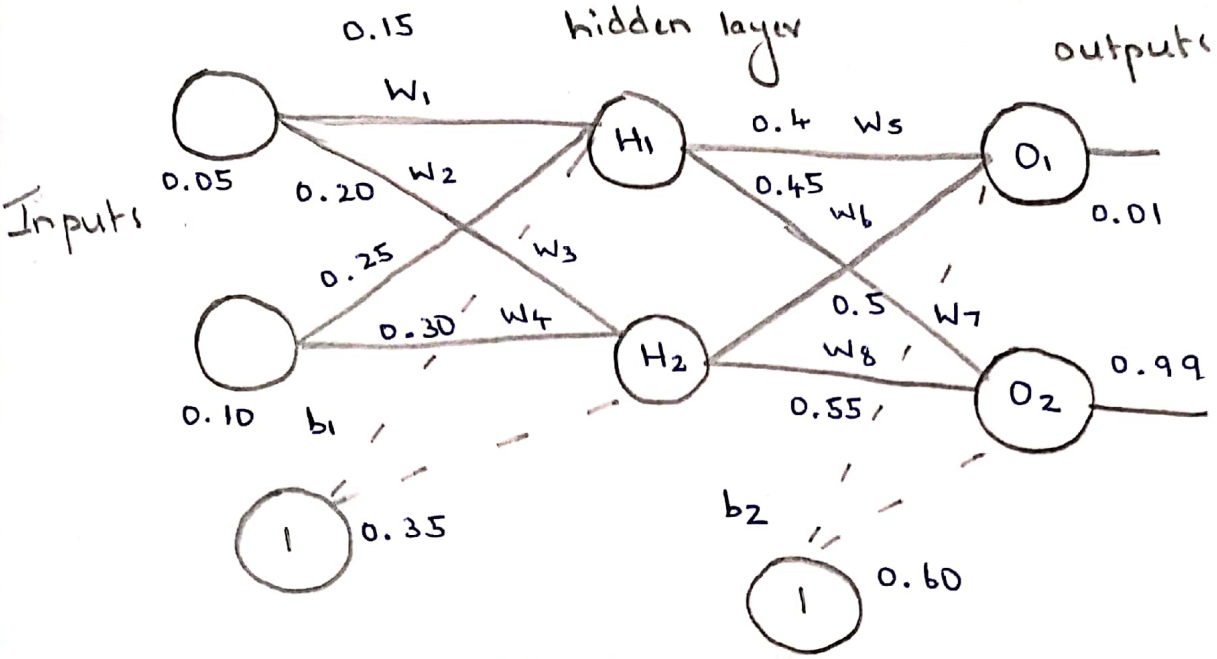
Backprop



0.025
0.015

Forward Pass

$$h_1 = w_1 i_1 + w_2 i_2 + b_1 \times 1$$

$$= (0.15 \times 0.05) + (0.20 \times 0.1) + (0.35 \times 1)$$

$$= 0.3775$$

+ Activation (Sigmoid)

$$\text{out } H_1 = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-0.3775}} = 0.593$$

$$h_2 = w_3 i_3 + w_4 i_4 + b_2 \times 1$$

$$= (0.25 \times 0.10) + (0.30 \times 0.05) + 0.35$$

$$= 0.39$$

+ Activation (Sigmoid)

$$\text{out } H_2 = \frac{1}{1 + e^{-0.39}} = 0.596$$

$$net_{o_1} = h_1 \times w_5 + h_2 \times w_6 + b_2 \times 1$$

$$= (0.4 \times 0.593) + (0.45 \times 0.596) + 0.6$$

$$= 1.105$$

$$out_{o_1} = \frac{1}{1 + e^{-1.105}} = 0.751$$

0.2965
0.3278
0.6243

$$net_{o_2} = h_1 \times w_7 + h_2 \times w_8 + b_2 \times 1$$

$$= (0.5 \times 0.593) + (0.55 \times 0.596) + 0.6$$

$$= 1.2243$$

$$out_{o_2} = \frac{1}{1 + e^{-1.2243}} = 0.7729$$

Now we have 2 outputs.

Calculate loss now.

$$E_{total_1} = \sum \frac{1}{2}(\text{true} - \text{predicted})^2$$

$$= \frac{1}{2}(0.01 - 0.751)^2 = 0.2748$$

$$E_{total_2} = \frac{1}{2}(0.99 - 0.7729)^2 = 0.0235$$

$$\text{Sum} = 0.2748 + 0.0235$$

$$= 0.2983$$

Backward Pass

For weight 5 $\frac{\partial \text{total}}{\partial w_5} = \frac{\partial \text{total}}{\partial \hat{o}_{o_1}^{out}} \times \frac{\partial \hat{o}_{o_1}^{out}}{\partial \text{net}_{o_1}} \times \frac{\partial \text{net}_{o_1}}{\partial w_5}$

$\frac{\partial \text{total}}{\partial \text{out}_{o_1}} = \text{total} = \frac{1}{2}(\text{target} - \text{pred})^2 + \frac{1}{2}(\text{target} - \text{pred})^2$

(o_1 ↑ ; o_2 ↑ ; partial derivative so 0)

piece 1 ↗

$\frac{\partial \text{total}}{\partial \text{out}_{o_1}} = (\text{target}_{o_1} - \text{out}_{o_1}) \times -1$

$= -(0.01 - 0.751) = 0.741$

$\text{out}_{o_1} = \frac{1}{1 + e^{-\text{net}_{o_1}}}$

$\frac{\partial \text{out}_{o_1}}{\partial \text{net}_{o_1}} = (\text{out}_{o_1})(1 - \text{out}_{o_1})$ $\boxed{\sigma(1-\sigma)}$

↑ piece 2

$= (0.751)(1 - 0.751) = 0.186$

$\text{net}_{o_1} = h_1 \times w_5 + h_2 \times w_6 + b_2 \times 1$

(w_6 ↑ 0 ; b_2 ↑ 0)

$\frac{\partial \text{net}_{o_1}}{\partial w_5} = h_1 = 0.593$

↑ piece 3.

Combine Now

$\frac{\partial \text{total}}{\partial w_5} = 0.741 \times 0.186 \times 0.593$

$= 0.08216$

After gradient calculation, we need to update the learning rate

$$W_{5\,new} = W_5 - lr \times \frac{\partial total}{W_5}$$

↑ hyperparameter

$$= 0.4 - 0.5 \times 0.0821$$

$$= 0.3589$$

Do the same for $W_1, W_2, W_3, W_4, W_6, W_7, W_8$.

XXX