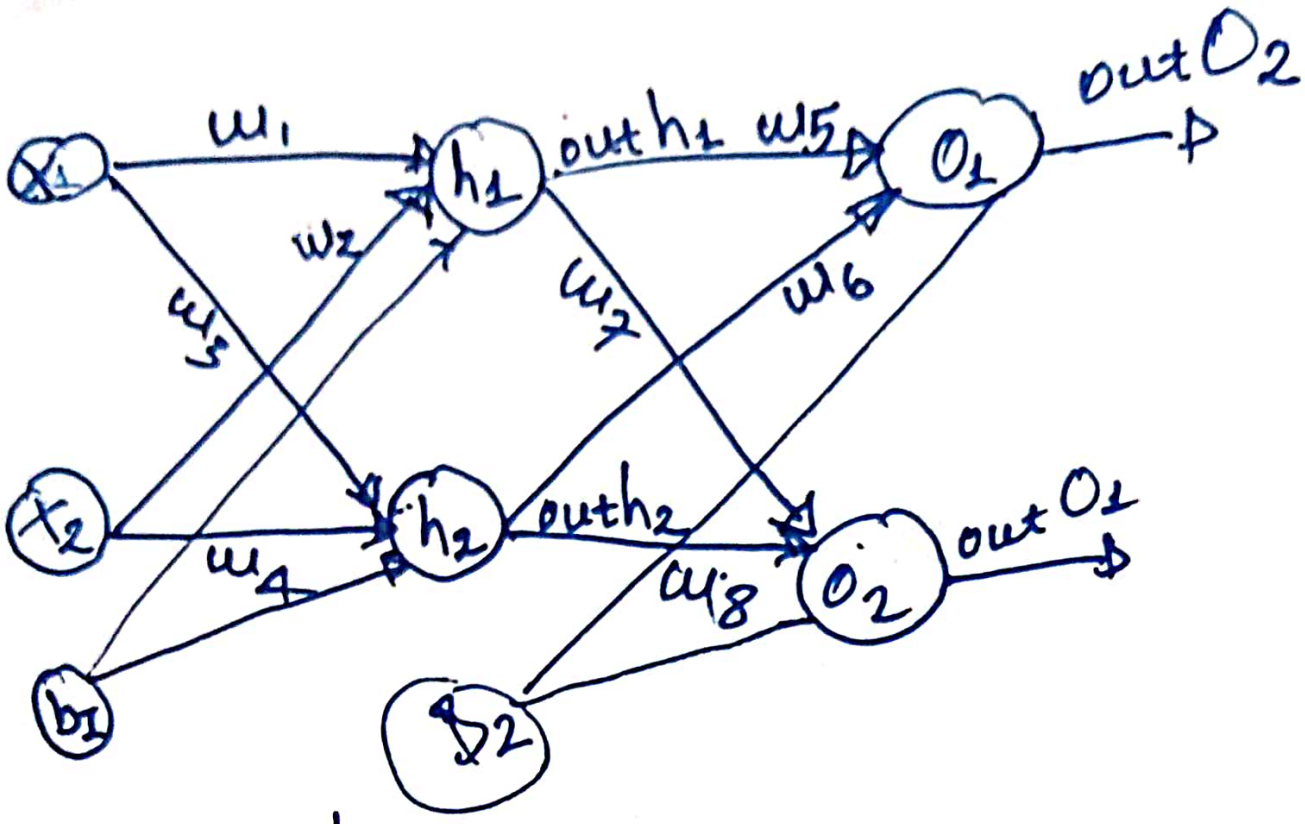
1



$$\underline{\text{Sigmoid}} = \frac{1}{1+e^{-x}}$$

$$h_1 = w_1 * x_1 + w_2 * x_2 + b_1$$

$$out h_1 = \frac{1}{1+e^{-h_1}}$$

$$h_2 = w_3 * x_1 + w_4 * x_2 + b_1$$

$$out h_2 = \frac{1}{1+e^{-h_1}}$$

$\underline{out h_1}$ & $\underline{out h_2}$ is the output at the hidden layer

$$O_1 = w_5 * out h_1 + w_6 * out h_2 + b_2$$

$$out O_1 = \frac{1}{1+e^{-O_1}}$$

$$O_2 = w_7 * out h_1 + w_8 * out h_2 + b_2$$

$$out O_2 = \frac{1}{1+e^{-O_2}}$$

$E_1 = \frac{1}{2}(d_1 - out O_1)^2$ where d_1 & d_2 is the target (desired output at O_1 & O_2 respectively

$E_2 = \frac{1}{2}(d_2 - out O_2)^2$

$$E_{tot} = \frac{1}{2}\sum_{i=1}^{n}(d_i - out O_i)^2$$

$\eta = 0.5$. η is Learning rate the value of η is 0-1

⊛ When We do weight update

$w5_{new} = w5 - \eta * E_{tot}$

$w6_{new} = w6 - \eta E_{tot}$

$w7_{new} = w7 - \eta E_{tot}$

$w8_{new} = w8 - \eta E_{tot}$

$w1_{new} = w1 - \eta E_{tot}$

$w2_{new} = w2 - \eta E_{tot}$

$w3_{new} = w3 - \eta E_{tot}$

$w4_{new} = w4 - \eta E_{tot}$

⊛ This weight update is performed by Backpropagation

To updat $w5$ → we derivate E_{tot} by $w5$

$$\frac{\partial E_{tot}}{\partial w5} = \frac{\partial E_{tot}}{\partial w5} * \frac{\partial out O_1}{\partial O_1} * \frac{\partial E_{tot}}{\partial out O_1}$$

$$\frac{\partial E_{tot}}{\partial w5} = \frac{\partial O_1}{\partial w5} * \frac{\partial out O_1}{\partial O_1} * 2\left[\frac{\frac{1}{2}(d_1 - out O_1)^2 + \frac{1}{2}(d_2 + out O_2)^2}{\partial out O_1}\right]$$

$$\frac{\partial E_1}{\partial w_5} = \frac{\partial O_1}{\partial w_5} * \frac{\partial out O_1}{\partial O_1} * \frac{\partial E_1}{\partial out O_1}$$

$$\frac{\partial O_1}{\partial w_5} = \frac{\partial (out h_1 * w_5 + out h_2 * w_6 + b_2)}{\partial w_5}$$

$$\frac{\partial O_1}{\partial w_5} = 1 * out h_1 + 0$$

$$= \underline{out h_1}$$

$$\frac{\partial out O_1}{\partial O_1} = \frac{\partial \left(\frac{1}{1+e^{-O_1}}\right)}{\partial O_1}$$

$$\underline{out O_1 (1 - out O_1)}$$

$$\frac{\partial E_1}{\partial out O_1} = \frac{\partial \left(\frac{1}{2}(d_1 - out O_1)^2\right)}{\partial out O_1}$$

$$= 2 * \frac{1}{2}(d_1 - out O_1)^{2-1} * -1$$

$$= -(d_1 - out O_1)$$

$$= -d_1 + out O_1$$

$$w_{5new} = w_5 - \eta \frac{\partial E_{tot}}{\partial w_5}$$

$$w_{5new} = w_5 - \eta * (out h_1 * out O_1 (1 - out O_1) (-d_1 + out O_1))$$

$$w_{5new} = w_5 - \eta * (outh_1 * outO_1(1 - outO_1) * (-d_1 + outO_1)) \quad (2)$$

$$\frac{\partial E_1}{\partial w_6} = \frac{\partial O_1}{\partial w_6} * \frac{\partial outO_1}{\partial O_1} * \frac{\partial E_1}{\partial outO_1}$$

$$= outh_2 * outO_1(1 - outO_1) * (-d_1 + outO_1)$$

$$w_{6new} = w_6 - \eta \frac{E_1}{\partial w_6}$$

$$w_{6new} = w_6 - \eta (outh_2 * outO_1(1 - outO_1) * (-d_1 + outO_1))$$

$$\frac{\partial E_2}{\partial w_7} = \frac{\partial O_2}{\partial w_7} * \frac{\partial outO_2}{\partial O_2} * \frac{E_2}{\partial outO_2}$$

$$= outh_1 * outO_2(1 - outO_2) * (-d_2 + outO_2)$$

$$w_{7new} = w_7 - \eta * \frac{\partial E_2}{\partial w_7}$$

$$= w_7 - \eta (outh_1 * outO_2(1 - outO_2) * (-d_2 + outO_2))$$

$$* (d_1 + outO_1))$$

$$w_{8new} = w_8 - \eta \frac{\partial E_2}{\partial w_8}$$

$$\frac{\partial E_2}{\partial w_8} = \frac{\partial O_2}{\partial w_8} * \frac{\partial out O_2}{\partial O_2} * \frac{\partial E_2}{\partial out O_2}$$

$$\frac{\partial O_2}{\partial w_8} = \frac{\partial (out h_2 * w_8 + out h_1 * w_7 + b_2)}{\partial w_8}$$

(∂w_8 → 1, ∂ constant → 0)

$$= out h_2$$

$$\frac{\partial E_2}{\partial w_8} = out h_2 * out O_2 (1 - out O_2) * (-d_2 - out O_2)$$

$$w_{8new} = w_8 - \eta \left(out h_2 * out O_2 (1 - out O_2) * (-d_2 + out O_2 \right)$$

This all about weight update from output to hidden layer. w_5, w_6, w_7, w_8

Now let start update the weight from hidden layer to input layer w_1, w_2, w_3, w_4

$$w_{1new} = w_1 - \eta * \frac{\partial E_{tot}}{\partial w_1}$$

$$\frac{\partial E_{tot}}{\partial w_1} = \frac{\partial h_1}{\partial w_1} * \frac{\partial out h_1}{\partial h_1} * \frac{\partial E_{tot}}{\partial out h_1}$$

$$\frac{\partial E_{tot}}{\partial out h_1} = \frac{\partial E_1}{\partial out h_1} + \frac{\partial E_2}{\partial out h_1}$$

$$\frac{\partial E_{tot}}{\partial w_1} = \frac{\partial h_1}{\partial w_1} * \frac{\partial out h_1}{\partial h_1} * \left(\frac{\partial E_1}{\partial out h_1} + \frac{\partial E_2}{\partial out h_1}\right)$$

$$\frac{E_1}{\partial out h_1} = \frac{\partial O_1}{\partial out h_1} * \frac{\partial out O_1}{\partial O_1} * \frac{E_1}{\partial out O_1}$$

$$\frac{\partial E_2}{\partial out h_1} = \frac{\partial O_2}{\partial out h_1} * \frac{\partial out O_2}{\partial O_2} * \frac{\partial E_2}{\partial out O_2}$$

$$\frac{\partial E_{tot}}{\partial w_1} = \frac{\partial h_1}{\partial w_1} * \frac{\partial out h_1}{\partial h_1} * \left(\frac{\partial O_1}{\partial out h_1} * \frac{\partial out O_1}{\partial O_1} * \frac{E_1}{\partial out O_1}\right) + \left(\frac{\partial O_2}{\partial out h_1} * \frac{\partial out O_2}{\partial O_2} * \frac{\partial E_2}{\partial out O_2}\right)$$

$$w_{1\,new} = w_1 - \eta * \frac{\partial E_{tot}}{\partial w_1}$$

$$\frac{\partial h_1}{\partial w_1} = \frac{\partial (w_1 * x_1 + w_2 * x_2 + b_1)}{\partial w_1}$$

$$= x_1$$

$$\frac{\partial out h_1}{\partial h_1} = \frac{\partial \left(\frac{1}{1+e^{-h_1}}\right)}{\partial h_1}$$

$$= out h_1 (1 - out h_1)$$

$$\frac{\partial O_1}{\partial out h_1} = \frac{\partial (out h_1 * w_5 + out h_2 * w_6 + b_2)}{\partial out h_1}$$

$$= w_5$$

$$\frac{\partial out O_1}{\partial O_1} = \frac{\partial \left(\frac{1}{1+e^{-O_1}}\right)}{\partial O_2}$$

$$= out O_1 (1 - out O_1)$$

$$\frac{E_1}{\partial out O_1} = \frac{\partial \left(\frac{1}{2}(d_1 - out O_1)^2\right)}{\partial out O_1}$$

$$= -d_1 + out O_1$$

$$\frac{\partial O_2}{\partial out_{h_1}} = \frac{(out_{h_1} * w_7 + out_{h_2} * w_8 + b_2)}{\partial out_{h_1}}$$

$$= w_7$$

$$\frac{\partial out O_2}{\partial O_2} = \frac{\partial \left(\frac{1}{1+e^{-O_2}}\right)}{\partial O_2}$$

$$= out O_2 (1 - out O_2)$$

$$\frac{\partial E_2}{\partial out O_2} = \frac{\partial \left(\frac{1}{2}(d_2 - out O_2)^2\right)}{\partial out O_2}$$

$$= -d_2 + out O_2$$

$$\frac{\partial E tot}{\partial w_1} = X_1 * out h_1 (1 - out h_1) * \Big(\big(w_5 * out O_1 (1 - out O_1) * (-d_1 + out O_1) \big) + \big(w_7 * out O_2 (1 - out O_2) * (-d_2 + out O_2) \big) \Big)$$

$$w_{1new} = w_1 - \eta * \frac{\partial E tot}{\partial w_1}$$

$$w_{2new} = w_2 - \eta * \frac{\partial E_{tot}}{\partial w_2}$$

$$\frac{\partial E_{tot}}{\partial w_2} = \frac{\partial h_1}{\partial w_2} * \frac{\partial out h_1}{\partial h_1} * \left(\frac{\partial E_1}{\partial out h_1} + \frac{\partial E_2}{\partial out h_1} \right)$$

$$\frac{\partial h_1}{\partial w_2} = x_2$$

$$\frac{\partial out h_1}{\partial h_1} = out h_1 (1 - out h_1)$$

$$\frac{\partial E_1}{\partial out h_1} = w_5 * out o_1 (1 - out o_1) * (-d_1 + out o_1)$$

$$\frac{\partial E_2}{\partial out h_1} = w_7 * out o_2 (1 - out o_2) * (-d_2 + out o_2)$$

$$\frac{\partial E_{tot}}{\partial w_2} = x_2 * out h_1 (1 - out h_1) * \Big(\big(w_5 * out o_1 (1 - out o_1) * (-d + out o_1) \big) + \big(w_7 * out o_2 (1 - out o_2) * (-d_2 + out o_2) \big) \Big)$$

$$w_{2new} = w_1 - \eta * \frac{\partial E_{tot}}{\partial w_2}$$

$$w_{3new} = w_3 - \eta * \frac{\partial E_{tot}}{\partial w_3}$$

$$\frac{\partial E_{tot}}{\partial w_3} = \frac{\partial h_2}{\partial w_3} * \frac{\partial out h_2}{\partial h_2} * \frac{\partial E_{tot}}{\partial out h_2}$$

$$\frac{\partial E_{tot}}{\partial w_3} = \frac{\partial h_2}{\partial w_3} * \frac{\partial out h_2}{\partial h_2} * \left(\frac{\partial E_1}{\partial out h_2} + \frac{\partial E_2}{\partial out h_2}\right)$$

$$\frac{\partial E_1}{\partial out h_2} = \frac{\partial o_1}{\partial out h_2} * \frac{\partial out o_1}{\partial o_1} \times \frac{\partial E_1}{\partial out o_1}$$

$$\frac{\partial E_2}{\partial out h_2} = \frac{\partial o_2}{\partial out h_2} * \frac{\partial out o_2}{\partial o_2} \times \frac{E_2}{\partial out o_2}$$

$$\frac{\partial h_2}{\partial w_3} = \frac{2(x_1 * w_3 + x_2 * w_4 + b_1)}{\partial w_3}$$

$$= \underline{x_1}$$

$$\frac{\partial out h_2}{\partial h_2} = \frac{\partial\left(\frac{1}{1+e^{-h_2}}\right)}{\partial h_2} = \underline{out h_2(1 - out h_2)}$$

$$\frac{\partial o_1}{\partial out h_2} = \frac{\partial(out h_2 * w_6 + out h_1 * w_5 + b_2)}{\partial out h_2}$$

$$= \underline{w_6}$$

$$\frac{\partial out o_1}{\partial o_1} = \underline{out o_1(1 - out o_1)}$$

$$\frac{\partial E_1}{\partial out_{O1}} = -d_1 + out_{O1}$$

$$\frac{\partial o_2}{\partial out_{h2}} = \frac{\partial (out_{h2} * w_8 + out_{h1} * w_7 + b_2)}{\partial out_{h2}}$$

$$= w_8$$

$$\frac{\partial out_2}{\partial o_2} = \frac{\partial \left(\frac{1}{1+e^{-o_2}}\right)}{\partial o_2}$$

$$= out_{O2}(1 - out_{O2})$$

$$\frac{\partial E_2}{\partial out_{O2}} = \frac{\partial \left(\frac{1}{2}(d_2 - out_{O2})^2\right)}{\partial out_{O2}}$$

$$= 2 * \frac{1}{2}(d_2 - out_{O2})^{2-1} * -1$$

$$= -d_2 + out_{O2}$$

$$\frac{\partial E_{tot}}{\partial w_3} = \frac{\partial h_2}{\partial w_3} * \frac{\partial out_{h2}}{\partial h_2} * \left(\left(\frac{\partial o_1}{\partial out_{h2}} * \frac{\partial out_{O1}}{\partial o_1} * \frac{\partial E_1}{\partial out_{O1}}\right) + \left(\frac{\partial o_2}{\partial out_{h2}} * \frac{\partial out_{O2}}{\partial o_2} * \frac{\partial E_2}{\partial out_{O2}}\right)\right)$$

$$\frac{2E_{tot}}{2w_3} = X_1 * outh_2(1-outh_2) * (((w_6 * outO_1(1-outO_1) * (-d_1 + outO_1) + (w_8 * outO_2(1-outO_2) * (-d_2 + outO_2))$$

$$w_{3new} = w_3 - \eta * \frac{2E_{tot}}{2w_3}$$

$$w_{4new} = w_4 - \eta * \frac{2E_{tot}}{2w_4}$$

for w_4 the procedure is the same as previus w_1, w_2, w_3

Then

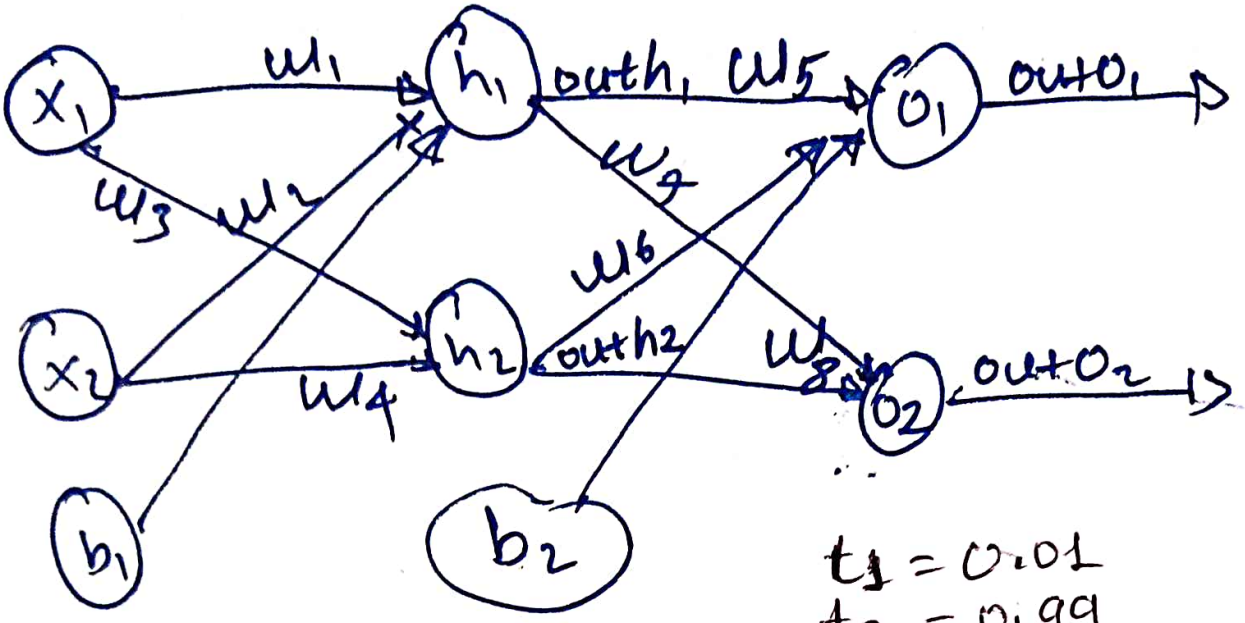
$$h_{1new} = w_{1new} * X_1 + w_{2new} * X_2 + b_1$$

$$h_{2new} = w_{3new} * X_1 + w_{4new} * X_2 + b_1$$

$$outh_{1new} = \frac{1}{1+e^{-h_{1new}}}$$

$$outh_{2new} = \frac{1}{1+e^{-h_{2new}}}$$

BP A Example



$t_1 = 0.01$
$t_2 = 0.99$

Given

$x_1 = 0.05$
$x_2 = 0.1$
$b_1 = 0.35$
$b_2 = 0.6$

$w_1 = 0.15$
$w_2 = 0.2$
$w_3 = 0.25$
$w_4 = 0.3$
$w_5 = 0.4$

$w_6 = 0.45$
$w_7 = 0.5$
$w_8 = 0.55$

η (learning rate) $= 0.5$

forward Pass

$h_1 = w_1 * x_1 + w_2 * x_2 + b_1$

$h_1 = 0.15 * 0.05 + 0.2 * 0.1 + 0.35$

$= 0.3775$

$$outh_1 = \frac{1}{1+e^{-h_1}} = \frac{1}{1+e^{-0.3775}}$$

$outh_1 = 0.596884378 = 593269992$

The Same procedur $outh_2 = 0.596884378$

$O_1 = w_5 * out h_1 + w_6 * out h_2 + b_2$ (2)

$O_1 = 1.105905967$

$out O_1 = \frac{1}{1+e^{-O_1}} = \frac{1}{1+e^{-1.105905967}} = 0.751365\,07$

⊛ The same process for out O_2

$out O_2 = 0.772928465$

$E_{total} = \sum_{i=1}^{n} \frac{1}{2} (target_i - output_i)^2$

$E_{O_1} = \frac{1}{2} (t_1 - out O_1)^2$

$= \frac{1}{2} (0.01 - 0.751365\,07)^2 = 0.274811083$

The same procedure for E_{O_2}

$E_{O_2} = \frac{1}{2} (t_2 - out O_2)^2 = 0.023560026$

$E_{tot} = E_{O_1} + E_{O_2} = 0.298371109$

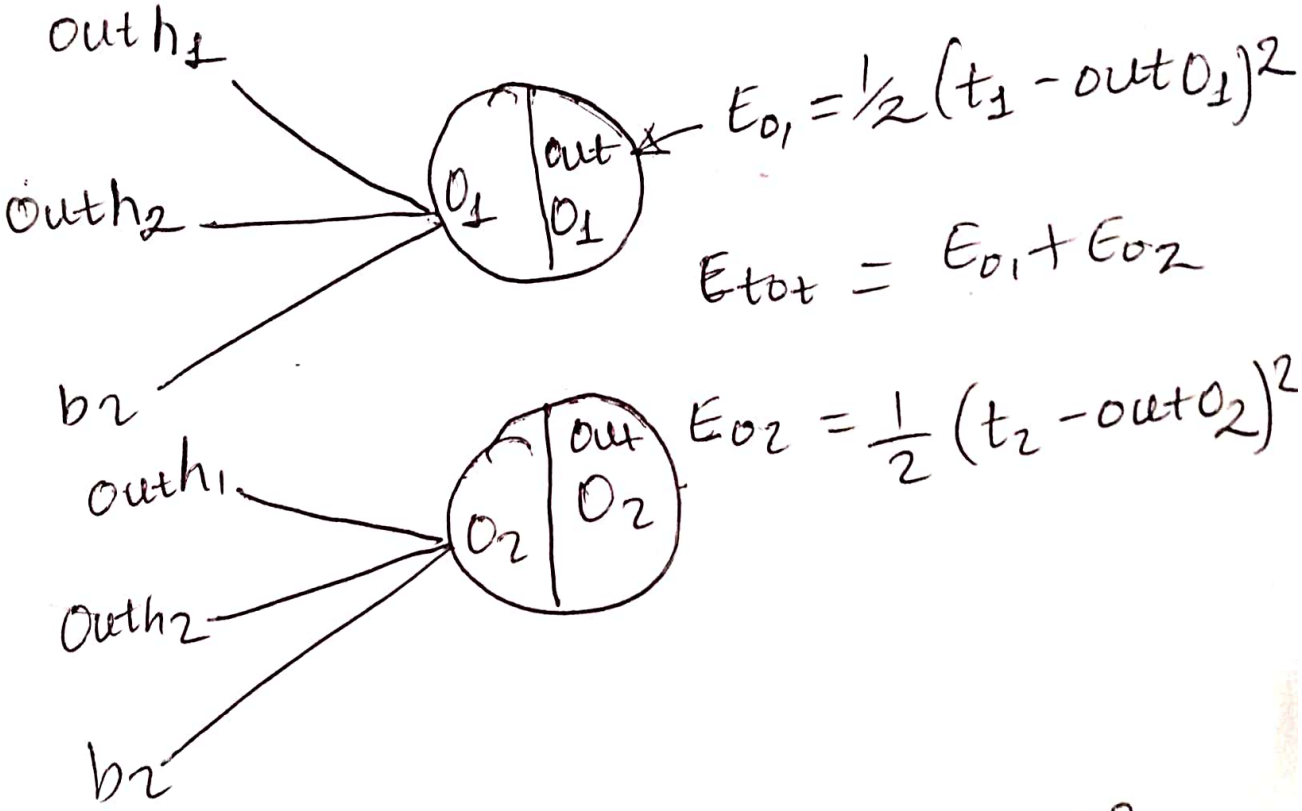
The error is huge, So, we have to update the weights, $w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8$

Hence we go to Backward propagation pass

6 First we start to update w_5

$$w_{5_{new}} = w_5 - \eta * \frac{\partial E_{tot}}{\partial w_5}$$

$$\frac{\partial E_{tot}}{\partial w_5} = \frac{\partial o_1}{\partial w_5} * \frac{\partial out_{o_1}}{\partial o_1} * \frac{\partial E_{tot}}{\partial out_1}$$



$$E_{tot} = \frac{1}{2}(t_1 - out_1)^2 + \frac{1}{2}(t_2 - out_{o_2})^2$$

$$\frac{\partial E_{tot}}{\partial out_1} = 2 \times \frac{1}{2}(t_1 - out_{o_1})^{2-1} * -1 + 0$$

$$= -(t_1 - out_1)$$

$$= -[0.01 - 0.75136507)$$

$$= 0.74136505$$

$$\frac{\partial out_{o_1}}{\partial o_1} = out_{o_1}(1 - out_{o_1})$$

$$= 0.751365 07(1 - 0.751365 07)$$

$$= 0.186315602$$

$$\frac{\partial o_1}{\partial w_5} = 1 * out_{h_1} * w_5^{1-1} + 0 + 0 = out_{h_1}$$

$$= 0.593269992$$

$$\frac{\partial E_{tot}}{\partial w_5} = 0.74136507 * 0.186315602 * 0.593269992$$

$$= 0.082167041$$

$$w_{5new} = w_5 - \eta * \frac{\partial E_{tot}}{\partial w_5}$$

$$= 0.4 - 0.5 * 0.082167041$$

$$= 0.35891648$$

using the same procedure as w_5

$$w_{6new} = w_6 - \eta \frac{\partial E_{tot}}{\partial w_6} = 0.408666186$$

$$w_{7new} = w_7 - \eta \frac{\partial E_{tot}}{\partial w_7} = 0.511301270$$

$$w_{8new} = w_8 - \eta \frac{\partial E_{tot}}{\partial w_8} = 0.561370121$$

Now go to adjust w_1, w_2, w_3, w_4,

5

$$w_{1new} = w_1 - \eta * \frac{\partial E_{tot}}{\partial w_1}$$

$$\frac{\partial E_{tot}}{\partial w_1} = \frac{\partial h_1}{\partial w_1} * \frac{\partial out\, h_1}{\partial h_1} * \frac{\partial E_{tot}}{\partial out\, h_1}$$

$$\frac{\partial E_{tot}}{\partial out\, h_1} = \frac{\partial E_1}{\partial out\, h_1} + \frac{\partial E_2}{\partial out\, h_1}$$

$$\frac{\partial E_{tot}}{\partial out\, h_1} = \frac{\partial o_1}{\partial out\, h_1} * \frac{\partial out\, o_1}{\partial o_1} * \frac{E_1}{\partial out\, O_1}$$

$$= 0.5 * out\, o_1(1 - out\, o_1) * (-d_1 + out\, O_1)$$

$$0.4 * 0.186815602 * 0.74136507$$

$$= 0.055399425$$

$$\frac{\partial E_2}{\partial out\, h_1} = \frac{\partial o_2}{\partial out\, h_1} * \frac{out\, o_2}{\partial o_2} * \frac{\partial E_2}{\partial out\, O_2}$$

$$= -0.019049119$$

$$\frac{\partial E_{tot}}{\partial out\, h_1} = \frac{\partial E_1}{\partial out\, h_1} + \frac{\partial E_2}{\partial out\, h_2}$$

$$\frac{\partial E_{tot}}{\partial out_{h_1}} = 0.055399425 + (-0.019049119)$$

$$= 0.036350306$$

6

$$\frac{\partial out_{h_1}}{\partial h_1} = out_{h_1}(1 - out_{h_1})$$

$$= 0.593269999(1 - 0.593269999)$$

$$= 0.241300709$$

$$\frac{\partial h_1}{\partial w_1} = x_1 = 0.05$$

$$\frac{\partial E_{tot}}{\partial w_1} = \frac{\partial h_1}{\partial w_1} * \frac{\partial out_{h_1}}{\partial h_1} * \frac{\partial E_{tot}}{\partial out_{h_1}}$$

$$= 0.05 * 0.241300709 * 0.036350306$$

$$= 0.000043856 8$$

$$w_{1new} = w_1 - \eta * \frac{\partial E_{tot}}{\partial w_1}$$

$$= 0.15 - 0.5 * 0.000043856 8$$

$$w_{1new} = 0.149780716$$

$$w_{2new} = w_2 - \eta * \frac{\partial E_{tot}}{\partial w_2} = 0.19956143$$

$$w_{3new} = w_3 - \eta * \frac{\partial E_{tot}}{\partial w_3} = 0.24975114$$

$$w_{4new} = w_4 - \eta * \frac{\partial E_{tot}}{\partial w_4} = 0.29950229$$