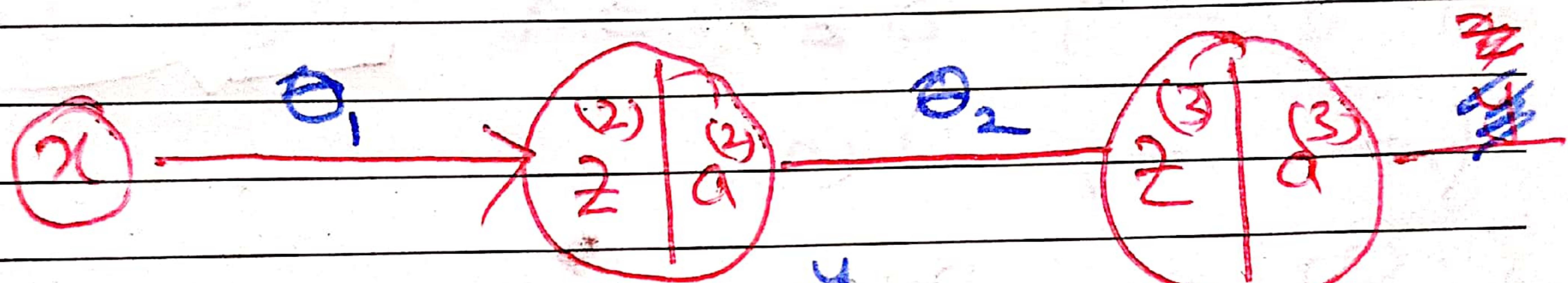
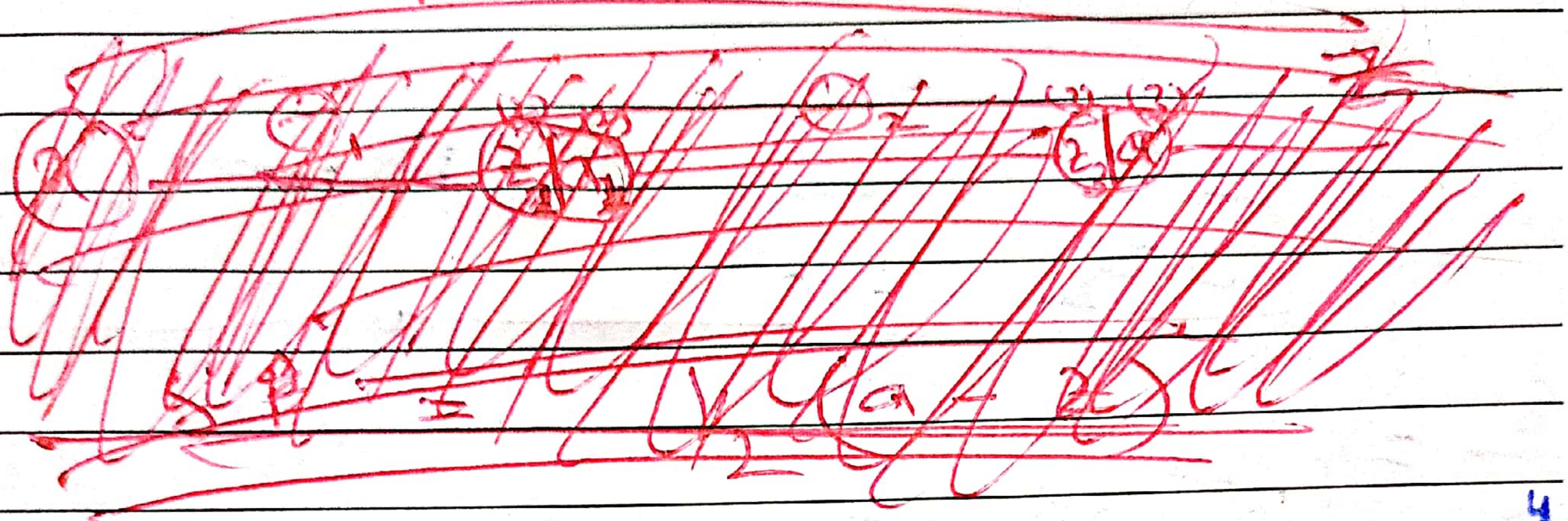


Date _____

To understand Simple Neural Net



$$\delta = \gamma_2 (y - q_2)^2$$

$$\frac{\partial \delta}{\partial \theta_2} = \frac{\partial \delta}{\partial a^{(3)}} \cdot \frac{\partial a^{(3)}}{\partial \theta_2}$$

it can further be elaborated

$$\frac{\partial \delta}{\partial \theta_2} = \frac{\partial \delta}{\partial a^{(3)}} \cdot \frac{\partial a^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial \theta_2} = \text{---}$$

$$\frac{\partial \delta}{\partial \theta_1} = \frac{\partial \delta}{\partial a^{(2)}} \cdot \frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial a^{(1)}} \cdot \frac{\partial a^{(1)}}{\partial \theta_1} = \text{---}$$

(ii)

Date _____

lets re-write eq. (i) & (ii)

$$\frac{\partial \delta}{\partial \theta_2} = \frac{\partial z^{(3)}}{\partial \theta_2} \cdot \frac{\partial g^{(3)}}{\partial z^{(3)}} \cdot \cancel{\frac{\partial \theta_2}{\partial \theta_2}} + \frac{\partial \delta}{\partial \theta_2}$$

$$\frac{\partial \delta}{\partial \theta_1} = \frac{\partial z^{(2)}}{\partial \theta_1} \cdot \frac{\partial g^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(3)}}{\partial \theta_1} \cdot \frac{\partial g^{(3)}}{\partial z^{(3)}} \cdot \cancel{\frac{\partial \theta_1}{\partial \theta_1}}$$

from ex i(a)

$$\frac{\partial z^{(3)}}{\partial \theta_2} = \gamma(a, \theta_2) = g^{(2)}$$

$$\frac{\partial g^{(3)}}{\partial z^{(3)}} = \frac{\gamma}{\partial z^{(3)}} \left(\frac{-1}{1 + e^{-z^{(3)}}} \right)$$

$$= \frac{\gamma}{\partial z^{(3)}} (1 + e^{-z^{(3)}})^{-1}$$

$$= -1 (1 + e^{-z^{(3)}})^{-2} \cdot \frac{\partial \gamma}{\partial z^{(3)}} \cdot 2e \cdot (-)$$

$$= (1 + e^{-z^{(3)}})^{-2} \cdot e^{-z^{(3)}}$$

$$= \frac{1}{(1 + e^{-z^{(3)}})^2} \cdot e^{-z^{(3)}}$$

Date _____

$$\begin{aligned} &= \frac{1}{1+e^{-x}} \cdot \frac{1+e^{-x}-1}{1+e^{-x}} \\ &= \frac{1}{1+e^{-x}} \cdot \left[\frac{1+e^{-x} - 1}{1+e^{-x}} - \frac{1}{1+e^{-x}} \right] \\ &= \delta^{(3)} \cdot (1 - \delta^{(2)}) \end{aligned}$$

$$\begin{aligned} \frac{\partial E}{\partial a^{(2)}} &= \frac{\partial}{\partial a^{(2)}} \left[\frac{1}{2} (\delta^{(3)} \cdot \delta^{(2)})^2 \right] \\ &= 2 \times \frac{1}{2} (\delta^{(2)} \cdot \delta^{(3)}) \cdot 1 \\ &= \delta^{(2)} \cdot \delta^{(3)} \end{aligned}$$

or (i)-a become

$$\frac{\partial E}{\partial \theta_2} = \delta^{(2)} \cdot \delta^{(3)} \cdot (1 - \delta^{(2)}) \cdot \delta^{(2)}$$

Similarly or ii-a being



Narsun Studios

Date _____

$$\frac{\partial \tau}{\partial \theta_1} = \alpha \cancel{x} n$$

$$\frac{\partial \alpha^{(2)}}{\partial \tau^{(2)}} = \alpha^{(2)} (1 - \alpha^{(2)})$$

$$\frac{\partial \tau}{\partial \theta_2} = \Theta_2$$

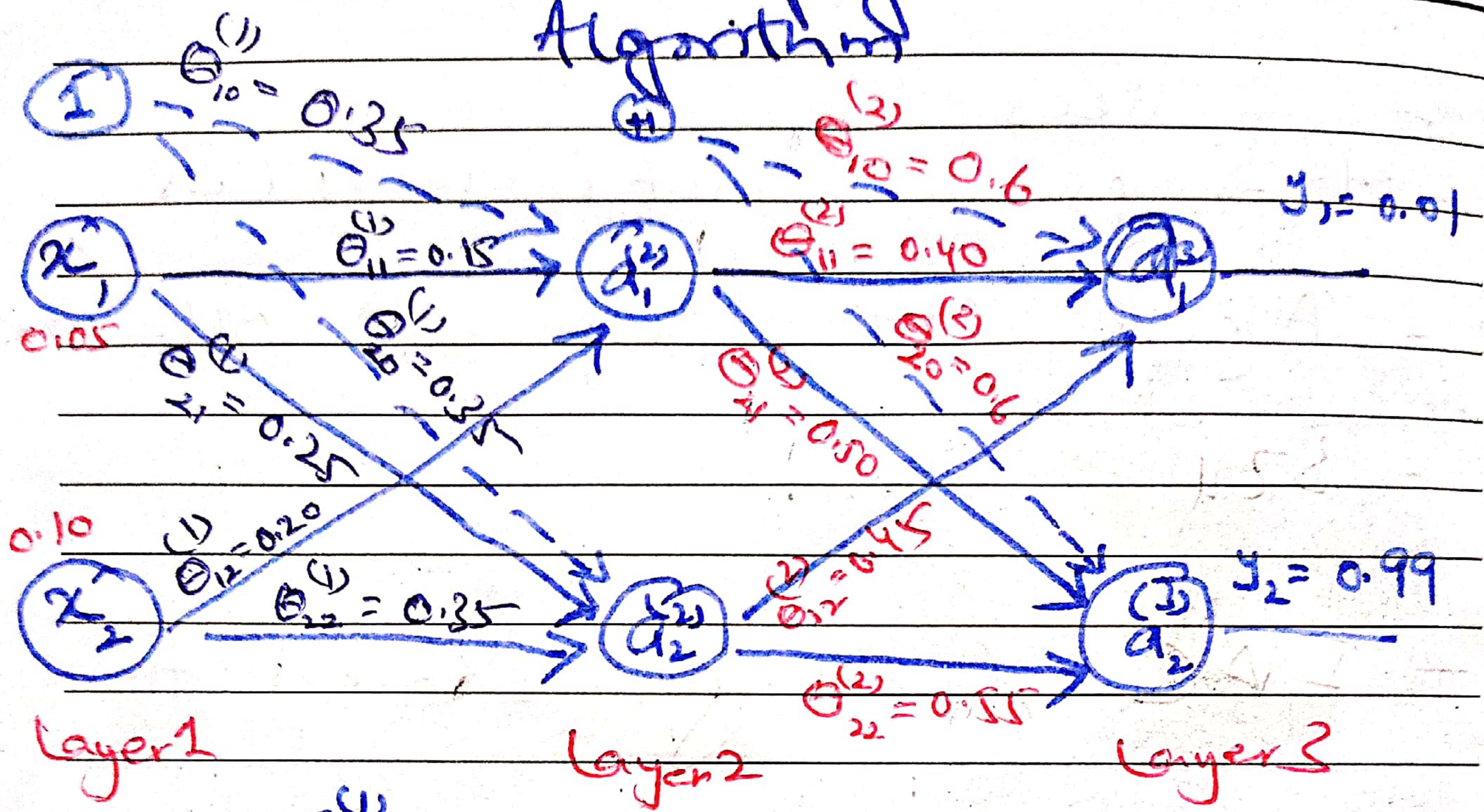
eq ii - q bounds

$$\frac{\partial \delta}{\partial \theta_1} = \cancel{x} * \alpha^{(2)} (1 - \alpha^{(2)}) \cdot \Theta_2$$

$$\delta^{(2)} (1 - \delta^{(2)}) \quad \delta^{(2)}$$

Date _____

Back Propagation Algorithm



Layer 1

Layer 2

Layer 3

$$\begin{aligned}
 z_1^{(2)} &= \theta_{10}^{(1)} + \theta_{11}^{(1)} x_1 + \theta_{12}^{(1)} x_2 \\
 &= 0.35 + 0.15 \times 0.05 + 0.20 \times 0.10 \\
 &= 0.377
 \end{aligned}$$

$$a_1^{(2)} = \frac{1}{1 + e^{-z_1^{(2)}}} = \frac{1}{1 + e^{-0.377}} = 0.59$$

Similarly

$$z_2^{(2)} = 0.35 + \theta_{21}^{(1)} x_1 + \theta_{22}^{(1)} x_2$$

$$a_2^{(2)} = \frac{1}{1 + e^{-z_2^{(2)}}} = 0.5968$$

Now for Layer 3,

$$\begin{aligned}
 z_1^{(3)} &= \theta_{10}^{(2)} + \theta_{11}^{(2)} a_1^{(2)} + \theta_{12}^{(2)} a_2^{(2)} \\
 &= 0.6 + 0.40 \times 0.593 + 0.45 \times 0.5968 \\
 &= 1.105
 \end{aligned}$$

Date _____

$$a_1^{(3)} = \frac{1}{1 + e^{-z_1^{(3)}}} = \frac{1}{1 + e^{-1.105}} = 0.7513$$

Similarly,

$$z_2 = 0.6 + \theta_{21} a_1^{(2)} + \theta_{22} a_2^{(2)}$$

$$a_2^{(3)} = \frac{1}{1 + e^{-z_2^{(3)}}} = 0.7729$$

Now,

$$\begin{aligned}\delta_1^{(3)} &= \gamma_1 (y_1 - a_1^{(3)})^2 \\ &= \gamma_1 (0.01 - 0.7513)^2 = 0.2747\end{aligned}$$

$$\delta_2^{(3)} = \gamma_2 (y_2 - a_2^{(3)})^2 = 0.0235$$

$$\delta_T = \delta_1^{(3)} + \delta_2^{(3)} = 0.2985$$

Goal is to make $\delta_T \approx 0$ (i)

$$\frac{\partial \delta_T}{\partial \theta_{11}^{(2)}} = \frac{\partial \delta_T}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(3)}} \times \frac{\partial z_1^{(3)}}{\partial \theta_{11}^{(2)}}$$

from eqn. (i), $\frac{\partial \delta_T}{\partial a_1^{(2)}} = \gamma_1 \left[\frac{\partial}{\partial a_1^{(2)}} \gamma_1 (y_1 - a_1^{(3)})^2 \right] = \gamma_1 (y_1 - a_1^{(3)})$

$$\frac{\partial \delta_T}{\partial a_1^{(2)}} = 0.25 \times 0.01 = 0.0025$$

Date _____

from eq. (i)

$$\begin{aligned}
 \frac{\partial S_T}{\partial a_1^{(3)}} &= \frac{1}{\partial a_1^{(3)}} \left[\frac{1}{2} (y_1 - a_1^{(3)})^2 + \frac{1}{2} (y_2 - a_2^{(3)})^2 \right] \\
 &= \frac{1}{2} \times 2 (y_1 - a_1^{(3)}) \cdot (-1) \\
 &= a_1^{(3)} - y_1 \\
 &= 0.7513 - 0.01 = 0.7412
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial a_1^{(3)}}{\partial z_1^{(3)}} &= \frac{\partial}{\partial z_1^{(3)}} \left(\frac{1}{1 + e^{-z_1^{(3)}}} \right) \\
 &= \frac{\partial}{\partial z_1^{(3)}} (1 + e^{-z_1^{(3)}})^{-1} \\
 &= -1 (1 + e^{-z_1^{(3)}})^{-2} \cdot e^{-z_1^{(3)}} \cdot (-1) \\
 &= -(1 + e^{-z_1^{(3)}})^{-2} \cdot e^{-z_1^{(3)}} \\
 &= \frac{1}{(1 + e^{-z_1^{(3)}})^2} \cdot e^{-z_1^{(3)}} \\
 &= \frac{1}{1 + e^{-z_1^{(3)}}} \cdot \frac{e^{-z_1^{(3)}}}{1 + e^{-z_1^{(3)}}}
 \end{aligned}$$

add & subtract 1

$$\begin{aligned}
 &= \frac{1 - e^{-z_1^{(3)}}}{1 + e^{-z_1^{(3)}}} \cdot \frac{1 + e^{-z_1^{(3)}} - 1}{1 + e^{-z_1^{(3)}}}
 \end{aligned}$$

Date _____

$$\begin{aligned}
 &= \frac{1}{1 + e^{-z_1^{(3)}}} \cdot \left(\frac{1 + e^{-z_1^{(3)}}}{1 + e^{-z_1^{(3)}}} - \frac{1}{1 + e^{-z_1^{(3)}}} \right) \\
 &= a_1^{(3)} (1 - a_1^{(3)}) \\
 &= 0.7513 (1 - 0.7513) \\
 &= 0.1868
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial \delta_T^{(3)}}{\partial \theta_{11}^{(2)}} &= \frac{\partial}{\partial \theta_{11}^{(2)}} (0.6 + \theta_{11}^{(2)} a_1^{(2)} + \theta_{12}^{(2)} a_2^{(2)}) \\
 &= a_1^{(2)} = 0.593
 \end{aligned}$$

Putting in eq (iv)

$$\begin{aligned}
 \frac{\partial \delta_T}{\partial \theta_{11}^{(2)}} &= 0.7413 * 0.1868 * 0.593 \\
 &= 0.0821
 \end{aligned}$$

Updating parameters $\theta_{ij}^{(2)}$ let $\alpha = 0.6$

$$\theta_{11}^{(2)} := \theta_{11}^{(2)} - \alpha \frac{\partial \delta_T}{\partial \theta_{11}^{(2)}}$$

$$= 0.4 - 0.6 \times 0.0821$$

$$= 0.3506$$

Similarly do for other parameters i.e.,

$$\theta_{12}^{(2)} := \theta_{12}^{(2)} - \alpha \frac{\partial \delta_T}{\partial \theta_{12}^{(2)}}$$

$$\theta_{21}^{(2)} := \theta_{21}^{(2)} - \alpha \frac{\partial \delta_T}{\partial \theta_{21}^{(2)}}$$

$$\theta_{22}^{(2)} := \theta_{22}^{(2)} - \alpha \frac{\partial \delta_T}{\partial \theta_{22}^{(2)}}$$



Narsun Studios

Date _____

Now for hidden layer i.e., layer 2

$$\frac{\partial \delta_j}{\partial \theta_{ii}} = \frac{\partial \delta_j}{\partial a_i^{(2)}} \times \frac{\partial a_i^{(2)}}{\partial z_i^{(2)}} \times \frac{\partial z_i^{(2)}}{\partial \theta_{ii}} \quad (\text{ii})$$

$$\frac{\partial \delta_j}{\partial a_i^{(2)}} = \frac{\partial \delta_j}{\partial a_i^{(2)}} + \frac{\partial \delta_{j+1}}{\partial a_i^{(2)}} \quad (\text{iii})$$

let's simplify eq. (iii)

$$\frac{\partial \delta_{j+1}}{\partial a_i^{(2)}} = \frac{\partial \delta_{j+1}}{\partial z_i^{(3)}} \times \frac{\partial z_i^{(3)}}{\partial a_i^{(2)}} \quad (\text{iv})$$

Similarly

$$\frac{\partial \delta_2}{\partial a_i^{(2)}} = \frac{\partial \delta_2}{\partial z_i^{(3)}} \times \frac{\partial z_i^{(3)}}{\partial a_i^{(2)}} \quad (\text{v})$$

eq. (iv) can further be elaborated

$$\frac{\partial \delta_j}{\partial z_i^{(3)}} = \frac{\partial \delta_j}{\partial a_i^{(2)}} \times \frac{\partial a_i^{(2)}}{\partial z_i^{(3)}}$$

$$= \frac{\partial}{\partial a_i^{(2)}} \left[\frac{1}{2} (y_i - a_i^{(3)})^2 \right] \times \frac{\partial}{\partial z_i^{(3)}} \cdot \frac{1}{1 + e^{-z_i^{(3)}}}$$

$$= (y_i - a_i^{(3)}) (-1) \times a_i^{(3)} (1 - a_i^{(3)})$$

$$= (0.7513 - 0.61) 0.7513 (1 - 0.7513) = 0.1388$$



Date _____

Similarly from (V)

$$\frac{\partial \delta_2^{(3)}}{\partial a_2^{(2)}} = (a_2^{(3)} - y_2) \times a_2^{(3)} (1 - a_2^{(3)}) \\ = -0.03809$$

lets work on the 2nd terms of
eq. (iv) & (v)

$$\frac{\partial z_1^{(3)}}{\partial a_1^{(2)}} = \frac{\partial}{\partial a_1^{(2)}} [0.6 + \theta_{11}^{(2)} a_1^{(3)} + \theta_{12}^{(2)} a_2^{(2)}] \\ = \theta_{11}^{(3)} = 0.40$$

Similarly for eq. (V)

$$\frac{\partial z_2^{(3)}}{\partial a_1^{(2)}} = \theta_{21}^{(2)} = 0.50$$

eq. (iv) & (v) become

$$\frac{\partial \delta_1^{(3)}}{\partial a_1^{(2)}} = 0.1385 \times 0.40 = 0.0554$$

$$\frac{\partial \delta_2^{(3)}}{\partial a_1^{(2)}} = -0.03809 \times 0.50 = -0.019045$$

Date _____

eq (iii) now becomes

$$\frac{\partial \delta_T}{\partial q_1^{(2)}} = 0.0554 + (-0.01904)$$
$$= 0.03635$$

first term of eq. (ii) is solved

2nd term of eq (ii) becomes

$$\frac{\partial Q_1^{(2)}}{\partial z_1^{(2)}} = q_1^{(2)} (1 - q_1^{(2)})$$
$$= 0.2413$$

Last (3rd) term of eq. (ii) becomes

$$\frac{\partial z_1^{(2)}}{\partial \theta_{11}^{(1)}} = x_1^{(1)} (0.35 + \theta_{11}^{(1)} x_1^{(1)} + \theta_{12}^{(1)} x_2^{(1)})$$
$$= x_1^{(1)} = 0.05$$

Eq (ii) now becomes

$$r\delta_T = 0.03635 \times 0.2413 \times 0.05$$

$$\frac{\partial \theta_{11}^{(1)}}{\partial \theta_{11}^{(1)}} = 0.0004385$$

$$\theta_{11}^{(1)} = \theta_{11}^{(1)} - \alpha \frac{\partial \delta_T}{\partial \theta_{11}^{(1)}} \Rightarrow 0.149236$$