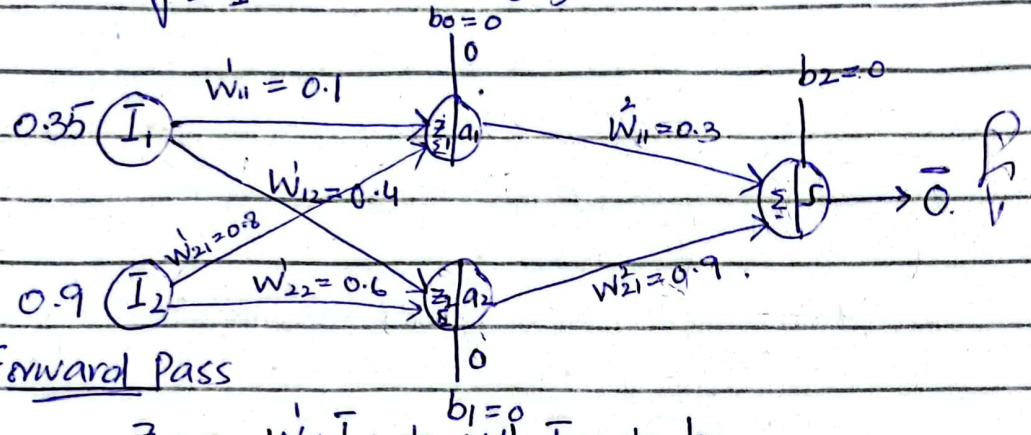
Forward pass & Backpropagation.

$$I = \begin{bmatrix} 0.35 \\ 0.9 \end{bmatrix} \quad \vec{W}^1 = \begin{bmatrix} w_{11}^1 & w_{12}^1 \\ w_{21}^1 & w_{22}^1 \\ 0.8 & 0.6 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.4 \\ 0.8 & 0.6 \end{bmatrix} \quad \vec{W}^2 = \begin{bmatrix} w_{11}^2 \\ w_{21}^2 \\ 0.9 \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0.9 \end{bmatrix}$$

$y = 1$ $\quad \alpha = 0.5$



Forward Pass

$$z_1 = w_{11}^1 I_1 + w_{21}^1 I_2 + b_0$$
$$= (0.1)(0.35) + (0.8)(0.9) + 0$$
$$\boxed{z_1 = 0.755}$$

$$a_1 = \sigma(z_1) = 1/(1+e^{-z_1}) = 1/1+0.47 = 1/1.47 = 0.68$$
$$\boxed{a_1 = 0.68}$$

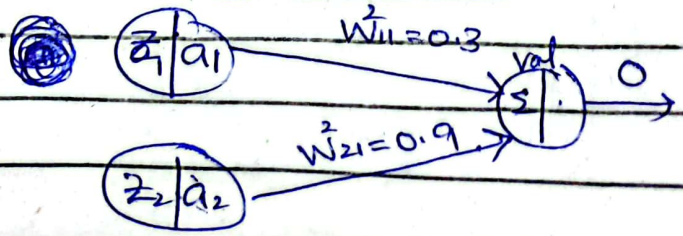
$$z_2 = w_{12}^1 I_1 + w_{22}^1 I_2 + b_1.$$
$$= (0.4)(0.35) + (0.6)(0.9) + 0.$$
$$= 0.68$$

$$a_2 = \sigma(0.68) = 1/1+e^{-0.68} = 1/1+0.5066 = 0.663.$$
$$\boxed{a_2 = 0.663}$$

$$val = b_2 + a_1 w_{11}^2 + a_2 w_{21}^2$$
$$val = 0 + 0.68(0.3) + 0.663(0.9) = 0.204 + 0.5967$$
$$0 = \sigma(0.8007) = 1/1+e^{-0.8007} = 1/1.449 = 0.6902$$
$$\boxed{0 = 0.6902} \qquad 0 = \sigma(val)$$

Back propagation :-

$$\text{Cost/Loss} = \frac{1}{2}(y-0)^2 = \frac{1}{2}(1-0.6902)^2 = 0.047988$$
$$\boxed{C = 0.04799}$$



Output layer :-

update W_{11}^2, W_{21}^2 :-

$$\frac{\partial e}{\partial w_{11}^2} = \frac{\partial e}{\partial o} * \frac{\partial o}{\partial val} * \frac{\partial val}{\partial w_{11}^2}$$

$o = \sigma(val) = \frac{1}{1+e^{-val}}$

$$\frac{\partial e}{\partial o} = \frac{1}{2} * 2 * (y-o)(o-1) = -(y-o)$$

$$\frac{\partial o}{\partial val} = \sigma(val) * (1 - \sigma(val)) = o * (1-o)$$

$$\frac{\partial val}{\partial w_{11}^2} = \frac{\partial (b_2 + a_1 w_{11}^2 + a_2 w_{21}^2)}{\partial w_{11}^2} = 0 + a_1 + 0 = a_1$$

$val = b_2 + a_1 w_{11}^2 + a_2 w_{21}^2$

$$\frac{\partial e}{\partial w_{11}^2} = 0 - (y-o) * o(1-o) * a_1$$

$$\frac{\partial e}{\partial w_{11}^2} = -(1-0.6902) * (0.6902)(1-0.6902) * 0.68$$

$$\boxed{\frac{\partial e}{\partial w_{11}^2} = -0.04505}$$

$$\frac{\partial e}{\partial w_{21}^2} = \frac{\partial e}{\partial o} * \frac{\partial o}{\partial val} * \frac{\partial val}{\partial w_{21}^2} = -(y-o) * o(1-o) * a_2$$

$$= -(1-0.6902) * 0.6902 (1-0.6902) * 0.663$$

$$\boxed{\frac{\partial e}{\partial w_{21}^2} = -0.04391}$$

New weights

$$w_{11}^2 = w_{11}^2 - 0.5\left(\frac{\partial e}{\partial w_{11}^2}\right) = 0.3 - 0.5(-0.04505) = 0.3225$$

$$\boxed{w_{11}^2 \text{new} = 0.3225}$$

$$w_{21}^2 = w_{21}^2 - \alpha\left(\frac{\partial e}{\partial w_{21}^2}\right) = 0.9 - 0.5(-0.04391) = 0.92195$$

$$\boxed{w_{21}^2 \text{new} = 0.92195}$$

hidden layers weights ($w_{11}, w_{12}, w_{21}, w_{22}$)

$$\frac{\partial e}{\partial w_{11}} = \frac{\partial e}{\partial o} * \frac{\partial o}{\partial val} * \frac{\partial val}{\partial a_1} * \frac{\partial a_1}{\partial z_1} * \frac{\partial z_1}{\partial w_{11}}$$

$$= -(y-o) * o(1-o) * w_{11}^2 * a_1(1-a_1) * I_1$$

$$= -(1-0.6902) * 0.6902 * (0.3) * 0.68(1-0.68) * 0.35$$

$$\boxed{\frac{\partial e}{\partial w_{11}} = -0.001535}$$

$$\frac{\partial C}{\partial w'_{12}} = \frac{\partial C}{\partial o} * \frac{\partial o}{\partial val} * \frac{\partial val}{a_2} * \frac{\partial a_2}{\partial z_2} * \frac{\partial z_2}{\partial w'_{12}}$$

$$= -(y-o) * o(1-o) * w^2_{21} * a_2(1-a_2) * I_1$$

$$= -(1-0.6902) * 0.6902(1-0.6902) * 0.9 * (0.663) * (1-0.663) * 0.35$$

$$= -0.004662$$

$$\frac{\partial C}{\partial w'_{21}} = \frac{\partial C}{\partial o} * \frac{\partial o}{\partial val} * \frac{\partial val}{\partial a_1} * \frac{\partial a_1}{\partial z_1} * \frac{\partial z_1}{\partial w'_{21}}$$

$$= -(y-o) * o(1-o) * w^2_{11} * a_1(1-a_1) * I_2 \quad *0.9$$

$$= -(1-0.6902) * 0.6902(1-0.6902) * 0.3(0.68)(1-0.68)$$

$$= -0.00389$$

$$\frac{\partial C}{\partial w'_{22}} = \frac{\partial C}{\partial o} * \frac{\partial o}{\partial val} * \frac{\partial val}{\partial a_2} * \frac{\partial a_2}{\partial z_2} * \frac{\partial z_2}{\partial w'_{22}}$$

$$= -(y-o) * o(1-o) * w^2_{21} * a_2(1-a_2) * I_2 \quad *0.9$$

$$= -(1-0.6902) * (0.6902)(1-0.6902) * 0.9 * (0.663)(1-0.663)$$

$$= -0.01199$$

Update weights (New $w'_{11}, w'_{12}, w'_{21}, w'_{22}$).

$$w'_{11new} = w'_{11old} - \alpha \frac{\partial C}{\partial w'_{11}} = 0.1 - 0.5(-0.00163) = 0.100765$$

$$w'_{12new} = w'_{12old} - \alpha \frac{\partial C}{\partial w'_{12}} = 0.4 - 0.5(-0.004662) = 0.4023$$

$$w'_{21new} = w'_{21old} - \alpha \frac{\partial C}{\partial w'_{21}} = 0.8 - 0.5(-0.00389) = 0.8015$$

$$w'_{22new} = w'_{22old} - \alpha \frac{\partial C}{\partial w'_{22}} = 0.6 - 0.5(-0.01199) = 0.605995$$

$$\vec{W'_1} = \begin{bmatrix} 0.100 & 0.402 \\ 0.8015 & 0.606 \end{bmatrix} \qquad \vec{W_2} = \begin{bmatrix} 0.3225 \\ 0.9215 \end{bmatrix}$$

$$\frac{\partial C}{\partial w_{11}^2} = \frac{\partial C}{\partial o} * \frac{\partial o}{\partial val} * \frac{\partial val}{\partial w_{11}^2}$$

$$\frac{\partial C}{\partial o} = \frac{d}{do}\left(\frac{1}{2}(y-\hat{y})^2\right) = \frac{1}{2}\frac{d}{do}(y-\hat{y})^2$$

$$= 2 \times \frac{1}{2}(y-o)\frac{d}{do}(y-o) = (y-o)(0-1) = -(y-o).$$

$$\boxed{\frac{\partial C}{\partial o} = -(y-o)}$$

$$\frac{\partial o}{\partial val} = \frac{d\,\sigma(val)}{dval} = \frac{d}{dval}\left(\frac{1}{1+e^{-val}}\right) = \frac{d}{dval}(1+e^{-val})^{-1}$$

$$= -1(1+e^{-val})^{-2}\frac{d}{dval}(1+e^{-val}).$$

$\frac{d}{dx} e^{x} = e^{x}\left(\frac{1}{x}x\right)$

$$= -1(1+e^{-val})^{-2}(0 + \frac{d}{dval}(e^{-val}))$$

$$= -(1+e^{-val})^{-2}(e^{-val})\frac{d}{dval}(-val).$$

$$= -(1+e^{-val})^{-2}(e^{-val})(-1)$$

$$= (1+e^{-val})^{-2}e^{-val} = \frac{e^{-val}}{(1+e^{-val})^2}$$

$$= \frac{1 \cdot e^{-val}}{(1+e^{-val})(1+e^{-val})} = \frac{1}{1+e^{-val}} \cdot \frac{e^{-val}+1-1}{1+e^{-val}}$$

$$= \frac{1}{1+e^{-val}} * \left(\frac{1+e^{-val}}{1+e^{-val}} - \frac{1}{1+e^{-val}}\right)$$

$$= \frac{1}{1+e^{-val}} * \left(1 - \frac{1}{1+e^{-val}}\right)$$

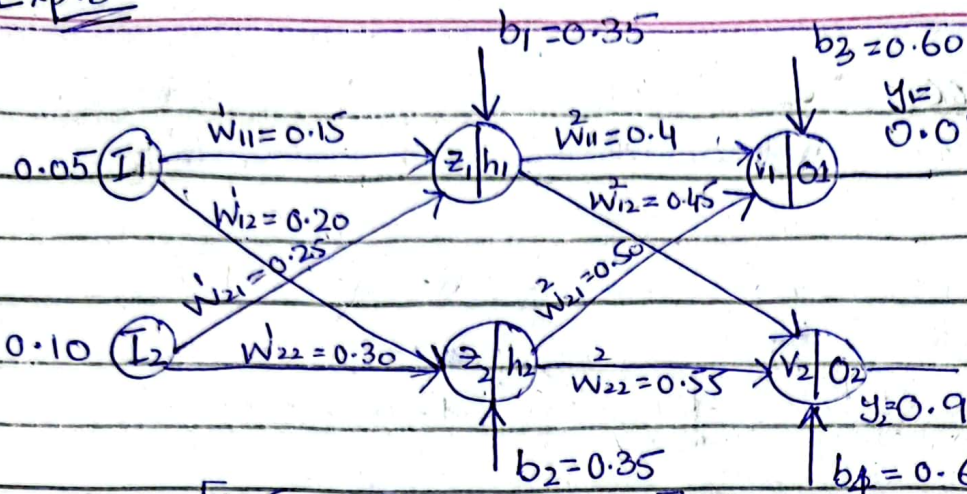
$$= \sigma(val)(1-\sigma(val))$$

$$= o(1-o)$$

$$\frac{\partial val}{\partial w_{11}^2} = \frac{d}{dw_{11}^2}(b_2 + a_1 w_{11}^2 + a_2 w_{21}^2) = \frac{d}{dw_{11}^2}(b_2) + \frac{d}{dw_{11}^2}(a_1 w_{11}^2) + \frac{d}{dw_{11}^2}(a_2 w_{21}^2)$$

$$= 0 + a_1 + 0 = a_1$$

Exp:2



$$W_1 = \begin{bmatrix} 0.15 & 0.20 \\ 0.25 & 0.30 \end{bmatrix} \quad W_2 = \begin{bmatrix} 0.4 & 0.45 \\ 0.50 & 0.55 \end{bmatrix} \quad I = \begin{bmatrix} 0.05 \\ 0.10 \end{bmatrix}$$

Forward pass :-

$z_1 = b_1 + w^1_{11} I_1 + w^1_{21} I_2 = 0.35 + (0.15)(0.05) + (0.25)(0.10)$

$z_1 = 0.3825$ $\quad h_1 = \sigma(z_1) = 1/1.6822 = 0.5945$

$h_1 = 0.5945$

$z_2 = b_2 + w^1_{12} I_1 + w^1_{22} I_2 = 0.35 + (0.2)(0.05) + (0.3)(0.1)$

$z_2 = 0.39$ $\quad h_2 = \sigma(z_2) = 1/1.6771 = 0.5963$

$h_2 = 0.5963$

$v_1 = b_3 + w^2_{11} h_1 + w^2_{21} h_2 = 0.60 + (0.4)(0.5945) + (0.50)(0.5963)$

$v_1 = 1.1359$ $\quad o_1 = \sigma(v_1) = 1/1.3211 = 0.7569$

$o_1 = 0.7569$

$v_2 = b_4 + w^2_{12} h_1 + w^2_{22} h_2 = 0.60 + (0.45)(0.5945) + (0.55)(0.5963)$

$v_2 = 1.1955$ $\quad o_2 = \sigma(v_2) = 1/1.3026 = 0.7677$

$o_2 = 0.7677$

Cost / Loss :-

$C = 1/2\{(y_1 - o_1)^2 + (y_2 - o_2)^2\} = 1/2\{(0.01 - 0.7569)^2 + (0.99 - 0.7677)^2\}$

$C = 1/2\{0.5579 + 0.0494\} = 0.30365$ $\quad C = 0.30365$

Update $w^2_{11}, w^2_{12}, w^2_{21}, w^2_{22}$ for output layer:

$$\frac{\partial C}{\partial w^2_{11}} = \frac{\partial C}{\partial o_1} * \frac{\partial o_1}{\partial v_1} * \frac{\partial v_1}{\partial w^2_{11}} = -(y_1 - o_1) * o_1(1 - o_1) * h_1$$

$$= -(0.01 - 0.7569) * (0.7569)(1 - 0.7569) * 0.5945$$

$$\frac{\partial C}{\partial w^2_{11}} = +0.0817$$

$\frac{\partial C}{\partial w_{21}^2} = \frac{\partial C}{\partial o_1} * \frac{\partial o_1}{\partial v_1} * \frac{\partial v_1}{\partial w_{21}^2} = -(y_1 - o_1) * o_1(1-o_1) * h_2$

$= -(0.01 - 0.7569) * (0.7569)(1-0.7569) * 0.5963 = 0.08195$

$\frac{\partial C}{\partial w_{12}^2} = \frac{\partial C}{\partial o_2} * \frac{\partial o_2}{\partial v_2} * \frac{\partial v_2}{\partial w_{12}^2} = -(y_2 - o_2) * o_2(1-o_2) * h_1$

$= -(0.99 - 0.7677) * 0.7677(1-0.7677) * 0.5945 = -0.02357$

$\frac{\partial C}{\partial w_{22}^2} = \frac{\partial C}{\partial o_2} * \frac{\partial o_2}{\partial v_2} * \frac{\partial v_2}{\partial w_{22}^2} = -(y_2 - o_2) * o_2(1-o_2) * h_2$

$= -(0.99 - 0.7677) * (0.7677)(1-0.7677) * 0.5945 = -0.02364$

New Weights :-

$w_{11}^{12} = w_{11}^2 - \alpha \partial C / \partial w_{11}^2 = 0.4 - 0.5(+0.0817) = 0.3592$

$w_{12}^{12} = w_{12}^2 - \alpha \partial C / \partial w_{12}^2 = 0.45 - 0.5(-0.02357) = 0.4618$

$w_{21}^{12} = w_{21}^2 - \alpha \partial C / \partial w_{21}^2 = 0.50 - 0.5(0.08195) = 0.459$

$w_{22}^{12} = w_{22}^2 - \alpha \partial C / \partial w_{22}^2 = 0.55 - 0.5(-0.02364) = 0.5618$

Update w_{11}^1, w_{12}^1, w_{21}^1, w_{22}^1 :

$\frac{\partial C}{\partial w_{11}^1} = \frac{\partial C_1}{\partial w_{11}^1} + \frac{\partial C_2}{\partial w_{11}^1} = \left(\frac{\partial C_1}{\partial o_1} * \frac{\partial o_1}{\partial v_1} * \frac{\partial v_1}{\partial h_1} * \frac{\partial h_1}{\partial z_1} * \frac{\partial z_1}{\partial w_{11}^1}\right) +$

$\left(\frac{\partial C_2}{\partial o_2} * \frac{\partial o_2}{\partial v_2} * \frac{\partial v_2}{\partial h_1} * \frac{\partial h_1}{\partial z_1} * \frac{\partial z_1}{\partial w_{11}^1}\right)$ taking common

$= \left\{\frac{\partial C_1}{\partial o_1} * \frac{\partial o_1}{\partial v_1} * \frac{\partial v_1}{\partial h_1} + \frac{\partial C_2}{\partial o_2} * \frac{\partial o_2}{\partial v_2} * \frac{\partial v_2}{\partial h_1}\right\} \left(\frac{\partial h_1}{\partial z_1} * \frac{\partial z_1}{\partial w_{11}^1}\right)$

$= \{(-(y_1 - o_1) * o_1(1-o_1) * w_{11}^2) + (-(y_2 - o_2) * o_2(1-o_2) * w_{12}^2)\}$

$h_1(1-h_1) * I_1$

$= \{[-(0.01 - 0.7569) * 0.7569(1-0.7569) * 0.4]$

$+ [-(0.99 - 0.7677) * 0.7677(1-0.7677) * 0.45]\} * 0.5945 *$

$(1-0.5945) * 0.05$

$\frac{\partial C}{\partial w_{11}^1} = \{(0.05497) + (-0.0178)\} * 0.01205 = 0.000448$

$w_{11}^1 = w_{11}^{1\,old} - \alpha \frac{\partial C}{\partial w_{11}^1} = 0.15 - 0.5(0.000448) = 0.1498$

$\frac{\partial C}{\partial w_{12}^1} = \frac{\partial C_1}{\partial w_{12}^1} + \frac{\partial C_2}{\partial w_{12}^1} = \frac{\partial C_1}{\partial o_1} * \frac{\partial o_1}{\partial v_1} * \frac{\partial v_1}{\partial h_2} * \frac{\partial h_2}{\partial z_2} * \frac{\partial z_2}{\partial w_{12}^1} +$

$\frac{\partial C_2}{\partial o_2} * \frac{\partial o_2}{\partial v_2} * \frac{\partial v_2}{\partial h_2} * \frac{\partial h_2}{\partial z_2} * \frac{\partial z_2}{\partial w_{12}^1} = \{(y_1 - o_1) * o_1(1-o_1) * w_{21}^2$

$+ -(y_2 - o_2) * o_2(1-o_2) * w_{22}^2\} * h_2(1-h_2) * I_1$

$$\frac{\partial C}{\partial w'_{12}} = \{-(0.01-0.7569)*0.7569(1-0.7569)*0.50 + -(0.99-0.7677)*0.7677(1-0.7677)*0.55\}* 0.5963(1-0.5963)*0.05$$

$$= (0.0687+(-0.0218))*0.01204 = 0.000565$$

$$w'_{12} = w^1_{12} - \alpha\, \partial C/\partial w'_{12} = 0.20 - 0.5(0.000567)$$

$$\boxed{w'_{12} = 0.1997}$$

$$\frac{\partial C}{\partial w'_{21}} = \frac{\partial C_1}{\partial o_1}*\frac{\partial o_1}{\partial v_1}*\frac{\partial v_1}{\partial h_1}*\frac{\partial h_1}{\partial z_1}*\frac{\partial z_1}{\partial w'_{21}} + \frac{\partial C_2}{\partial o_2}*\frac{\partial o_2}{\partial v_2}*\frac{\partial v_2}{\partial h_1}*\frac{\partial h_1}{\partial z_1}*\frac{\partial z_1}{\partial w'_{21}}$$

$$= \left\{\frac{\partial C_1}{\partial o_1}*\frac{\partial o_1}{\partial v_1}*\frac{\partial v_1}{\partial h_1} + \frac{\partial C_2}{\partial o_2}*\frac{\partial o_2}{\partial v_2}*\frac{\partial v_2}{\partial h_1}\right\}*\frac{\partial h_1}{\partial z_1}*\frac{\partial z_1}{\partial w'_{21}}$$

$$= \{-(y_1-o_1)*o_1(1-o_1)*w^2_{11} + (-(y_2-o_2)*o_2(1-o_2)*w^2_{12}\}$$

$$*h_1(1-h_1)*I_2$$

$$= (-(0.01-0.7569)*0.7569(1-0.7569)*0.4) + (-(0.99-0.7677)*0.7677(1-0.7677)*0.45)*$$

$$0.5945*(1-0.5945)*0.10$$

$$= (0.05497+(-0.01784))*0.0241 = 0.0008948$$

$$w'_{21} = w'_{21} - \alpha\frac{\partial C}{\partial w'_{21}} = 0.25 - 0.5(0.0008948) = 0.2496$$

$$\boxed{w'_{21} = 0.2496}$$

$$\frac{\partial C}{\partial w'_{22}} = \left\{\frac{\partial C_1}{\partial o_1}*\frac{\partial o_1}{\partial v_1}*\frac{\partial v_1}{\partial h_2} + \frac{\partial C_2}{\partial o_2}*\frac{\partial o_2}{\partial v_2}*\frac{\partial v_2}{\partial h_2}\right\}*\frac{\partial h_2}{\partial z_2}*\frac{\partial z_2}{\partial w'_{22}}$$

$$= \{-(y_1-o_1)o_1(1-o_1)w^2_{21} + (-(y_2-o_2)o_2(1-o_2)w^2_{22}\}*$$

$$h_2(1-h_2)I_2$$

$$= (-(0.01-0.7569)0.7569(1-0.7569)*0.5 + -(0.99-0.7677)0.7677(1-0.7677)0.55)*$$

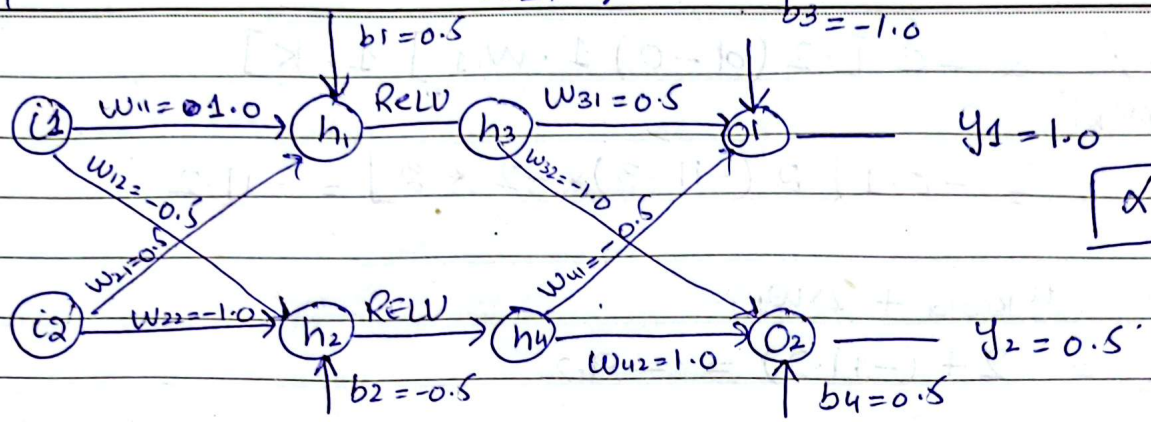
$$0.5963(1-0.5963)0.10$$

$$= (0.0687+(-0.0218))*0.0249 = 0.00113.$$

$$w'_{22} = w'_{22} - \alpha\frac{\partial C}{\partial w'_{22}} = 0.3 - 0.5(0.00113) = 0.2994$$

$$\boxed{w'_{22} = 0.2994}$$

Update w_{31} and w_{21} ?



$\alpha = 0.1$

Forward pass :-

$h_1 = w_{11} i_1 + w_{21} i_2 + b_1 = 2$

$h_3 = ReLU(h_1) = 2.$ $\quad (h_3 = h_1)$

$h_2 = w_{12} i_1 + w_{22} i_2 + b_2 = -0.5$

$h_4 = REW(h_2) = 0$

$O_1 = w_{31} h_3 + w_{41} h_4 + b_3 = 0$

$O_2 = w_{32} h_3 + w_{42} h_4 + b_4 = -1.5$

Find Loss / cost :-

$$C = \frac{1}{2}\left[(1-0)^2 + (0.5-(-1.5))^2\right] = 2.5$$

update w_{31} :-

$$\frac{\partial C}{\partial w_{31}} = \frac{\partial C}{\partial O_1} * \frac{\partial O_1}{\partial w_{31}}$$

$$\boxed{\frac{\partial C}{\partial O_1} = -(y_1 - O_1)} \qquad \boxed{\frac{\partial O_1}{\partial w_{31}} = h_3}$$

$$\frac{\partial C}{\partial w_{31}} = -(y_1 - O_1) * h_3 = -2$$

New weight of w_{31} :-

$$w_{31_{new}} = w_{31_{old}} - \alpha \frac{\partial C}{\partial w_{31}} = 0.5 - (0.1)(-2) = 0.7$$

Now update W_{21} :-

$$\frac{\partial C}{\partial w_{21}} = \frac{\partial C_1}{\partial w_{21}} + \frac{\partial C_2}{\partial w_{21}}$$

$$= \frac{\partial C_1}{\partial o_1} * \frac{\partial o_1}{\partial h_3} * \frac{\partial h_3}{\partial h_1} * \frac{\partial h_1}{\partial w_{21}} + \frac{\partial C_2}{\partial o_2} * \frac{\partial o_2}{\partial h_3} * \frac{\partial h_3}{\partial h_1} * \frac{\partial h_1}{\partial w_{21}}$$

$\frac{\partial C_1}{\partial o_1} = -(y_1 - o_1)$ $\quad$ $\frac{\partial C_2}{\partial o_2} = -(y_2 - o_2)$.

$\frac{\partial o_1}{\partial h_3} = w_{31}$ $\quad$ $\frac{\partial o_2}{\partial h_3} = w_{32}$.

$\frac{\partial h_3}{\partial h_1} = \frac{\partial}{\partial h_1}(h_1) = 1$. $\quad$ $\frac{\partial h_1}{\partial w_{21}} = i_2$.

$$\frac{\partial C}{\partial w_{21}} = \left(\frac{\partial C_1}{\partial o_1} * \frac{\partial o_1}{\partial h_3} + \frac{\partial C_2}{\partial o_2} * \frac{\partial o_2}{\partial h_3} \right) \frac{\partial h_3}{\partial h_1} * \frac{\partial h_1}{\partial w_{21}}.$$

$$= \left(-(y_1 - o_1) * w_{31} - (y_2 - o_2) * w_{32} \right) (1 * i_2)$$

$= -1.5$ after solving.

New weight of w_{21} :-

$$w_{21_{new}} = w_{21_{old}} - \alpha \frac{\partial C}{\partial w_{21}}$$

$$= 0.5 - (0.1)(-1.5) = 0.65$$