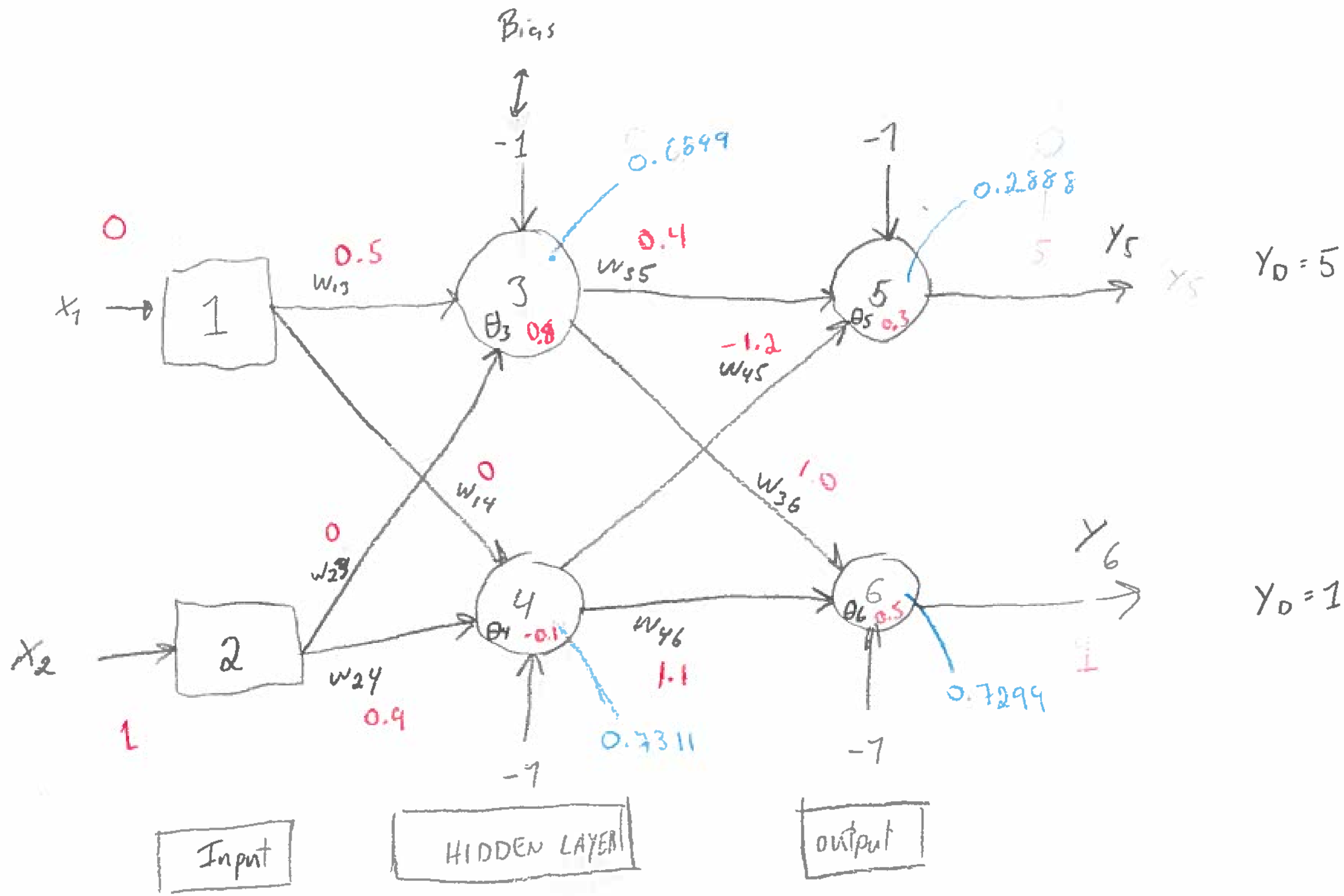
Bias
-1
-1
-1
-1
x_1
x_2
0
1
1
2
3
4
5
6
w_{13}
0.5
w_{23}
0
w_{14}
0
w_{24}
0.4
w_{35}
0.4
w_{45}
-1.2
w_{36}
1.0
w_{46}
1.1
θ_3 0.8
θ_4 -0.1
θ_5 0.3
θ_6 0.5
0.6649
0.7311
0.2888
0.7299
Y_5
Y_6
$Y_D = 5$
$Y_D = 1$
Input
HIDDEN LAYER
output

Tip To Self:

Gjør som i artikkel finn alle verdier,
Så oppdater vekter og θ.
På Denne måten er det lettere å Ikke bruke de nye vektene når man
skal bruke de "gamle" verdiene.

STEP 1 : Done

$\boxed{\text{Sigmoid: } \frac{1}{1+e^{-x}}}$ $\quad B \cdot \theta$ $\quad x'$

~~Step 2~~

HL: 3

$$Y_3 = \text{Sigmoid}\left(\underbrace{x_1 w_{13} + x_2 w_{23}}_{x'} - \theta_3\right) = 0 \times 0.5 + 1 \cdot 0 = 0$$

denne kan være fin for å slippe å regne "e^{-x}" hele => $e^{-x' + B\cdot\theta}$

$$Y_3 = \frac{1}{1+e^{-\left(\underbrace{(0 \cdot 0.5)+(1 \cdot 0)}_{x' + B\theta}+(-1)\cdot 0.8\right)}} = \frac{1}{1+e^{-(0.8)}} = \underline{0.6899}$$

HL: 4

$$Y_4 = \text{Sigmoid}\left(x_2 \cdot w_{24} + x_1 \cdot w_{14} - \theta_4\right) = 1 \times 0.9 + 0 \cdot 0 = 0.9$$

$\cancel{= 0.9 >= \theta_4}$

$\cancel{= 0.9 >= -0.1}$

$\cancel{= 1}$

$$Y_4 = \frac{1}{1+e^{-((1 \times 0.9)+(0 \times 0)+(-1)\times(-0.1))}} = \frac{1}{1+e^{-(1)}} \approx \underline{0.7311}$$

$0.9 + 0 + 0.1$

OL: 5

$$Y_5 = \text{Sigmoid}\left(Y_3 \cdot w_{35} + Y_4 \cdot w_{45} - \theta_5\right) = 0.6899 \cdot 0.4 + 0.7311 \cdot (-1.2) = -0.6014$$

$\cancel{= -0.6014 >= \theta_5}$

$\cancel{= -0.6014 >= 0.3}$

$\cancel{= 0}$ (IKKE GTE)

$$Y_5 = \frac{1}{1+e^{-(0.6899 \times 0.4 + 0.7311 \times (-1.2) + (-1) \times 0.3)}} = \frac{1}{1+e^{-(-0.9014)}} = \underline{0.2888}$$

$x' = 0.7311 \cdot 1.1 + 0.6899 \cdot 1.0$

$= 0.8042 + 0.6899 = 1.4941$

OL: 6

$$Y_6 = \text{Sigmoid}\left(\overbrace{Y_4 \cdot w_{46} + Y_3 \cdot w_{36}}^{x'} - \theta\right) = \frac{1}{1+e^{-x' + B\cdot\theta}}$$

$$= \frac{1}{1+e^{-(1.4941 + (-1 \cdot 0.5))}} = \frac{1}{1+e^{-(0.9941)}} \approx \underline{0.7299}$$

STEP 3

The first thing we need to do is calculate the 'error gradient' (derivative of sigmoid).

$\delta_k = Y_k(1-Y_k) \cdot error$

error = $Y_{Desired} - Y_{Predicted}$

$\delta_5 = Y_5(1-Y_5)e_5 = 0.2888(1-0.288) \cdot 4.7112 = \underline{0.9677}$

$e_5 = 5 - 0.2888 = 4.7112$

$\delta_6 = Y_6(1-Y_6)e_6 = 0.7299 \cdot (1-0.7299) \cdot 0.2701 = \underline{0.0532}$

$e_6 = 1 - 0.7299 = 0.2701$

Next up is the actuall weight training / Correction

learning rate $(\alpha) = 0.1$

$\boxed{\begin{array}{l} W_{35} = W_{35} + \Delta W_{35} = 0.4 + 0.0668 = \underline{0.4668} \\ \Delta W_{35} = \alpha \cdot Y_3 \cdot \delta_5 = 0.1 \cdot 0.6899 \cdot 0.9677 = 0.0668 \end{array}}$

$\boxed{\begin{array}{l} W_{45} = W_{45} + \Delta W_{45} = -1.2 + 0.0707 = \underline{-1.1293} \\ \Delta W_{45} = \alpha \cdot Y_4 \cdot \delta_5 = 0.1 \cdot 0.7311 \cdot 0.9677 = 0.0707 \end{array}}$

$\boxed{\begin{array}{l} W_{36} = W_{36} + \Delta W_{36} = 1.0 + 0.0037 = \underline{1.0037} \\ \Delta W_{36} = \alpha Y_3 \cdot \delta_6 = 0.1 \cdot 0.6899 \cdot 0.0532 = 0.0037 \end{array}}$

$\boxed{\begin{array}{l} W_{46} = W_{46} + \Delta W_{46} = 1.1 + 0.0039 = \underline{1.1039} \\ \Delta W_{46} = \alpha \cdot Y_4 \cdot \delta_6 = 0.1 \cdot 0.7311 \cdot 0.0532 = 0.0039 \end{array}}$

$\boxed{\begin{array}{l} \theta_5 = \theta_5 + \Delta\theta_5 = 0.3 + -0.0 \\ \qquad = \underline{0.2032} \\ \Delta\theta_5 = \alpha \cdot BIAS \cdot \delta_5 = \\ \qquad = 0.1 \cdot (-1) \cdot 0.9677 = -0.096 \end{array}}$

$\boxed{\begin{array}{l} \theta_6 = \theta_6 + \Delta\theta_6 \\ \qquad = 0.5 + -0.0053 \\ \qquad = \underline{0.4947} \\ \Delta\theta_6 = \alpha \cdot BIAS \cdot \delta_6 \\ \qquad = 0.1 \cdot (-1) \cdot 0.0532 \\ \qquad = -0.0053 \end{array}}$

Now we are done with output layer incoming weights for this Iteration, Last thing to do, befor this iteration is over, is to upate the weights between Input & Hidden Layers

There are some minor changes to calculating 'error gradient' for hidden layer

The Algo changes from $y_k(1-y_k)e_k$ to $y_j(1-y_j)\sum_{k=1}^{l}\delta_k \cdot w_{jk}$

$$\underline{\delta_3} = y_3(1-y_3)\cdot\sum_{k=1}^{n}\delta_k\, w_{jk} \quad , n = \text{antall outputs nodes}$$

$$= 0.6899\,(1-0.6899)\cdot\left[\delta_5\cdot w_{35} + \delta_6\cdot w_{36}\right]$$

$$\cdot\left((0.9677\cdot 0.4) + (0.0532\cdot 1.0)\right)$$

$$= 0.6899\cdot(0.3101) \cdot 0.4403$$

$$= \underline{0.0942}$$

$$\underline{\delta_4} = y_4(1-y_4)\sum_{k=1}^{n}\delta_k\, w_{jk}$$

$$= 0.7311\cdot(1-0.7311)\cdot\left[\delta_5\cdot w_{45} + \delta_6\cdot w_{46}\right]$$

$$= \cdot\left[(0.9677\cdot(-1.2)) + (0.0532\cdot 1.1)\right]$$

$$= 0.7311\times 0.2689 \times(-1.1027)$$

$$= \underline{-0.2168}$$

$$W_{13} = w_{13} + \Delta W_{13} = 0.5 + 0 = \underline{0.5}$$
$$\Delta W_{13} = \alpha\cdot x_i\cdot\delta_j = \alpha\cdot x_1\cdot\delta_3 = 0.1\times 0\times 0.0942 = 0$$

$$W_{14} = W_{14} + \Delta W_{14} = 0 + 0 = \underline{0}$$
$$\Delta W_{14} = \alpha\cdot x_1\cdot\delta_j = 0.1\times 0\times\ldots = 0$$

$$W_{23} = W_{23} + \Delta W_{23} = 0 + 0.00942 = \underline{0.00942}$$
$$\Delta W_{23} = \alpha\cdot x_2\cdot\delta_3 = 0.1\times 1\times 0.0942 = 0.00942$$

$$W_{24} = W_{24}\cdot\Delta W_{24} = 0.9 + (-0.02168) = \underline{0.8783}$$
$$\Delta W_{24} = \alpha\cdot x_2\cdot\delta_4 = 0.1\times 1\times(-0.2168) = -0.02168$$

$$\underline{\theta_3} = \theta_3 + \Delta\theta_3 = 0.8 + (-0.00942) = \underline{0.7906}$$
$$\Delta\theta_3 = \alpha\cdot B\cdot\delta_3$$
$$= 0.1\cdot(-1)\cdot 0.0942$$
$$= -0.00942$$

$$\underline{\theta_4} = \theta_4 + \Delta\theta_4 = -0.1 + 0.0216 = \underline{0.07832}$$
$$\Delta\theta_4 = \alpha\cdot B\cdot\delta_4$$
$$= 0.1\times(-1)\times(-0.2168)$$
$$= 0.02168$$