HomeWork-2

CSC 578
Name: Milin Desai

4.7 two layer feedforward ANN

input : a, b
hidden layer : c
output layer : d

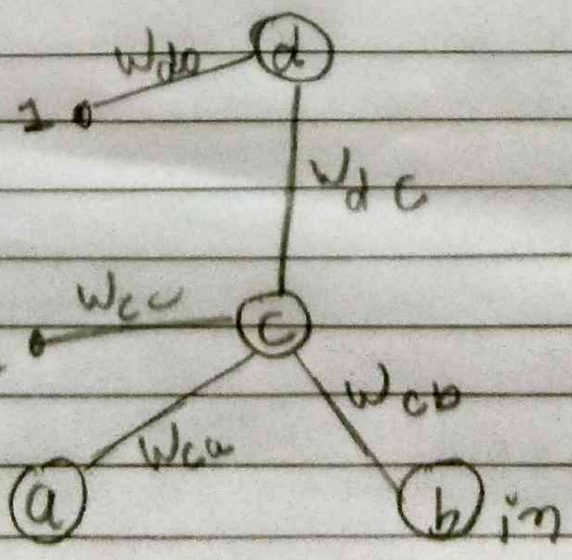
Weights : 5

$w_{ca}, w_{cb}, w_{c0}, w_{dc}, w_{d0}$

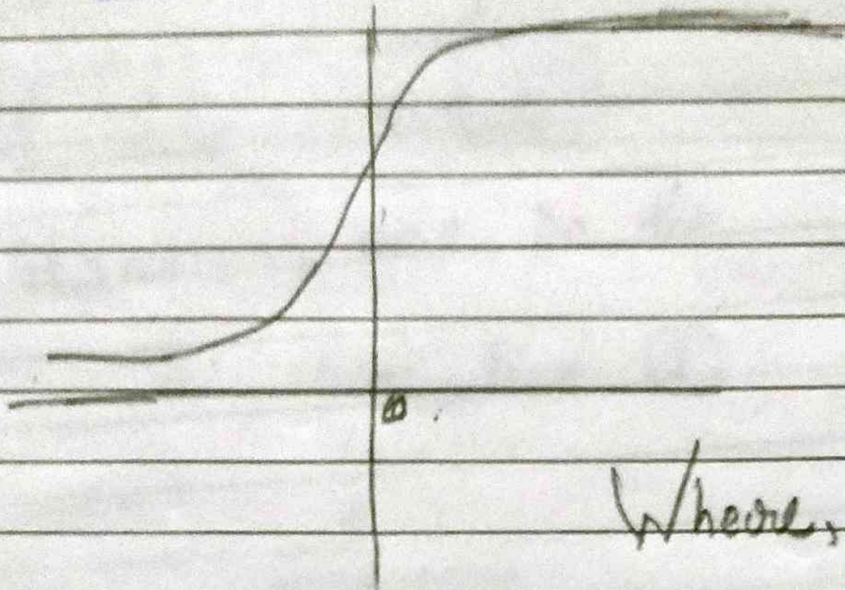
Threshold weight w_{x0}

$\eta = 0.3$
$\alpha = 0.9$ (momentum)

	a	b	d
iteration 1	1	0	1
iteration 2	0	1	0



As Sigmoid Activation function:-



Where,

$$\phi(z) = \frac{1}{1+e^{-z}}$$

In Our Case,

$$net = \sum_{i=0}^{n} w_i x_i$$

Where

w = weight of inputs

x = Input

here,

$$\boxed{O = \sigma(net) = \frac{1}{1+e^{-net}} \quad (1)}$$

From eq - 1,

$$\boxed{\sigma(d) = \frac{1}{1 + e^{-y}}} \quad — \ (2)$$

Applying iteration one in the equation (2)

$$O_c = \frac{1}{1 + e^{-y}}$$

where $y = w_i \cdot x_i$

$$\therefore \frac{1}{1 + e^{-\{1(0.1) + 0(0.1) + 0.1(0.1)\}}}$$

$$= \frac{1}{1 + e^{-(0.2)}} = \frac{1}{1 + 0.82}$$

$$= \frac{1}{1.82} = \underline{0.55}$$

Same for O_d,

$$O_d = 0.54$$

Note:- we will use O_c as input of the O_d

Step 2:-

Calculating error term δ_k for each input

$$\delta_k \rightarrow O_k(1-O_k)(t_k-O_k)$$

for us,

$$\delta_d = O_d(1-O_d)(t_d-O_d) \quad \because t_d = 1$$

$$= 0.54(1-0.54)(1-0.54)$$

$$\boxed{= 0.114}$$

Step 3:-

Calculate error term δ_c

$$\delta_c \leftarrow O_c(1-O_c)\sum w_{dc}\delta_d$$

$$= (0.55)(1-0.55)(0.1)(0.114)$$

$$\boxed{= 0.0028}$$

Step 4 :-

$$\omega_{ji} \leftarrow \omega_{ji} + \Delta \omega_{ji} \rightarrow \text{(I)}$$

Where

$$\Delta \cdot \omega_{ji} = \eta \delta_j x_{ji} + \alpha \Delta \omega_{ji}(n-1) \rightarrow \text{(P)}$$

~~(Add~~.

for Our Case,

$$\Delta \cdot \omega_{dc} = \eta \delta_d O_c + \alpha \Delta \omega_{dc}(n-1)$$

$$\because \alpha \Delta \omega_{dc}(n-1) = 0$$

So,

$$\Delta \omega_{dc} = \eta \delta_d O_c \quad (\because \eta = 0.3)$$

$$= (0.3)(0.114)(0.55)$$

$$\boxed{\Delta \omega_{dc} = 0.018}$$

from (I)

$$\omega_{dc} = \omega_{dc} + \Delta \cdot \omega_{dc}$$

$$= 0.1 + 0.018$$

$$\boxed{\omega_{dc} = 0.118}$$

DATE / /

From equation (I) & (P), ~~We Can~~ After Computing

we will get below Values

$\Delta W_{do} = 0.034$, $W_{do} = 0.134$

$\Delta W_{co} = 0.0008$, $W_{co} = 0.1008$

$\Delta W_{ca} = 0.0008$, $W_{ca} = 0.1008$

$\Delta W_{cb} = 0$, $W_{cb} = 0.1$

For iteration 2 $a = 0$, $b = 1$, $d = 0$

Applying in equation (2)

$$O_c = \frac{1}{1 + e^{-\{(0.1008)(0) + (0.1008)(0) + (0.1)(1)\}}}$$

$$= 0.55$$

$$O_d = \frac{1}{1 + e^{-\{(0.134)(1) + (0.118)(0.55)\}}}$$

$$= 0.54$$

Step 2 :-

Calculating error term δ_d

$\delta_d = O_d (1 - O_d)(t_d - O_d)$ ∴ $t_d = 0 = d$

$= (0.54)(0.45) \cdot (-0.54)$

$= -0.136$

For δ_c

$\delta_c = O_c (1 - O_c) \sum w_{dc} \cdot \delta_d$

$= (0.55)(0.45)(-0.136)(0.118)$

$= -0.004$

updating weights of each network

from equation (P)

$\Delta w_{do} = \eta \delta_d O_{do} + \alpha \Delta w_{do}(n-1)$

$= (0.3)(-0.136)(1) + (0.9)(0.034)$

$= -0.009$

$w_{do} = w_{do} + \Delta w_{do}$

$= 0.134 + (-0.009) = 0.125$

Likewise for

$\Delta w_{dc} = -0.005$, $w_{dc} = 0.113$

$\Delta w_{ca} = 0.0007$, $w_{ca} = 0.101$

$\Delta w_{c0} = -0.0004$, $w_{c0} = 0.100$

$\Delta w_{cb} = -0.0012$, $w_{cb} =$ 0.099

4.8 Backpropagation => tanh

where

$$O = \tanh(\vec{w} \cdot \vec{x})$$

Weight update rule output & hidden layers ?

→

$$O = \tanh(\vec{w} \cdot \vec{x})$$

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

Simple form for derivative

$$\tanh'(x) = 1 - \tanh^2(x)$$

step 2 Calculating error term δ_k

$$\delta_k \leftarrow O_k (1 - O_k)(t_k - O_k) \quad \text{— (1)}$$

From equation (1)

$$\delta_k = \tanh^2(x)\,(t_k - o_k)$$

Step 3:

$$\delta_a \leftarrow o_a\,(1 - o_a) \sum w_{ka}\,\delta_k$$

$$= \tanh^2(x) \sum w_{ka}\,\delta_k$$

Step 4:

$$\Delta w = \eta\,\delta_k \cdot x_j$$

$$= \eta\left(\tanh^2(x)\,(t_k - o_k)\right)$$

$$w \leftarrow w_j + \Delta w$$

$$= w_j + \eta\left(\tanh^2(x)\,(t_k - o_k)\right)$$