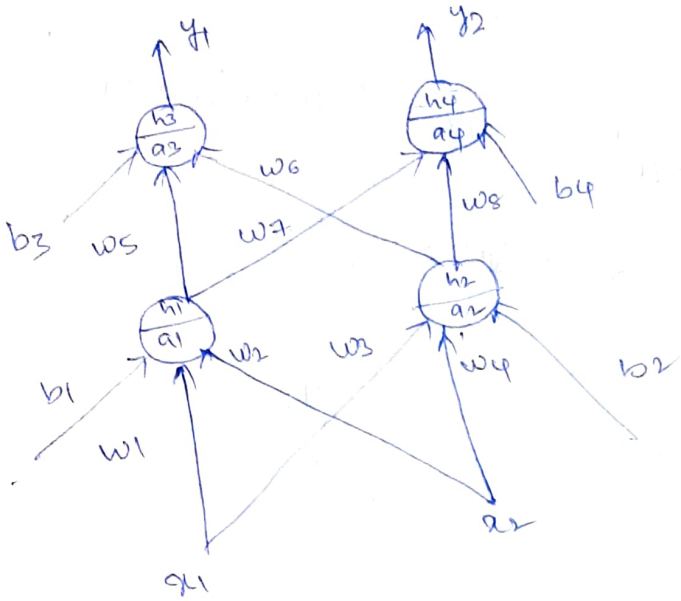
Team:

G R Uday Kumar Reddy:213170007

Mohit Punasiya:213170006

Google colab link :
https://colab.research.google.com/drive/1aVRQqZZS44FhjQtgMtlPq9HCf3T8hG_5?usp=sharing

Given Neural Network.



Loss = MSE + Cross-Entropy

$$L = \sum_{i=1}^{2} \left[(y_i - \hat{y}_i)^2 - (y_i \log \hat{y}_i + (1-y_i) \log(1-\hat{y}_i)) \right]$$

Softmax output layer :-

$$h_3 = y_1 = \frac{e^{a_3}}{e^{a_3} + e^{a_4}} \quad ; \quad h_4 = y_2 = \frac{e^{a_4}}{e^{a_3} + e^{a_4}}$$

Hidden layers

$$h_1 = \tanh(a_1)$$

$$h_2 = \tanh(a_2)$$

Where

$$a_1 = w_1 x_1 + w_2 x_2 + b_1$$

$$a_2 = w_3 x_1 + w_4 x_4 + b_2$$

$$h_1 = \tanh(a_1)$$

$$h_2 = \tanh(a_2)$$

$$a_3 = w_5 h_1 + w_6 h_2 + b_3$$

$$a_4 = w_7 h_1 + w_8 h_2 + b_4$$

$$h_4 = \frac{e^{a_4}}{e^{a_3} + e^{a_4}} = y_2 \quad ; \quad h_3 = \frac{e^{a_3}}{e^{a_3} + e^{a_4}} = y_1$$

Backward pass

$$\frac{\partial L}{\partial h_3} = \frac{\partial L}{\partial \hat{y}_1} = \frac{\partial}{\partial \hat{y}_1}\left[(y_1-\hat{y}_1)^2 - (y_1 \log \hat{y}_1 + (1-y_1)\log(1-\hat{y}_1))\right]$$

$$= -2(y_1-\hat{y}_1) - \left(\frac{y_1}{\hat{y}_1} + \frac{(y_1-1)}{(1-\hat{y}_1)}\right)$$

at output layer

$$\frac{\partial L}{\partial \hat{y}_1} = -\left[2(y_1-\hat{y}_1) + \frac{y_1-\hat{y}_1}{\hat{y}_1(1-\hat{y}_1)}\right]$$

11y

$$\frac{\partial L}{\partial \hat{y}_2} = -\left[2(y_2-\hat{y}_2) + \frac{(y_2-\hat{y}_2)}{\hat{y}_2(1-\hat{y}_2)}\right]$$

$$\frac{\partial L}{\partial a_3} = \frac{\partial L}{\partial h_3}\cdot\frac{\partial h_3}{\partial a_3}$$

$$\frac{\partial L}{\partial a_3} = -\left[2(y_1-h_3) + \frac{(y_1-h_3)}{h_3(1-h_3)}\right] \times \left[h_3(1-h_3)\right]$$

$$\frac{\partial L}{\partial a_3} = -\left[(y_1-h_3)(2h_3(1-h_3)+1)\right]$$

$$\frac{\partial L}{\partial a_3} = (h_3-y_1)(2h_3(1-h_3)+1)$$

$$\frac{\partial L}{\partial a_4} = (h_4-y_2)(2h_4(1-h_4)+1)$$

$$\frac{\partial L}{\partial w_5} = \frac{\partial L}{\partial a_3}\cdot\frac{\partial a_3}{\partial w_5}$$

$$= \left(\frac{\partial L}{\partial a_3}\right)\cdot h_1$$

$$\frac{\partial L}{\partial w_6} = \frac{\partial L}{\partial a_3} \times h_2$$

$$\frac{\partial L}{\partial b_3} = \frac{\partial L}{\partial a_3}\cdot 1$$

$$\frac{\partial L}{\partial w_7} = \frac{\partial L}{\partial a_4} \times h_1$$

$$\frac{\partial L}{\partial w_8} = \frac{\partial L}{\partial a_4} \times h_2$$

$$\frac{\partial L}{\partial b_4} = \frac{\partial L}{\partial a_4}$$

$$\frac{\partial L}{\partial h_1} = \frac{\partial L}{\partial a_3} \times \frac{\partial a_3}{\partial h_1}$$

$$= \frac{\partial L}{\partial a_3} \times w_5$$

$$\frac{\partial L}{\partial h_2} = \frac{\partial L}{\partial a_4} \cdot w_6$$

$$\frac{\partial L}{\partial a_1} = \frac{\partial L}{\partial h_1} \times \frac{\partial h_1}{\partial a_1}$$

$$\frac{\partial L}{\partial a_1} = \frac{\partial}{\partial a_1}(\tanh(a_1))$$

$$= \frac{\partial}{\partial a_1}\left(\frac{2}{1+e^{-2a_1}} - 1\right)$$

$$= 1 - h_1^2$$

$$\boxed{\frac{\partial h_1}{\partial a_1} = 1 - h_1^2}$$

$$\boxed{\frac{\partial h_2}{\partial a_2} = 1 - h_2^2}$$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial a_1} \times \frac{\partial a_1}{\partial w_1}$$

$$= \frac{\partial L}{\partial a_1} \times x_1$$

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial a_1} \times \frac{\partial a_1}{\partial w_2}$$

$$= \frac{\partial L}{\partial a_1} \cdot x_2$$

$$\frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial a_2} \cdot x_1$$

$$\frac{\partial L}{\partial w_4} = \frac{\partial L}{\partial a_2} \cdot x_2$$

$$\frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial a_2} \cdot 1$$

$$\frac{\partial L}{\partial a_3} = (h_3 - y_1)\left(h_3(1-h_3)+1\right)$$

$$\frac{\partial L}{\partial a_4} = (h_4 - y_2)\left(h_4(1-h_4)+1\right)$$

(1)

$$\frac{\partial L}{\partial w_5} = \frac{\partial L}{\partial a_3} \times h_1$$

$$\frac{\partial L}{\partial w_6} = \frac{\partial L}{\partial a_3} \times h_2$$

(2)

$$\frac{\partial L}{\partial w_7} = \frac{\partial L}{\partial a_4} \times h_1$$

$$\frac{\partial L}{\partial w_8} = \frac{\partial L}{\partial a_4} \times h_2$$

$$\frac{\partial L}{\partial b_4} = \frac{\partial L}{\partial a_4} \cdot 1$$

③ $\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial a_1} \cdot x_1$

$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial a_1} \cdot x_2$

$\frac{\partial L}{\partial b_1} = \frac{\partial L}{\partial a_1}$

$\frac{\partial L}{\partial a_1} = (1-h_1^2) \times \frac{\partial L}{\partial a_3} \times w_5 + (1-h_1^2) \times \frac{\partial L}{\partial a_4} \times w_7$

$= (1-h_1^2)\left[\frac{\partial L}{\partial a_3} w_5 + \frac{\partial L}{\partial a_4} w_7\right]$

$\frac{\partial L}{\partial a_2} = (1-h_2^2)\left[\frac{\partial L}{\partial a_3} w_6 + \frac{\partial L}{\partial a_4} w_8\right]$

④ $\frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial a_2} \cdot x_1$

$\frac{\partial L}{\partial w_4} = \frac{\partial L}{\partial a_2} \cdot x_2$

$\frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial a_2} \cdot 1$