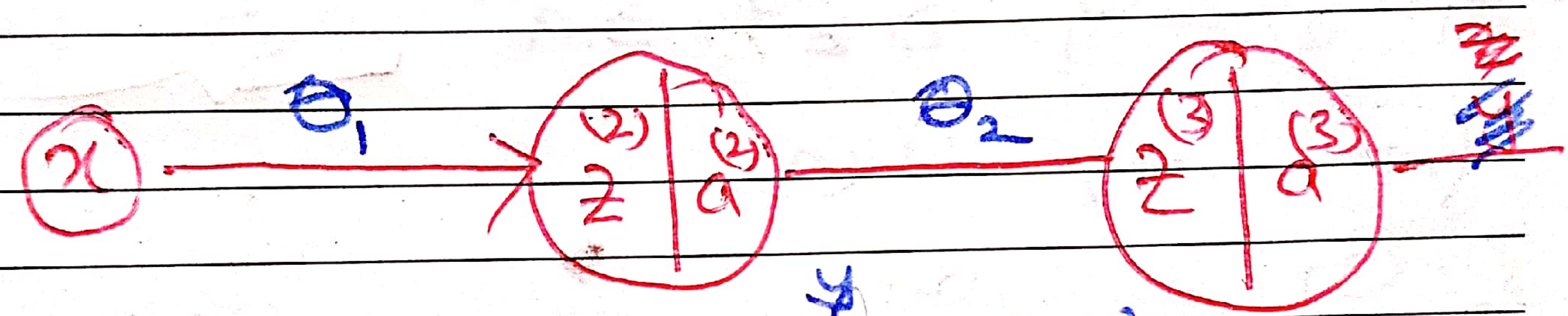
Date

To understand Simple Neural Net



$$\delta = \frac{1}{2}\left(a^{(3)} - y\right)^2$$

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

it can further be elaborated

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

$$\frac{\partial \delta}{\partial \theta_1} = \frac{\partial \delta}{\partial a^{(3)}} \cdot \frac{\partial a^{(3)}}{\partial z^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial a^{(2)}} \cdot \frac{\partial a^{(2)}}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial \theta_1} \quad \text{— (ii)}$$

Date

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

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

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

from eq i (a)

$$\frac{\partial z^{(3)}}{\partial \theta_2} = \frac{\partial}{\partial \theta_2}\left(a^{(2)} \cdot \theta_2\right) = a^{(2)}$$

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

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

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

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

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

Date

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

$$= \frac{1}{1+e^{-z^{(3)}}} \cdot \left[\frac{1+e^{-z^{(3)}}}{1+e^{-z^{(3)}}} - \frac{1}{1+e^{-z^{(3)}}} \right]$$

$$= a^{(3)} \cdot \left(1 - a^{(3)}\right)$$

$$\frac{\partial \varepsilon}{\partial a^{(3)}} = \frac{\partial}{a^{(3)}} \left[\frac{1}{2} \left(\cancel{\ } a^{(3)} - y \right)^2 \right]$$

$$= 2 \times \frac{1}{2} \left(a^{(3)} - y \right) \cdot 1$$

$$= a^{(3)}$$

eq (i)-a become

$$\frac{\partial \varepsilon}{\partial \theta_2} = a^{(2)} \cdot a^{(3)} \left(1 - a^{(3)}\right) \cdot a^{(3)}$$

Similarly eq ii-a being

Date ____________

$$\frac{\partial z^{(2)}}{\partial \theta_1} = \cancel{\theta} \; x$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = a^{(2)}\left(1 - a^{(2)}\right)$$

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

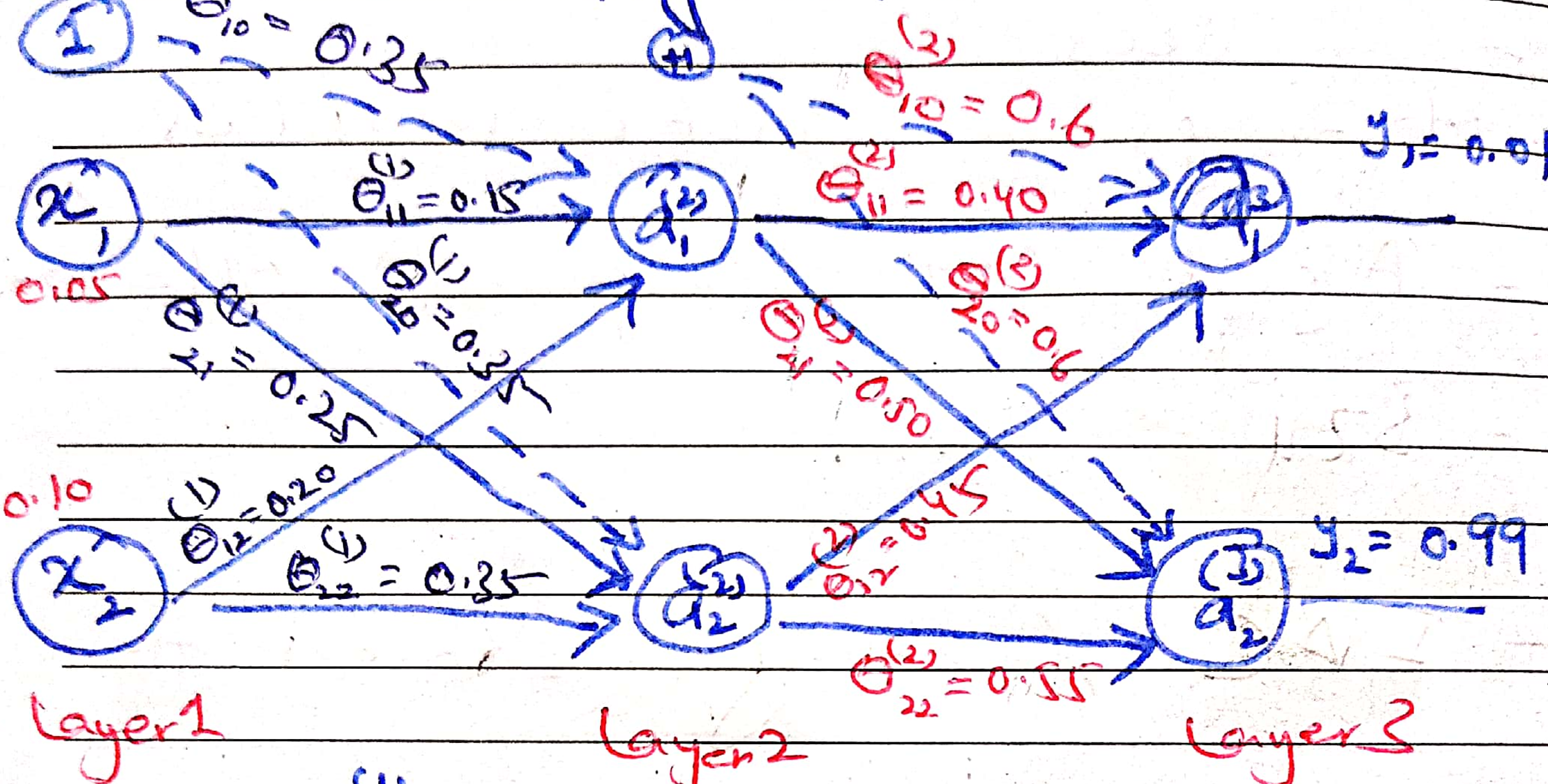
[illegible]

$$\frac{\partial J}{\partial \theta_1} = \cancel{\;} \; x \cdot a^{(2)}\left(1 - a^{(2)}\right) \cdot \theta_2$$

$$a^{(3)}\left(1 - a^{(3)}\right) \quad a^{(3)}$$

Date

Back Propagation Algorithm



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

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

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,

$$z_1^{(3)} = \theta_{10}^{(2)} + \theta_{11}^{(2)} a_1^{(2)} + \theta_{12}^{(2)} a_2^{(2)}$$
$$= 0.6 + 0.40 \times 0.593 + 0.45 \times 0.59$$
$$= 1.105$$

Date ______

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

Similarly,

$$z_2^{(3)} = 0.6 + \Theta_{21}^{(2)} a_1^{(2)} + \Theta_{22}^{(2)} a_2^{(2)}$$

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

Now,

$$\delta_1^{(3)} = \frac{1}{2}\left(y_1 - a_1^{(3)}\right)^2$$

$$= \frac{1}{2}\left(0.01 - 0.7513\right)^2 = 0.2747$$

$$\delta_2^{(3)} = \frac{1}{2}\left(y_2 - a_2^{(3)}\right)^2 = 0.0235$$

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

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

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

from eq. (i)

$$\cancel{\frac{\partial \delta_T}{\partial a_1^{(3)}} = \frac{\partial}{\partial a_1^{(3)}}\left[\frac{1}{2}\left(y_1 - a_1^{(3)}\right)^2 + \frac{1}{2}\left(y_2 - a_2^{(3)}\right)^2\right]}$$

$$\cancel{= 0.75 \times (0.01) = 0.7413}$$

Date

from eq (i)

$$\frac{\partial E_T}{\partial a_1^{(3)}} = \frac{\partial}{\partial a_1^{(3)}}\left[\frac{1}{2}(y_1 - a_1^{(3)})^2 + \frac{1}{2}(y_2 - a_2^{(3)})^2\right]$$

$$= \frac{1}{2} \times 2 (y_1 - a_1^{(3)}) \cdot (-1)$$

$$= a_1^{(3)} - y_1$$

$$= 0.7513 - 0.01 = 0.7412$$

$$\frac{\partial a_1^{(3)}}{\partial z_1^{(3)}} = \frac{\partial}{\partial z_1^{(3)}}\left(\frac{1}{1+e^{-z_1^{(3)}}}\right)$$

$$= \frac{\partial}{\partial z_1^{(3)}}\left(1+e^{-z_1^{(3)}}\right)^{-1}$$

$$= -1\left(1+e^{-z_1^{(3)}}\right)^{-2} \cdot e^{-z_1^{(3)}} \cdot (-1)$$

$$= \left(1+e^{-z_1^{(3)}}\right)^{-2} \cdot e^{-z_1^{(3)}}$$

$$= \frac{1}{\left(1+e^{-z_1^{(3)}}\right)^2} \cdot e^{-z_1^{(3)}}$$

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

add & sbtract 1

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

Date

$$= \frac{1}{1+e^{-z_1^{(3)}}} \cdot \left(\frac{1+e^{-z_1^{(3)}}}{1+e^{-z_1^{(3)}}} - \frac{1}{1+e^{-z_1^{(3)}}} \right)$$

$$= a_1^{(3)} \left(1 - a_1^{(3)} \right)$$

$$= 0.7513 (1 - 0.7513)$$

$$= 0.1868$$

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

$$= a_1^{(2)} = 0.593$$

putting in eq (i)

$$\frac{\partial \delta_T}{\partial \theta_{11}^{(2)}} = 0.7412 \times 0.1868 \times 0.593$$

$$= 0.0821$$

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.,

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

Narsun Studios

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

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

Date

Now for hidden layer i.e., layer 2

$$\frac{\partial \delta_T}{\partial \theta_{11}} = \frac{\partial \delta_T}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial \theta_{11}^{(1)}} \text{ --- (ii)}$$

$$\frac{\partial \delta_T}{\partial a_1^{(2)}} = \frac{\partial \delta_1^{(3)}}{\partial a_1^{(2)}} + \frac{\partial \delta_2^{(3)}}{\partial a_1^{(2)}} \text{ --- (iii)}$$

let's simplify eq. (iii)

$$\frac{\partial \delta_1^{(3)}}{\partial a_1^{(2)}} = \frac{\partial \delta_1^{(3)}}{\partial z_1^{(3)}} \times \frac{\partial z_1^{(3)}}{\partial a_1^{(2)}} \text{ --- (IV)}$$

Similarly

$$\frac{\partial \delta_2^{(3)}}{\partial a_1^{(2)}} = \frac{\partial \delta_2^{(3)}}{\partial z_2^{(3)}} \times \frac{\partial z_2^{(3)}}{\partial a_1^{(2)}} \text{ --- (V)}$$

eq. (IV) can further be elaborated

$$\frac{\partial \delta_1^{(3)}}{\partial z_1^{(3)}} = \frac{\partial \delta_1^{(3)}}{\partial a_1^{(3)}} \times \frac{\partial a_1^{(3)}}{\partial z_1^{(3)}}$$

$$= \frac{\partial}{\partial a_1^{(3)}}\left[\tfrac{1}{2}\left(y_1 - a_1^{(3)}\right)^2\right] \times \frac{\partial}{\partial z_1^{(3)}} \cdot \frac{1}{1+e^{-z_1^{(3)}}}$$

$$= \left(y_1 - a_1^{(3)}\right)(-1) \times a_1^{(3)}\left(1 - a_1^{(3)}\right)$$

$$= \left(a_1^{(3)} - y_1\right) a_1^{(3)}\left(1 - a_1^{(3)}\right)$$

$$= (0.7513 - 0.01)\, 0.7513(1 - 0.7513) = 0.138$$

Date

Similarly from (v)

$$\frac{\partial \delta_2^{(3)}}{\partial z_2^{(3)}} = (a_2^{(3)} - y_2) \times a_2^{(3)}(1-a_2^{(3)})$$

$$= -0.03809$$

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

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

$$= \theta_{11}^{(2)} = 0.40$$

Similary 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

$$\frac{\partial \delta_T}{\partial a_1^{(2)}} = 0.0554 + (-0.01904)$$

$$= 0.03625$$

first term of eq. (ii) is solved

2nd term of eq (ii) becomes

$$\frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} = a_1^{(2)}(1 - a_1^{(2)})$$

$$= 0.2413$$

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

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

$$= x_1 = 0.05$$

Eq (ii) now becomes

$$\frac{\partial \delta_T}{\partial \theta_{11}} = 0.03635 \times 0.2413 \times 0.05$$

$$= 0.0004385$$

$$\theta_{11}^{(1)} = \theta_{11}^{(1)} - \alpha \frac{\partial \delta_T}{\partial \theta_{11}^{(1)}} \Rightarrow 0.149736$$