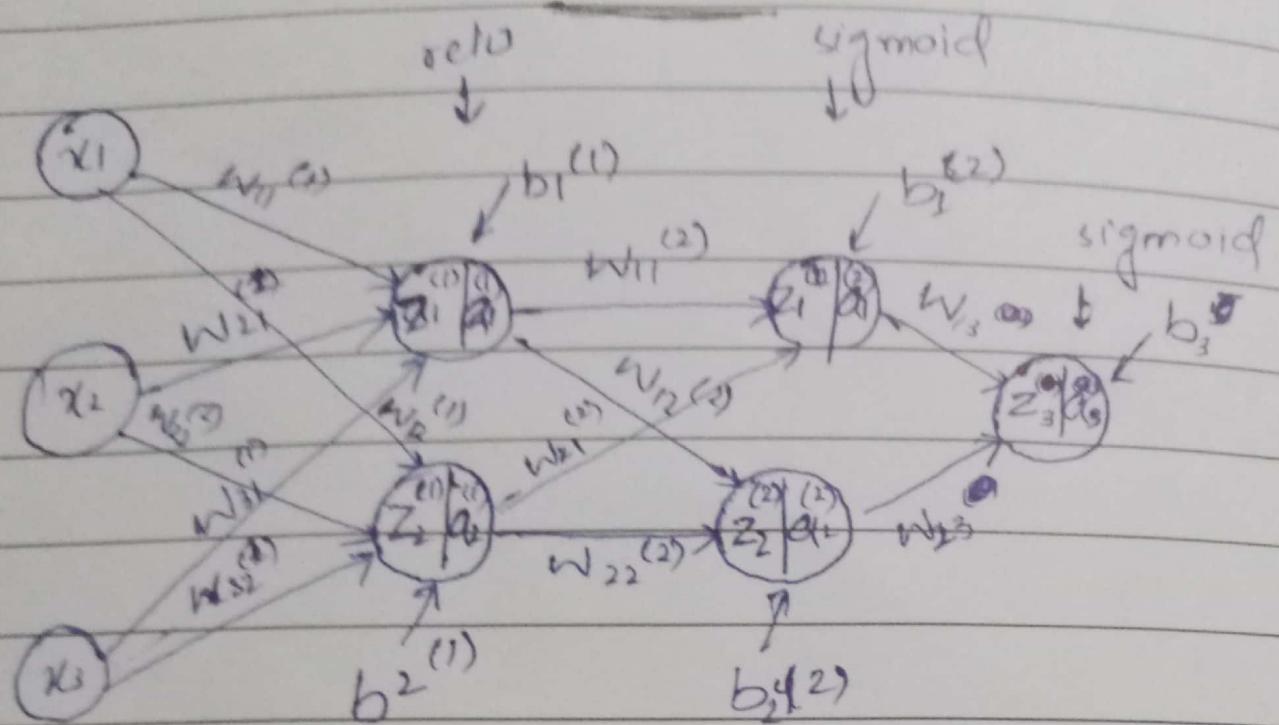


Question 2



FORWARD PASS

hidden layer 1

$$z_1^{(1)} = w_{11}^{(1)}x_1 + w_{21}^{(1)}x_2 + w_{31}^{(1)}x_3 + b_1^{(1)}$$

$$z_2^{(1)} = w_{21}^{(1)}x_1 + w_{22}^{(1)}x_2 + w_{32}^{(1)}x_3 + b_2^{(1)}$$

$$a_1^{(1)} = \text{ReLU}(z_1^{(1)}) = \max(0, z_1^{(1)})$$

$$a_2^{(1)} = \text{ReLU}(z_2^{(1)}) = \max(0, z_2^{(1)})$$

hidden layer 2

$$z_1^{(2)} = w_{11}^{(2)}a_1^{(1)} + w_{21}^{(2)}a_2^{(1)} + b_1^{(2)}$$

$$z_2^{(2)} = w_{22}^{(2)}a_2^{(1)} + w_{32}^{(2)}a_3^{(1)} + b_2^{(2)}$$

Date: _____

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

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

Output layer

$$z_3^{(3)} = w_{13}^{(3)} a_1^{(2)} + w_{23}^{(3)} a_2^{(2)} + b_3$$

$$\hat{y} = a_3^{(3)} = \sigma(z_3^{(3)}) = \frac{1}{1+e^{-z_3^{(3)}}}$$

BACKWARD PASS

Backward pass

MSE

$$\frac{1}{2} (a_3 - y)^2$$

$$\frac{\partial L}{\partial a_3} = a_3 - y$$

Output layer

$$\frac{\partial L}{\partial b_3} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial b_3}$$

$$\frac{\partial L}{\partial w_{13}} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial w_{13}}$$

$$\frac{\partial L}{\partial w_{23}} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_1} \times \frac{\partial z_1}{\partial w_{23}}$$

Hidden layer 2

$$\frac{\partial L}{\partial b_1^{(1)}} = \frac{\partial L}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial a_1^{(1)}} \times \frac{\partial a_1^{(1)}}{\partial b_1^{(1)}}$$

$$\frac{\partial L}{\partial b_2^{(1)}} = \frac{\partial L}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial a_1^{(1)}} \times \frac{\partial a_1^{(1)}}{\partial z_2^{(1)}} \times \frac{\partial z_2}{\partial b_2^{(1)}}$$

$$\frac{\partial L}{\partial w_0^{(1)}} = \frac{\partial L}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial a_1^{(1)}} \times \frac{\partial a_1^{(1)}}{\partial z_2^{(1)}} \times \frac{\partial z_2}{\partial w_0^{(1)}}$$

$$\frac{\partial L}{\partial w_{11}^{(1)}} = \frac{\partial L}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial a_1^{(1)}} \times \frac{\partial a_1^{(1)}}{\partial z_2^{(1)}} \times \frac{\partial z_2}{\partial w_{11}^{(1)}}$$

$$\frac{\partial L}{\partial w_{12}^{(1)}} = \frac{\partial L}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial a_1^{(1)}} \times \frac{\partial a_1^{(1)}}{\partial z_2^{(1)}} \times \frac{\partial z_2}{\partial w_{12}^{(1)}}$$

$$\frac{\partial L}{\partial w_{21}^{(1)}} = \frac{\partial L}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial a_1^{(1)}} \times \frac{\partial a_1^{(1)}}{\partial z_2^{(1)}} \times \frac{\partial z_2}{\partial w_{21}^{(1)}}$$

Hidden layer 1

$$\frac{\partial L}{\partial b_1^{(2)}} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial a_1^{(1)}} \times \frac{\partial a_1^{(1)}}{\partial z_1^{(1)}} \times \frac{\partial z_1^{(1)}}{\partial b_1^{(1)}}$$

+

$$\frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_2^{(2)}} \times \frac{\partial a_2^{(2)}}{\partial z_2^{(2)}} \times \frac{\partial z_2^{(2)}}{\partial a_1^{(1)}} \times \frac{\partial a_1^{(1)}}{\partial z_1^{(1)}} \times \frac{\partial z_1^{(1)}}{\partial b_1^{(1)}}$$

$$\frac{\partial L}{\partial b_2^{(2)}} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial a_2^{(1)}} \times \frac{\partial a_2^{(1)}}{\partial z_2^{(1)}} \times \frac{\partial z_2^{(1)}}{\partial b_2^{(1)}}$$

+

$$\frac{\partial L}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial a_2^{(1)}} \times \frac{\partial a_2^{(1)}}{\partial z_2^{(1)}} \times \frac{\partial z_2^{(1)}}{\partial a_2^{(0)}} \times \frac{\partial a_2^{(0)}}{\partial z_2^{(0)}} \times \frac{\partial z_2^{(0)}}{\partial b_2^{(0)}}$$

$$\frac{d}{dt} \left(\frac{d\mathbf{r}}{dt} \times \frac{d\mathbf{v}}{dt} \right) = \frac{d^2\mathbf{r}}{dt^2} \times \frac{d\mathbf{v}}{dt} + \frac{d\mathbf{r}}{dt} \times \frac{d^2\mathbf{v}}{dt^2}$$

$$\frac{D_L \times D_H \times D_B \times D_{AV}^{(1)} \times D_{Z_2}^{(1)} \times D_{AV}^{(1)} \times D_{Z_1}^{(1)}}{D_H \quad D_B \quad D_A^{(1)} \quad D_{Z_1}^{(1)} \quad D_{Z_2}^{(1)} \quad D_{AV}^{(1)} \quad D_{AV}^{(1)}}$$

$$\frac{\partial L}{\partial w_{21}^{(0)}} = \frac{\partial L}{\partial t} \times \frac{\partial a_3}{\partial a_2} \times \frac{\partial a_2^{(2)}}{\partial a_1^{(0)}} + \frac{\partial L}{\partial t} \times \frac{\partial a_3^{(0)}}{\partial a_2^{(0)}} \times \frac{\partial a_2^{(0)}}{\partial a_1^{(0)} \partial a_2^{(0)}}$$

$$\frac{\partial L \times \partial a_3 \times \partial_{23} \times \partial a_1^{(2)} \times \partial_{23}^{(2)} \times \partial a_1^{(1)} \times \partial_{23}^{(1)}}{\partial a_3 \quad \partial_{23} \quad \partial a_1^{(2)} \quad \partial_{23}^{(2)} \quad \partial a_1^{(1)} \quad \partial_{23}^{(1)} \quad \partial a_1^{(1)}}$$

$$\frac{\partial L}{\partial w_3^{(1)}} = \frac{\partial L}{\partial x} \times \frac{\partial a_3}{\partial x} + \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_1^{(2)}} + \frac{\partial a_3}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial a_1^{(1)}} + \frac{\partial a_3}{\partial z_2^{(2)}} \times \frac{\partial z_2^{(2)}}{\partial a_1^{(1)}} + \frac{\partial a_3}{\partial z_1^{(1)}} \times \frac{\partial z_1^{(1)}}{\partial w_3^{(1)}}$$

$$\frac{\partial L \times \partial a_3 \times \partial_{23} \times \cancel{\partial a_2^{(2)}} \times \cancel{\partial_{22}^{(2)}} \times \partial a_1^{(4)} \times \cancel{\partial_{21}^{(4)}}}{\partial a_3 \quad \partial_{23} \quad \cancel{\partial a_2^{(2)}} \quad \cancel{\partial_{22}^{(2)}} \quad \partial a_1^{(4)} \quad \cancel{\partial_{21}^{(4)}} \quad \cancel{\partial a_{31}^{(4)}}}$$

$$\frac{\partial L}{W_{12}^{(1)}} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial a_2^{(1)}} \times \frac{\partial a_2^{(1)}}{\partial z_2^{(1)}} \times \frac{\partial z_2^{(1)}}{\partial w_2}$$

$$\frac{\partial L}{\partial a_3} = \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_2^{(2)}} \times \frac{\partial a_2^{(2)}}{\partial z_2^{(2)}} \times \frac{\partial z_2^{(2)}}{\partial a_2^{(1)}} \times \frac{\partial a_2^{(1)}}{\partial z_2^{(1)}} \times \frac{\partial z_2^{(1)}}{\partial a_{12}}$$

$$\frac{\partial L}{\partial w_2^{(1)}} = \frac{\partial L}{\partial z} + \frac{\partial a_3}{\partial z} \times \frac{\partial z_3}{\partial a_3} + \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial a_2^{(1)}} + \frac{\partial a_2^{(1)}}{\partial z_2^{(1)}} \times \frac{\partial z_2^{(1)}}{\partial w_2^{(1)}}$$

$\frac{2L}{2a_3} \times \frac{2a_3}{2z_3} \times \frac{2z_3}{2a_1^{(2)}} \times \frac{2a_2^{(2)}}{2z_2^{(2)}} \times \frac{2z_2^{(2)}}{2a_2^{(1)}} \times \frac{2a_2^{(1)}}{2z_2^{(1)}} \times \frac{2z_2^{(1)}}{2a_1^{(1)}}$

URBANE PAPER PRODUCT

$$\frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial w_3}$$

$$= (a_3 - y)(a_3 + 1)a_3 a_1^{(2)}$$

$$\frac{\partial L}{\partial w_3} = (a_3 - y)(a_3 + 1)a_3 a_2^{(2)} \frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial w_3}$$

hidden layer 2

$$\frac{\partial L}{\partial a_1^{(2)}} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_1^{(2)}}$$

$$= (a_3 - y)a_3(a_3 - a_3) \times 2 \left(W_{13}a_1^{(2)} + W_{23}a_2^{(2)} \right)$$

$$= (a_3 - y)a_3(1 - a_3)W_{13}$$

$$\frac{\partial L}{\partial a_2^{(2)}} = (a_3 - y)a_3(1 - a_3)W_{23} \quad \frac{\partial L}{\partial a_2^{(2)}} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_2^{(2)}}$$

$$\frac{\partial L}{\partial a_1^{(2)}} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}}$$

$$= (a_3 - y)a_3(1 - a_3) \times W_{13} \times a_1^{(2)}(1 - a_1^{(2)})$$

$$\frac{\partial L}{\partial z_1^{(2)}} = (a_3 - y)a_3(1 - a_3) \times \frac{\partial z_3}{\partial a_2^{(2)}} \times \frac{\partial a_2^{(2)}}{\partial z_1^{(2)}}$$

$$= (a_3 - y)a_3(1 - a_3)W_{23} \times a_2(1 - a_2)$$

$$\frac{\partial L}{\partial b_1^{(2)}} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial b_1^{(2)}}$$

$$= (a_3 - y)a_3(1 - a_3)W_{13}a_1(1 - a_1)$$

Date:

$$\frac{\partial L}{\partial z_1^{(2)}} = \frac{\partial L}{\partial a_3} \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_2^{(2)}} \times \frac{\partial a_2^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial b_L^{(2)}}$$

$$= (a_3 - y) a_3 (1 - a_3) \times W_{23} \times a_2 (1 - a_2)$$

$$\frac{\partial L}{\partial a_1^{(2)}} = \frac{\partial L}{\partial a_3} \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_2^{(2)}} \times \frac{\partial a_2^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial W_{11}^{(2)}}$$

$$= (a_3 - y) a_3 (1 - a_3) \times W_{13} \times a_1 (1 - a_1) a_1^{(1)}$$

$$\frac{\partial L}{\partial W_{21}^{(2)}} = \frac{\partial L}{\partial a_3} \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_2^{(2)}} \times \frac{\partial a_2^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial W_{21}^{(2)}}$$

$$= (a_3 - y) a_3 (1 - a_3) \times W_{13} \times a_1 (1 - a_1) \times a_2^{(1)}$$

$$\frac{\partial L}{\partial W_{12}^{(2)}} = \frac{\partial L}{\partial a_3} \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_2^{(2)}} \times \frac{\partial a_2^{(2)}}{\partial z_2^{(2)}} \times \frac{\partial z_2^{(2)}}{\partial W_{12}^{(2)}}$$

$$= (a_3 - y) \times a_3 (1 - a_3) \times W_{23} \times a_2 (1 - a_2) \times a_1^{(1)}$$

$$\frac{\partial L}{\partial W_{22}^{(2)}} = \frac{\partial L}{\partial a_3} \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_2^{(2)}} \times \frac{\partial a_2^{(2)}}{\partial z_2^{(2)}} \times \frac{\partial z_2^{(2)}}{\partial W_{22}^{(2)}} \times \frac{\partial W_{22}^{(2)}}{\partial a_2^{(1)}}$$

$$= (a_3 - y) a_3 (1 - a_3) \times W_{23} \times a_2 (1 - a_2) \times a_2^{(1)}$$

a_{b_2}

Hidden layer 1

$$\frac{\partial L}{\partial b_1^{(1)}} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial a_1^{(1)}} \times \frac{\partial a_1^{(1)}}{\partial z_2^{(1)}} +$$

$$\frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_2^{(2)}} \times \frac{\partial a_2^{(2)}}{\partial z_2^{(2)}} \times \frac{\partial z_2^{(2)}}{\partial a_1^{(1)}} \times \frac{\partial a_1^{(1)}}{\partial z_2^{(1)}}$$

$$= (a_3 - y) a_3 (1-a_3) \times w_{13} \times a_1 (1-a_1) \times w_{11}^{(2)} \times a_1 (1-a_1)$$

$$+ (a_2 - y) a_2 (1-a_2) \times w_{23} \times a_2 (1-a_2) \times w_{12}^{(2)} \times a_1 (1-a_1)$$

Date:

$$\frac{\partial L}{\partial b_2^{(1)}} = (a_3 - y) a_3(1-a_3) w_{13} a_1(1-a_1) w_{21}^{(2)} \times a_2(1-a_2)$$
$$+ (a_3 - y) a_3(1-a_3) w_{13} a_2(1-a_2) w_{22}^{(2)} \times a_2(1-a_2)$$

Date:

$$\frac{2a_2^{(1)}}{2w_3^{(1)}} \cdot \frac{2L}{2a_3} \times \frac{2z_3}{2a_1^{(2)}} \times \frac{2a_1^{(2)}}{2z_1^{(2)}} \times \frac{2z_1^{(2)}}{2a_2^{(1)}} \times \frac{2a_2^{(1)}}{2z_2^{(1)}} \times \frac{2z_2^{(1)}}{2w_2^{(1)}}$$

+

$$\frac{2L}{2a_3} \times \frac{2a_3}{2z_3} \times \frac{2z_3}{2a_2^{(2)}} \times \frac{2a_2^{(2)}}{2z_2^{(2)}} \times \frac{2z_2^{(2)}}{2a_2^{(1)}} \times \frac{2a_2^{(1)}}{2z_2^{(1)}} \times \frac{2z_2^{(1)}}{2w_3^{(1)}}$$

$$\frac{\partial L}{\partial w_{11}^{(1)}} = (a_3 - y) a_3 (1-a_3) w_{13} a_1^{(2)} (1-a_1^{(2)}) \times w_{21}^{(2)} \times 1 \times x_1^{(1)}$$

$$+ (a_3 - y) a_3 (1-a_3) w_{23} a_2^{(2)} (1-a_2^{(2)}) \times w_{12}^{(2)} \times 1 \times x_1^{(1)}$$

$$\frac{\partial L}{\partial w_{21}^{(1)}} = (a_3 - y) a_3 (1-a_3) w_{13} a_1^{(2)} (1-a_1^{(2)}) \times w_{11}^{(2)} \times 1 \times x_2^{(1)}$$

$$+ (a_3 - y) a_3 (1-a_3) w_{23} a_2^{(2)} (1-a_2^{(2)}) \times w_{12}^{(2)} \times 1 \times x_2^{(1)}$$

$$\frac{\partial L}{\partial w_{31}^{(1)}} = (a_3 - y) a_3 (1-a_3) w_{13} a_1^{(2)} (1-a_1^{(2)}) \times w_{11}^{(2)} \times 1 \times x_3^{(1)}$$

$$+ (a_3 - y) a_3 (1-a_3) w_{23} a_2^{(2)} (1-a_2^{(2)}) \times w_{12}^{(2)} \times 1 \times x_3^{(1)}$$

URBANE PAPER PRODUCT

4

Date:

$$\frac{\partial L}{\partial W_{12}^{(1)}} = (a_3 - y) a_3 (1-a_3) w_{13} a_1^{(1)} (1-a_1^{(1)}) \times w_{21}^{(1)} \times 1 \times x_1^{(1)}$$

$$+ (a_3 - y) a_3 (1-a_3) w_{23} a_2^{(1)} (1-a_2^{(1)}) \times w_{21}^{(1)} \times 1 \times x_1^{(1)}$$

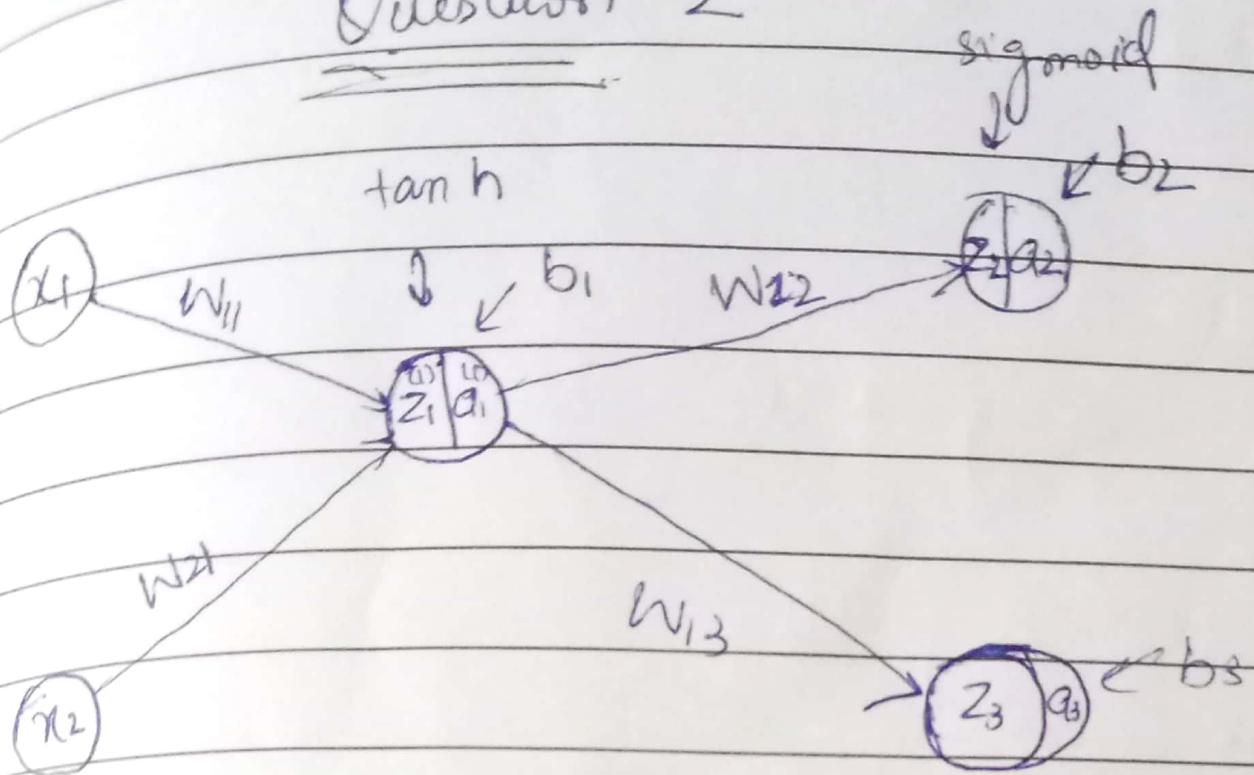
$$\frac{\partial L}{\partial W_{22}^{(2)}} = (a_3 - y) a_3 (1-a_3) w_{13} a_1^{(2)} (1-a_1^{(2)}) \times w_{21}^{(2)} \times 1 \times x_2^{(2)}$$

$$+ (a_3 - y) a_3 (1-a_3) w_{23} a_2^{(2)} (1-a_2^{(2)}) \times w_{22}^{(2)} \times 1 \times x_2^{(2)}$$

$$\frac{\partial L}{\partial W_{32}^{(2)}} = (a_3 - y) a_3 (1-a_3) w_{13} a_1^{(2)} (1-a_1^{(2)}) \times w_{21}^{(2)} \times 1 \times x_3^{(2)}$$

$$+ (a_3 - y) a_3 (1-a_3) w_{23} a_2^{(2)} (1-a_2^{(2)}) \times w_{22}^{(2)} \times 1 \times x_3^{(2)}$$

Question 2



$$\partial L^2 = [y \log a_2 + (1-y) \log(1-a_2)]$$

$$\frac{\partial L^2}{\partial a_2} = \frac{-y}{a_2} - \frac{(1-y)}{1-a_2}$$

$$\frac{\partial L^2}{\partial a_2} = \frac{-y}{a_2} + \frac{1-y}{1-a_2}$$

$$\frac{\partial L}{\partial a_3} = \frac{-y + 1 - y}{a_3 + 1 - a_3}$$

$$\begin{aligned}\frac{\partial L}{\partial b_2} &= \frac{\partial L}{\partial a_2} \times \frac{\partial a_2}{\partial z_2} \times \frac{\partial z_2}{\partial b_2} \\ &= \frac{-y + 1 - y}{a_2 + 1 - a_2} \times a_2(1 - a_2) * \\ &\quad \frac{-y + a_2(y + a_2 - a_2y) * a_2(1 - a_2)}{a_2(1 - a_2)} \\ &= a_2 - y.\end{aligned}$$

$$\frac{\partial L}{\partial b_3} = a_3 - y$$

$$\begin{aligned}\frac{\partial L}{\partial w_{12}} &= \frac{\partial L}{\partial a_2} \times \frac{\partial a_2}{\partial z_2} \times \frac{\partial z_2}{\partial w_{12}} \\ &= (a_2 - y) \approx a_2\end{aligned}$$

$$\frac{\partial L}{\partial w_{23}} = (a_3 - y) a_2$$

$$\begin{aligned}\frac{\partial L}{\partial b_1} &= \frac{\partial L}{\partial a_2} \times \frac{\partial a_2}{\partial z_2} \times \frac{\partial z_2}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial b_1} \\ &+ \end{aligned}$$

$$\frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial b_1}$$

$$\frac{\partial L}{\partial b_1} \cdot (a_1 - y) \times a_2 (1-a_2) \times w_{12} \times 1 - \tanh^2 z_1 *$$

$$\frac{\partial L}{\partial b_2} + (a_2 - y) \times$$

$$(a_2 - y) a_3 (1-a_3) \times w_{13} \times 1 - \tanh^2 z_1$$

$$\frac{\partial L}{\partial b_3} = (a_3 - y) a_2 (1-a_2) \times w_{23} \times 1 - \tanh^2 z_2$$

$$z_1 = (a_1 - y) a_2 (1-a_2)$$

$$\frac{\partial L}{\partial w_{11}} = \frac{\partial L}{\partial a_2} \times \frac{\partial a_2}{\partial z_2} \times \frac{\partial z_2}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial w_{11}}$$

$$= \frac{\partial L}{\partial z_2} +$$

$$\frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial z_3} \times \frac{\partial z_3}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial w_{11}}$$

$$+ (a_2 - y) a_2 (1-a_2) w_{12} \times \frac{1 - \tanh^2 z_1}{1 + \tanh^2 z_1} \times x_1 +$$

$$(a_3 - y) a_3 (1-a_3) w_{13} \times \frac{1 - \tanh^2 z_1}{1 + \tanh^2 z_1} \times x_1$$

$$\frac{\partial L}{\partial w_{12}} \cdot (a_2 - y) a_2 (1-a_2) w_{12} \times x_2 +$$

$$\frac{\partial L}{\partial w_{13}} \cdot (a_3 - y) a_3 (1-a_3) w_{13} \times x_2 \times \frac{1 - \tanh^2 z_1}{1 + \tanh^2 z_1}$$

forward pass

$$z_1 = w_{11}x_1 + w_{21}x_2 + b_1$$

$$a_1 = \tanh(z_1) = \frac{e^{z_1} - e^{-z_1}}{e^{z_1} + e^{-z_1}}$$

$$z_2 = a_1 w_{12} + b_2$$

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

$$z_3 = a_2 w_{13} + b_3$$

$$a_3 = \sigma(z_3) = \frac{1}{1+e^{-z_3}}$$

i) squared loss function.

$$x_1 = 6$$

$$x_2 = 2$$

$$x_3 = 5 \rightarrow 3^{\text{rd}} \text{ it}$$

$$x_1 = 5$$

$$x_2 = 8$$

$$x_3 = 6 \rightarrow 2^{\text{nd}} \text{ it}$$

forward prop.

$$z_1^{(1)} = 6(0.42) + 2(0.72) + 5(0.21) + 0.01$$

$$a_1^{(1)} = 5.02 \quad \text{ReLU}^{\max}(0; z_1, w)$$

10)

$$z_2^{(1)} = 6(0.11) + 2(0.3) + 5(0.65) + 0.01$$

$$a_2^{(1)} = 4.52 \quad \text{ReLU} = \max(0, 4.52) = 4.52$$

11)

$$z_3^{(1)} = 5(0.2(0.19) + 4.52(0.29) + 0.05 = 2.21$$

$$a_3^{(1)} = 1 / (1 + e^{-2.21}) = 0.9011$$

11)

$$z_1^{(2)} = 5(0.2(0.91) + 4.52(0.9011) + 0.05 = 8.69$$

$$a_1^{(2)} = 1 / (1 + e^{-8.69}) = 0.999$$

10)

15)

$$z_2^{(2)} = 0.9011(0.61) + 0.999(0.39) + 0.08$$

$$= 1.01928$$

$$a_2^{(2)} = 0.734$$

$$\text{LF} = \frac{1}{2} (1 - 0.734)^2 \Rightarrow 0.0353$$

10)

15)

Backward Prop Calculation:

$$\frac{\partial L}{\partial} = (0.734 - 1) 0.734 (1 - 0.734) \times 0.9011$$

$$\frac{\partial W_{13}}{\partial} = -0.04680$$

$$W_{13} = W_{13} - \alpha \frac{\partial L}{\partial W_{13}}$$

$$\frac{\partial}{\partial W_{13}}$$

$$\therefore 0.61 - 0.7(-0.0468) = 0.642$$

$$\frac{\partial L}{\partial} = (0.734 - 1) 0.734 (1 - 0.734) \times 0.999 \Rightarrow$$

$$W_{23} = 0.14317 - 0.05188$$

$$W_{23} = W_{23} - \alpha \frac{\partial L}{\partial W_{23}} = 0.39 - 0.7(-0.051)$$

$$\frac{\partial}{\partial W_{23}} = 0.4257$$

2nd HL

$$\frac{\partial L}{\partial W_{12}} = (0.734 - 1) 0.734 (1 - 0.734) \times 0.61 \times 0.9011 (1 - 0.9011) \times 5.02$$

$$\Rightarrow -0.014$$

$$W_{11} = 0.17 - 0.7(-0.014) = 0.1798$$

$$\frac{\partial L}{\partial W_{21}}$$

$$\frac{\partial L}{\partial W_{21}} = -0.014 \quad \frac{\partial L}{\partial W_{21}} = -0.014$$

$$W_{22}^{(2)} = W_{22}^{(1)} \alpha \frac{\partial L}{\partial W_{21}}$$

$$\frac{\partial}{\partial W_{21}}$$

$$= 0.29 - 0.7(-0.014)$$

$$= 0.3001$$

Date:

$$\frac{\partial L}{\partial w_{12}^{(2)}} = (0.734 - 1) 0.734 (1 - 0.734) \times 0.39 \times 0.999 (1 - 0.999) \times 5.02 \\ = -0.00010$$

$$w_{12}^{(2)} = 0.91 - 0.7 \times 0.00010 \\ = 0.910$$

$$\frac{\partial L}{\partial w_{22}^{(2)}} = -0.00010$$

$$w_{22}^{(2)} = w_{22}^{(2)} - \frac{\partial L}{\partial w_{22}^{(2)}} = 0.08 \times 0.7 (-0.00010) \\ = 0.080$$

1st hidden layer

$$\frac{\partial L}{\partial w_{11}^{(1)}} = (0.734 - 1)(0.734)(1 - 0.734)(0.61) \times 0.9011 \\ \times (1 - 0.9011) \times 0.17 \times 1 \times 6 \Rightarrow + (0.734 - 1) \\ (0.734)(1 - 0.734) \times 0.39 \times 0.999 \\ (1 - 0.999) \Rightarrow -0.00305$$

$$\frac{\partial L}{\partial w_{11}} = w_{11} - \alpha \frac{\partial L}{\partial w_{11}} = 0.42 - 0.7 \frac{(-0.00305)}{305} \\ = 0.422$$

$$\frac{\partial L}{\partial w_{21}} = -0.0010$$

$$w_{21} = 0.72 - 0.7 (-0.0010) \\ = 0.720$$

$$\frac{\partial L}{\partial w_{31}} = -0.0025$$

$\frac{\partial w_{31}}{\partial w_{31}}$

$$w_{31} = w_{31} - \alpha \frac{\partial L}{\partial w_{31}} = 0.211$$

$$\frac{\partial L}{\partial w_{12}} = -0.015 \quad w_{12\text{new}} = 0.12$$

$\frac{\partial w_{12}}{\partial w_{12}}$

$$\frac{\partial L}{\partial w_{22}} = -0.005116 \quad w_{22\text{new}} = 0.30$$

$\frac{\partial w_{22}}{\partial w_{32}}$

$$\frac{\partial L}{\partial w_{32}} = -0.012 \Rightarrow w_{32\text{new}} = 0.65$$

New biases: ~~b_1, b_2~~

$$\frac{\partial L}{\partial b_3} = -0.05 \quad b_{3\text{new}} = 0.116$$

$$\frac{\partial L}{\partial b_2^{(2)}} = -0.0028 \quad b_{2\text{new}}^{(2)} = 0.05$$

$\frac{\partial b_2^{(2)}}{\partial b_2^{(1)}}$

$$\frac{\partial L}{\partial b_2^{(1)}} = -0.0001 \quad b_{2\text{new}}^{(1)} = 0.05$$

$$\frac{\partial L}{\partial b_1} = -0.0005 \quad b_{1\text{new}} = 0.011$$

$$\frac{\partial L}{\partial b_1} = -0.0025 \quad b_{1\text{new}} = 0.01$$

for iteration 2

$$z_1^{(1)} = 9.13 \quad a_1^{(1)} = 9.13$$

$$z_2^{(1)} = 6.86 \quad a_2^{(1)} = 6.86$$

$$z_1^{(2)} = 3.59 \quad a_1^{(2)} = 3.59 \quad 0.973$$

$$z_2^{(2)} = 8.90 \quad a_2^{(2)} = 8.90 \quad 0.999$$

$$z_3 = 1.0636 \rightarrow 0.743$$

$$y = \frac{1}{2} (0.743)^2, D = 296$$

Date:

$$W_{13}^{(1)} = 0.5$$

$$W_{13}^{(2)} = 0.3$$

$$W_{12}^{(1)} = 0.16$$

$$W_{12}^{(2)} = 0.28$$

$$W_{12}^{(3)} = 0.9$$

$$W_{22}^{(1)} = 0.08$$

$$W_{13} = 0.04$$

$$W_{21} = 0.21$$

$$W_{31} = 0.21$$

$$W_{12} = 0.10$$

$$W_{22} = 0.29$$

$$W_{32} = 0.65$$

$$b^3 = 0.02 \quad b_1^{(1)} = 0.05 \quad b_1^{(2)} = 0.05, b_2^{(1)} = 0.01$$

Gross entropy loss.

$$L_2 = [1 - \log(0.743) + (1-1) \log(1-0.743)] \\ = 0.129$$

$$\frac{\partial L}{\partial W_{13}} = (0.743 - 1) 0.9011 = -0.2 \quad W_{13} = 0.7$$

$$\frac{\partial L}{\partial W_{23}} = (0.743 - 1) 0.999 = 0.25 \quad W_{23} = 0.5$$

$$\frac{\partial L}{\partial W_{31}} = (0.743 - 1) \times 0.6 \times 0.9011 (1 - 0.9011) \times 0.1$$

$$\frac{\partial L}{\partial W_{12}^{(1)}} = -0.07$$

$$W_{11} = 0.2$$

$$\frac{\partial L}{\partial W_{21}^{(1)}} = 0.07 \quad W_{21}^{(1)} = 0.3$$

$$\frac{\partial L}{\partial W_{12}^{(2)}} = -0.006 \quad W_{12}^{(2)} = 0.9$$

$$\frac{\partial L}{\partial W_{12}^{(3)}} = -0.006 \quad W_{12}^{(3)} = 0.68$$

$$W_{11}^{(1)} \text{ new } = 0.4$$

$$b_3^{\text{new}} = 0.2$$

$$W_{21}^{(1)} = 0.7$$

$$b_1^{\text{new}} = 0.06$$

$$W_{31}^{(1)} = 0.2$$

$$b_2^{\text{new}} = 0.05$$

$$W_{12}^{(1)} = 0.12$$

$$b_3^{\text{new}} = 0.01$$

$$W_{22}^{(1)} = 0.3$$

$$b_2^{\text{new}} = 0.01$$

$$W_{32}^{(1)} = 0.66$$

$$Z_{AS} = 1.6 \quad a_1 = 0.8$$

$$\text{Loss} = 0.7$$

$$W_{13}^{\text{new}} = -0.2$$

$$W_{23}^{\text{new}} = -0.007$$

$$W_{11}^{(2)} \text{ new } = 0.14$$

$$W_{21}^{(2)} \text{ new } = 0.2$$

$$W_{12}^{(2)} \text{ new } = 0.91$$

$$W_{22}^{(2)} \text{ new } = 0.08$$

$$W_{11}^{(1)} \text{ new } = 0.87$$

$$W_{21}^{(1)} \text{ new } = 0.7$$

$$W_{31}^{(1)} \text{ new } = 0.21$$

$$W_{21}^{(1)} \text{ new } = 0.11$$

$$W_{22}^{(1)} \text{ new } = 0.28$$

$$W_{32}^{(1)} \text{ new } = 0.65$$

$$b_3^{\text{new}} = -0.31$$

$$b_1^{\text{new}} = 0.05$$

$$b_2^{\text{new}} = 0.05$$

$$b_1^{\text{new}} = 0.01$$

$$b_2^{\text{new}} = 0.01$$

URBANE PAPER PRODUCT