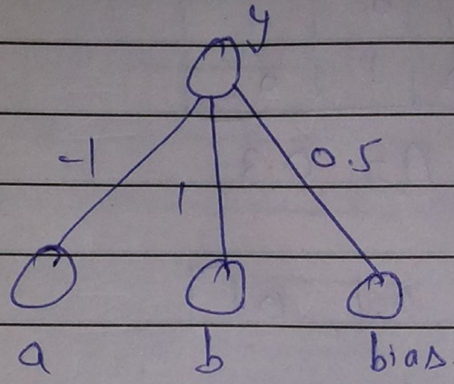
Equation of decision boundary is $a - b - 0.5 = 0$

or $b - a + 0.5 = 0$

for $b - a + 0.5 > 0$ class is positive

and for $b - a + 0.5 < 0$ class is negative

Hence the perceptron model for the given function is



Q 4

(a) $Z = \tanh(\cancel{\omega x^T}\ X W^T)$

$$y^i = \frac{\exp(Z V^T)}{\sum_{i=1}^{K} \exp(Z V_K^T)} = \sigma(Z V^T)$$

σ: softmax

Error function, $E = -\sum_{n=1}^{N} \sum_{i=1}^{K} t_{ni} \log(y^i_{ni})$

Let us calculate weight updates for single forward pass.

$$\frac{\partial E}{\partial V} = -\sum_{i=1}^{K} \frac{t_i}{y_i} \frac{\partial (y_i)}{\partial V}$$

$$= -\sum_{i=1}^{K} \frac{t_i}{y_i} \frac{\partial(\sigma(Z V^T))}{\partial V}$$

$$= \sum_{i=1}^{K} \partial\left(\frac{\sigma(Z V^T)}{\partial (Z V^T)} \times \frac{\partial (Z V^T)}{\partial V}\right)$$

$$= -\sum_{i=1}^{K} \frac{t_i}{y_i} y_i (1-y_i) Z$$

Let us calculate weight updates for $k=2$

$$\frac{\partial E}{\partial V_{ih}} = -t_i(1-y_i)z_h + (1-t_i)y_i z_h$$

$$= (y_i - t_i) z_h$$

Weigh update equation of V

~~$\Delta V / V = V_{ih} = V_{ih} +$~~ $\eta\left(\frac{-\partial E}{\partial V_{ih}}\right)$

$$\Delta V_{ih} = \eta\left(\frac{-\partial E}{\partial V_{ih}}\right)$$

$$= \eta(t_i - y_i) z_h$$

Now let us calculate weight updates for w

$$y = \sigma(zV^T),$$

$$z = \tanh(XW^T)$$

$$\frac{\partial E}{\partial w} = \frac{\partial E}{\partial y} \times \frac{\partial E}{\partial z} \times \frac{\partial z}{\partial w}$$

$$= \left[\frac{t}{y} \cdot y(1-y)V + \frac{(1-t)(1-y)(-y)V}{1-y}\right] \times \left(\frac{\partial z}{\partial(xw^T)}\right) \times \left(\frac{\partial xw^T}{\partial w}\right)$$

$$\Delta w_{hj} = \sum_{i=1}^{K} (t_i - y_i) V_h \left\{ \frac{\partial \tanh(xw^T)}{\partial (xw^T)} \right\} \times \frac{\partial x w^T}{\partial w}$$

$$= \sum_{i=1}^{K} (t_i - y_i) V_h \left[1 - \tanh(x w_{hj}^T)^2 \right] . x_j$$

$$= \sum_{i=1}^{K} (t_i - y_i) V_h \cdot \left[1 - Z_h^2 \right] . x_j$$

$$\Delta w_{hj} = \eta \left(\sum_{i=1}^{K} t_i - y_i \right) \cancel{V_{ih}}\, V_{ih} \left[1 - Z_h^2 \right] x_j$$