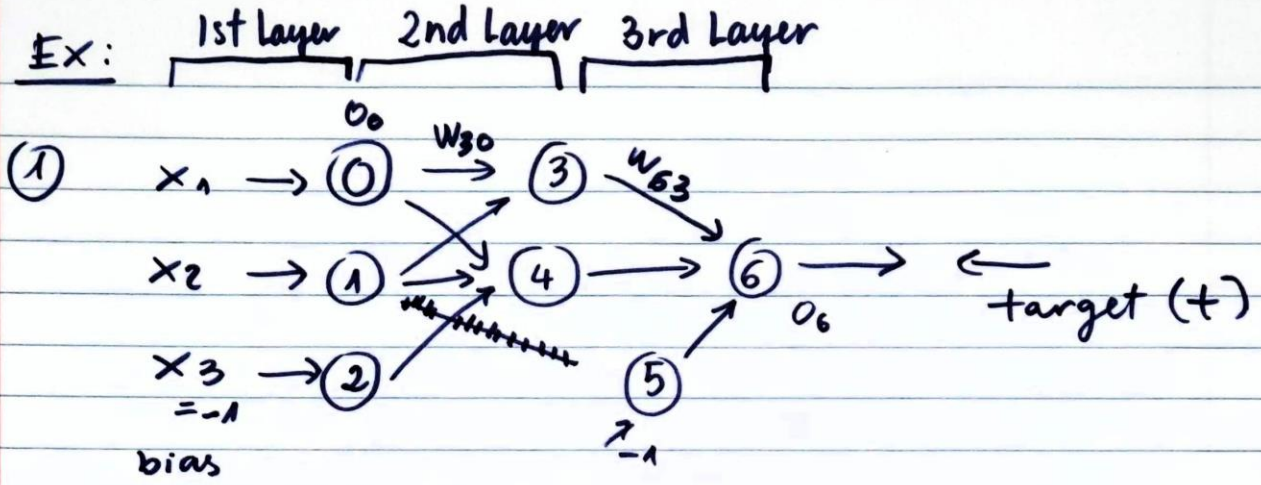
MLP with Backpropagation - ①

Ex:



9 weights variables

② Initializing:

$w_{all} = 0.2$; η : learning rate $= 0.2$

$E_{max} = 0.0001\ (0.01\%)$ $E=0$, $k=1$

② Apply input:

$x^{(1)} = (0.3, 0.4)$, $t(1) = 0.88$
$x^{(2)} = (0.1, 0.6)$, $t(2) = 0.82$
$x^{(3)} = (0.9, 0.4)$, $t(3) = 0.57$

Bias $= -1$
sigmoid

$$f(tot) = \frac{1}{1+e^{-\lambda\, tot}}$$

* if $\lambda \neq 1$ we have to take derivative and find $f'(tot)$

$\lambda = 1$
$\rightarrow f'(tot) = f(tot)(1-f(tot))$

$x_1, x_2, x_3 \rightarrow O_0, O_1, O_2$ does not have weights $\rightarrow$ $O_1 = f(x_1)\ x_1 = 0.3$
$O_2 = x_3 = -1$ $O_1 = x_2 = 0.4$

MLP with Backpropagation-(2)

(3) Forward :

$o_3 = f(tot_3)$

$tot_3 = w_{30} \cdot o_0 + w_{31} o_1 + w_{32} o_2 = 0.2(o_0 + o_1 + o_2)$

$= 0.2(-0.3) = -0.06$

$f(tot_3) = \frac{1}{1 + e^{-(-0.06)}} = \boxed{0.485}$

$o_4 = f(tot_4)$

$tot_4 = w_{40} o_0 + w_{41} o_1 + w_{42} o_2 = 0.2(-0.3) = -0.06$

$\rightarrow f(tot_4) = \boxed{0.485}$

$o_5 = \boxed{-1}$

$o_6 = f(w_{63} \cdot o_3 + w_{64} o_4 + w_{65} o_5)$

$= f(0.2(0.485 \cdot 2 - 1)) = f(-0.006) = 0.4985$

(4) Output error measure

$E = \frac{1}{2}(t - o_6)^2 = \frac{1}{2}(0.88 - 0.4985)^2$

$= \boxed{0.0728}$

(5) error signal

$\delta_6 = f'(tot_6)(t - o_6)$

$= o_6(1 - o_6)(t - o_6)$

$= 0.4985(1 - 0.4985)(0.88 - 0.4985)$

$\rightarrow = \boxed{0.0954}$

MLP with Backpropagation - (2)

(5) Now we go back $\delta \leftarrow \Delta w \leftarrow \delta_6$

- Use signal error δ_6 to update weights in 2nd layer

$$\Delta w_{63} = \eta \delta_6 o_3 = 0.2 . 0.0954 . 0.485 = 0.0093$$

$$w_{63}^{new} = w_{63}^{old} + \Delta w_{63} = 0.2 + 0.0093 = 0.2093$$

$$\Delta w_{64} = 0.0093 \rightarrow w_{64}^{new} = 0.2093$$

$$\Delta w_{65} = -0.01908 \rightarrow w_{65}^{new} = 0.1809$$

- error signal in 2nd layers, δ_3, δ_4

$$\delta_3 = f_3'(tot) \sum_{i=6}^{6} w_{i3} \delta_i$$

$$= o_3 (1 - o_3) w_{63} \delta_6 = 0.485 . (1 - 0.485) . 0.2093 \quad (0.0954 \checkmark)$$

$$= 0.0048$$

* Note: we use w_{63}^{old}

$$\delta_4 = 0.0048$$

(6) Update weights for 2nd layer

$$\begin{bmatrix} \Delta w_{32} \\ \Delta w_{30} \\ \Delta w_{31} \\ \Delta w_{40} \\ \Delta w_{41} \\ \Delta w_{42} \end{bmatrix} = \eta \cdot \begin{bmatrix} \delta_3 \\ \delta_3 \\ \delta_3 \\ \delta_4 \\ \delta_4 \\ \delta_4 \end{bmatrix} \cdot \begin{bmatrix} o_2 \\ o_0 \\ o_1 \\ o_0 \\ o_1 \\ o_2 \end{bmatrix} = \begin{bmatrix} \quad \\ \quad \\ \quad \\ \quad \\ \quad \\ \quad \end{bmatrix} \rightarrow \begin{bmatrix} w_{32} \\ w_{30} \\ w_{31} \\ w_{40} \\ w_{41} \\ w_{42} \end{bmatrix}^{new} = \begin{bmatrix} \quad \\ \quad \\ \quad \\ \quad \\ \quad \\ \quad \end{bmatrix}^{old} + \begin{bmatrix} \quad \\ \quad \\ \quad \\ \quad \\ \quad \\ \quad \end{bmatrix}^{\Delta}$$