DEEP LEARNING

1.

Network diagram:

x_1, $w_1 = -1.7$
x_2, $w_2 = 0.1$ → Σ → S → σ → h_1

h_1 → $w_3 = -0.2$ → Σ → S_3 → σ → $\hat{y}$

x_3, $w_3 = -0.6$
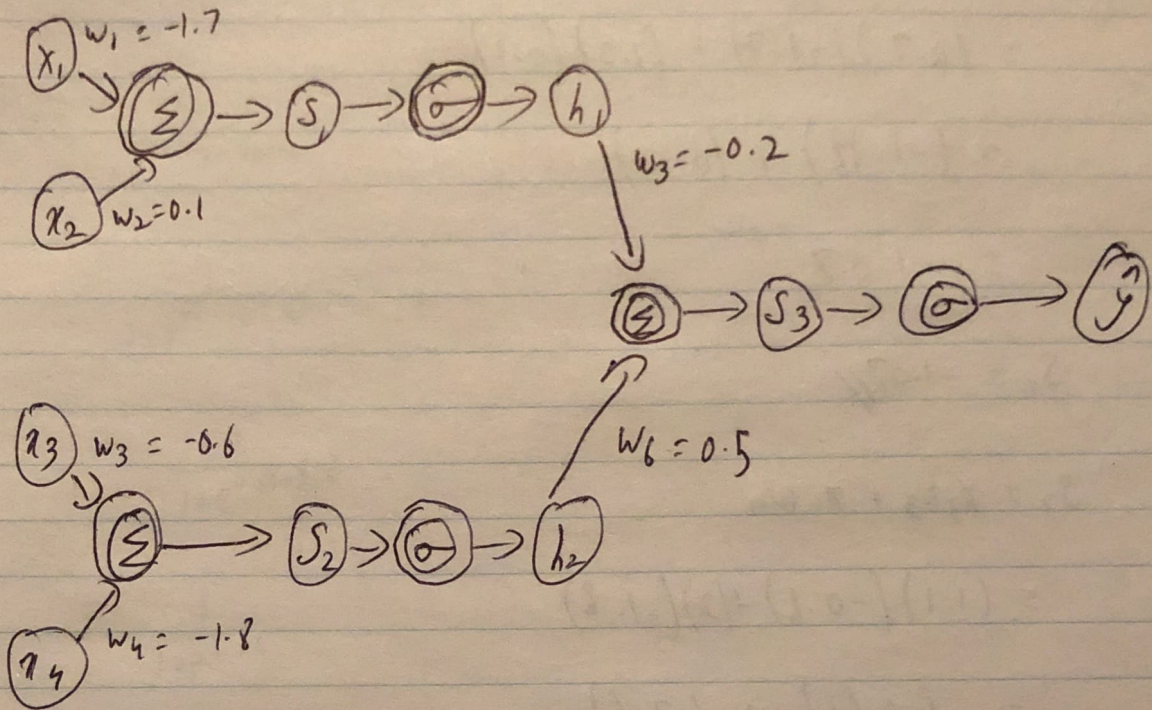x_4, $w_4 = -1.8$ → Σ → S_2 → σ → h_2

h_2 → $w_6 = 0.5$ → Σ

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

$$h_1 = \frac{1}{1 + e^{-w_1 x_1 - x_2 w_2}}$$

L_2 loss $L(y, \hat{y}) = \| \hat{y} - y \|^2$

$x_1 = 0.7$, $x_2 = 1.2$, $x_3 = 1.1$, $x_4 = 2$

$y = 0.5$

$$S_1 = x_1 w_1 + x_2 w_2$$

$$= (0.7)(-1.7) + (1.2)(0.1)$$

$$= (-1.19) + (0.12)$$

$$= -1.07$$

$$S_1 = -1.07 //$$

$$S_2 = x_3 w_3 + x_4 w_4$$

$$= (1.1)(-0.6) + (2)(-1.8)$$

$$= (-0.66) + (-3.6)$$

$$S_2 = -4.26 //$$

$$h_1 = \frac{1}{1 + e^{-x_1 w_1 - x_2 w_2}}$$

$$= \frac{1}{1 + e^{-(0.7)(-1.7) - (1.2)(0.1)}}$$

$$= \frac{1}{1 + e^{(1.19) - (0.12)}}$$

$$= \frac{1}{1 + e^{1.07}}$$

$$= \frac{1}{3.915}$$

$h_1 = 0.255 \; //$

$h_2 = \dfrac{1}{1+e^{-x_3 w_3 - x_4 w_4}}$

$= \dfrac{1}{1+e^{-(1.1)(-0.6)-(2)(-1.8)}}$

$= \dfrac{1}{1+e^{0.66+3.6}}$

$= \dfrac{1}{1+e^{4.26}}$

$= \dfrac{1}{1+70.5}$

$= \dfrac{1}{71.5}$

$h_2 = 0.0139 \; //$

$S_3 = h_1 w_5 + h_2 w_6$

$= (0.255)(-0.2) + (0.0139)(0.5)$

$= -0.051 + 0.00695$

$S_3 = -0.0445 \; //$

$$\hat{y} = \frac{1}{1+e^{-h_1 w_5 - h_2 w_6}}$$

$$= \frac{1}{1+e^{-(0.255)(-0.2) - (0.0139)(0.5)}}$$

$$= \frac{1}{1+e^{-0.051 - 0.0695}}$$

$$= \frac{1}{1+e^{0.40440 5}}$$

$$= \frac{1}{1+0.653}$$

$$\hat{y} = 0.4889 //$$

Gradient of L2 loss function

$||\hat{y}-y||^2$ is $2||\hat{y}-y|| - \frac{\partial E}{\partial y}$

Using backward propagation

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial S_3} \times \frac{\partial S_3}{\partial h_n} \times \frac{\partial h_n}{\partial S_1} \times \frac{\partial S_1}{\partial w_1}$$

$$= 2||\hat{y}-y|| \times \sigma'(S_3) \times w_5 \times \sigma'(S_1) \times x_1$$

because $w_5 = \frac{\partial S_3}{\partial h_n}$ and $\frac{\partial S_1}{\partial w_1}$

$$\sigma'(S_3) = \sigma(S_3)(1 - \sigma S_3)$$
$$\sigma' S_1 = \sigma(S_1)(1 - \sigma(S_1))$$

$$\frac{\partial E}{\partial w_1} = 2[(0.4889 - 0.511) \times (\sigma(S_3)(1 - \sigma(S_3)))] \times (-0.2) \times (0.7) \times [\sigma(S_1)(1-\sigma(S_1))]$$

$$= 2(0.107) \times [(0.2559)(1 - 0.2554)] \times 0.2 \times [(0.4884)(1 - 0.4884)] \times 0.7$$

$$\frac{\partial E}{\partial w_1} = -0.00114 \quad //$$