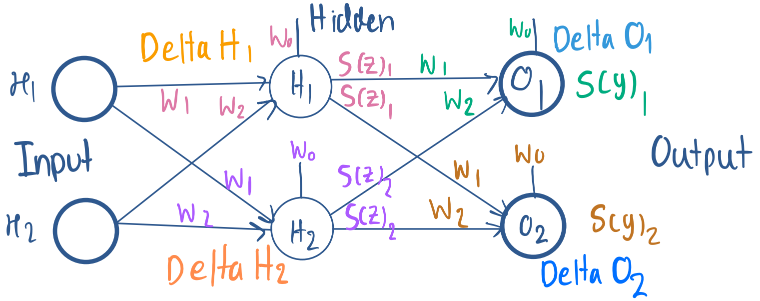
BNN – Backpropagation Neural Network

output 2 ตัว
target 2 ตัว



$\eta = 0.3$

At first iteration W_0, W_1, W_2 have been randomized

Following iterations $W_i \leftarrow W_i + \Delta W_i$

Jutatuch Sathitviriyapun

1 ตลอด

xOR

not Target 1

					H1					H2					O1					O2				
x_0	X1	X2	Target 1	Target 2	W0	W1	W2	Z_1	$S(Z)_1$	W0	W1	W2	Z_2	$S(Z)_2$	W0	W1	W2	Y_1	$S(Y)_1$	W0	W1	W2	Y_2	$S(Y)_2$
1	0	0	0	1	0.0789	0.0636	0.0923	0.0789	0.5197	0.0964	0.0581	0.0983	0.0964	0.5241	0.0609	0.0727	0.0109	0.1044	0.5261	0.0382	0.0932	0.0485	0.1121	0.528
1	0	1	1	0	0.079	0.0636	0.0923	0.1713	0.5427	0.0967	0.0581	0.0983	0.1951	0.5486	0.0215	0.0522	-0.0097	0.0445	0.5111	0.0735	0.1116	0.067	0.1708	0.5426
1	1	0	1	0	0.0784	0.0636	0.0916	0.1419	0.5354	0.096	0.0581	0.0976	0.1541	0.5384	0.0582	0.0721	0.0104	0.1024	0.5256	0.0331	0.0897	0.0449	0.1052	0.5263
1	1	1	0	1	0.0781	0.0633	0.0916	0.2331	0.558	0.0956	0.0577	0.0976	0.2509	0.5624	0.0936	0.0911	0.0295	0.1611	0.5402	-0.0063	0.0686	0.0237	0.0453	0.5113

$Z_1 = W_0 + W_1 x_1 + W_2 x_2$ $\qquad$ $S(Z)_1 = \dfrac{1}{1 + e^{-Z_1}}$

$Z_2 = W_0 + W_1 x_1 + W_2 x_2$ $\qquad$ $S(Z)_2 = \dfrac{1}{1 + e^{-Z_2}}$

$Y_1 = W_0 + S(Z)_1 W_1 + S(Z)_2 W_2$ $\qquad$ $S(Y)_1 = \dfrac{1}{1 + e^{-Y_1}}$

$Y_2 = W_0 + S(Z)_1 W_1 + S(Z)_2 W_2$ $\qquad$ $S(Y)_2 = \dfrac{1}{1 + e^{-Y_2}}$

δ $\qquad$ Δ

η = 0.3							
Delta O1	Delta O2	Delta W0 (O1)	Delta W1 (O1)	Delta W2 (O1)	Delta W0 (O2)	Delta W1 (O2)	Delta W2 (O2)
-0.131159031	0.117632619	-0.039347709	-0.020449524	-0.020621512	0.035289786	0.018340567	0.018494818
0.122156205	-0.134664834	0.036646861	0.019888917	0.020105062	-0.04039945	-0.021925515	-0.022163793

(target − output) $\qquad$ diff sigmoid = output (1 − output)

$\text{Delta } O_1 = (\text{Target1} - S(Y)_1) \times S(Y)_1 \times (1 - S(Y)_1)$

$\text{Delta } O_2 = (\text{Target2} - S(Y)_2) \times S(Y)_2 \times (1 - S(Y)_2)$

$\text{Delta } W_0(O_1) = \eta \times \text{Delta } O_1 \times 1$

$\text{Delta } W_1(O_1) = \eta \times \text{Delta } O_1 \times S(Z_1)$

$\text{Delta } W_2(O_1) = \eta \times \text{Delta } O_1 \times S(Z_2)$

$\text{Delta } W_0(O_2) = \eta \times \text{Delta } O_2 \times 1$

$\text{Delta } W_1(O_2) = \eta \times \text{Delta } O_2 \times S(Z_1)$

$\text{Delta } W_2(O_2) = \eta \times \text{Delta } O_2 \times S(Z_2)$

Delta H1	δ Delta H2	Δ Delta W0 (H1)	Delta W1 (H1)	Delta W2 (H1)	Delta W0 (H2)	Delta W1 (H2)	Delta W2 (H2)
0.000358065	0.001067162	0.00010742	0	0	0.000320149	0	0
-0.002145474	-0.002529187	-0.000643642	0	-0.000643642	-0.000758756	0	-0.000758756

O_1 $\qquad$ O_2

$\text{Delta } H_1 = (\text{Delta } O_1 \times W_1 + \text{Delta } O_2 \times W_1) \times S(Z_1) \times (1 - S(Z_1))$

$\text{Delta } H_2 = (\text{Delta } O_1 \times W_2 + \text{Delta } O_2 \times W_2) \times S(Z_2) \times (1 - S(Z_2))$

เหมือนพุ่งออกจาก H_1/H_2

$\text{Delta } W_0(H1) = \eta \times \text{Delta } H_1 \times 1$

$\text{Delta } W_1(H1) = \eta \times \text{Delta } H_1 \times x_1$

$\text{Delta } W_2(H1) = \eta \times \text{Delta } H_1 \times x_2$

$\text{Delta } W_0(H2) = \eta \times \text{Delta } H_2 \times 1$

$\text{Delta } W_1(H2) = \eta \times \text{Delta } H_2 \times x_1$

$\text{Delta } W_2(H2) = \eta \times \text{Delta } H_2 \times x_2$

https://youtu.be/dQw4w9WgXcQ?si=X5Nj5TTkI2t8lS5h

Jutatuch Sattavirigupan

Backpropagation Algorithm: Operational Summary

Given	A training set $\mathcal{T}$ comprising vectors $X_k \in \mathbb{R}^n$ and desired output vectors $D_k \in \mathbb{R}^p$ and an n–q–p architecture neural network $\mathcal{N}$
	เริ่มต้น random weight ทั้งหมด
Initialize	↬ Randomize weights w^1_{ih} to small values, set $\Delta w^0_{ih} = 0$, $i = 0, \ldots, n$; $h = 1, \ldots, q$ ↬ Randomize weights w^1_{hj} to small values, set $\Delta w^0_{hj} = 0$, $h = 0, \ldots, q$; $j = 1, \ldots, p$ ↬ Set $k = 1$, η, α, and the error tolerance τ as desired.
Iterate	⟳ Repeat { ⇝ Select a training pair $(X_k, D_k) \in \mathcal{T}$ forward pass ⇝ Compute signals on forward pass in the following sequence: $\mathcal{S}(x^k_i) = x^k_i, \quad i = 1, \ldots, n$ $\mathcal{S}(x^k_0) = 1$ $z^k_h = \sum_{i=0}^{n} w^k_{ih} x^k_i, \quad h = 1, \ldots, q$ Hidden $\mathcal{S}(z^k_h) = \frac{1}{1+\exp(-z^k_h)}, \quad h = 1, \ldots, q$ $\mathcal{S}(z^k_0) = 1$ Output $y^k_j = \sum_{h=0}^{q} w^k_{hj} \mathcal{S}(z^k_h), \quad j = 1, \ldots, p$ $\mathcal{S}(y^k_j) = \frac{1}{1+\exp(-y^k_j)}, \quad j = 1, \ldots, p$

Backpropagation Algorithm: Operational Summary(contd.)

$\delta(y) = \frac{1}{1+e^{-y}}$

$\mathcal{S}'(y_j) = \mathcal{S}(y_j)(1-\mathcal{S}(y_j))$

$\boxed{\frac{d\delta(y)}{dy}} = \boxed{\delta(y)(1-\delta(y))}$

target − output

⇝ Compute deltas/errors at output neurons:

Delta Output

$$\delta_j^k = (d_j^k - \mathcal{S}(y_j^k))\mathcal{S}'(y_j^k) \qquad j = 1, \ldots, p$$

$$\Delta w_{hj}^k = \eta \delta_j^k \mathcal{S}(\cancel{x}_h^k) \; z_h \qquad h = 0, \ldots, q; \; j = 1, \ldots, p$$

(learning Rate × delta × input ขาเข้า)

⇝ Compute deltas/errors at hidden neurons:

delta output

Delta Hidden

$$\delta_h^k = \left(\sum_{j=1}^{p} \delta_j^k w_{hj}^k\right) \mathcal{S}'(z_h^k) \qquad h = 1, \ldots, q$$

$$\Delta w_{ih}^k = \eta \delta_h^k x_i^k \qquad i = 0, \ldots, n; \; h = 1, \ldots, q$$

W จาก hidden ไป Output

diff สัญญาณ

q จำนวน hidden node

p จำนวน Output node

⇝ Update weights:

$$w_{hj}^{k+1} = w_{hj}^k + \Delta w_{hj}^k + \alpha \Delta w_{hj}^{k-1} \qquad h = 0, \ldots, q; \; j = 1, \ldots, p$$

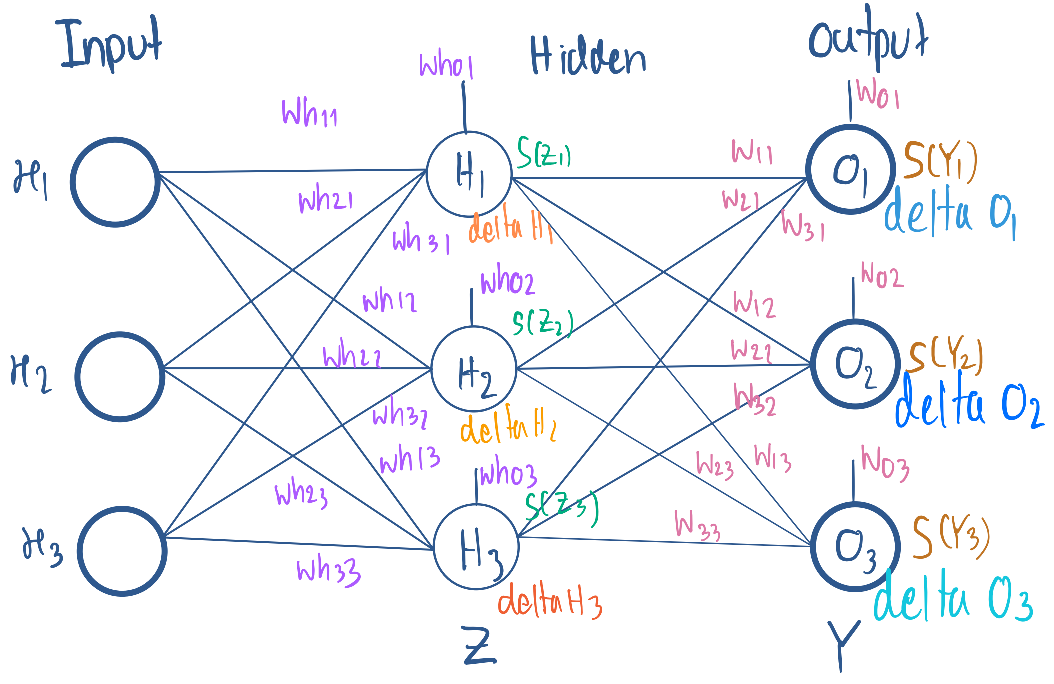
$$w_{ih}^{k+1} = w_{ih}^k + \Delta w_{ih}^k + \underbrace{\alpha \Delta w_{ih}^{k-1}}_{\text{term momentum (ไม่สำคัญ)}} \qquad i = 0, \ldots, n; \; h = 1, \ldots, q$$

⇝ Collect pattern error $\mathcal{E}_k$

} until $(\mathcal{E}_{av} = \frac{1}{Q} \sum_{k=1}^{Q} \mathcal{E}_k < \tau)$

Learning Rate × Delta Output × Input ขาเข้า

Jutatuch Sattaviriyakun



$$Z_1 = wh_{01} + wh_{11} \cdot \mathcal{H}_1 + wh_{21} \cdot \mathcal{H}_2 + wh_{31} \cdot \mathcal{H}_3$$
$$Z_2 = wh_{02} + wh_{12} \cdot \mathcal{H}_1 + wh_{22} \cdot \mathcal{H}_2 + wh_{32} \cdot \mathcal{H}_3$$
$$Z_3 = wh_{03} + wh_{13} \cdot \mathcal{H}_1 + wh_{23} \cdot \mathcal{H}_2 + wh_{33} \cdot \mathcal{H}_3$$

$$S(Z_1) = \frac{1}{1+e^{-Z_1}} \; ; \; S(Z_2) = \frac{1}{1+e^{-Z_2}} \; ; \; S(Z_3) = \frac{1}{1+e^{-Z_3}}$$

$$Y_1 = W_{01} + S(Z_1) W_{11} + S(Z_2) W_{21} + S(Z_3) W_{31}$$
$$Y_2 = W_{02} + S(Z_1) W_{12} + S(Z_2) W_{22} + S(Z_3) W_{32}$$
$$Y_3 = W_{03} + S(Z_1) W_{13} + S(Z_2) W_{23} + S(Z_3) W_{33}$$

$$S(Y_1) = \frac{1}{1+e^{-Y_1}} \; ; \; S(Y_2) = \frac{1}{1+e^{-Y_2}} \; ; \; S(Y_3) = \frac{1}{1+e^{-Y_3}}$$

δ

$$\text{delta } O_1 = (\text{target1} - S(Y_1)) \times S(Y_1) \times (1 - S(Y_1))$$
$$\text{delta } O_2 = (\text{target2} - S(Y_2)) \times S(Y_2) \times (1 - S(Y_2))$$
$$\text{delta } O_3 = (\text{target3} - S(Y_3)) \times S(Y_3) \times (1 - S(Y_3))$$

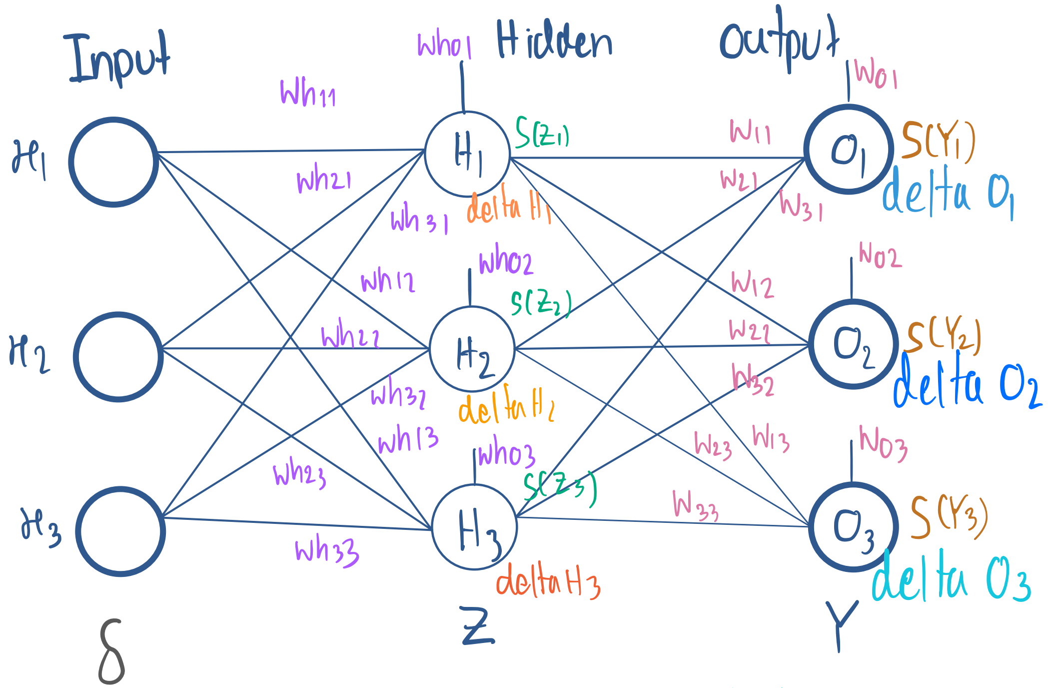
Δ

$$\text{delta } W_{01}(O_1) = \eta \times \text{delta } O_1 \times 1$$
$$\text{delta } W_{11}(O_1) = \eta \times \text{delta } O_1 \times S(Z_1)$$
$$\text{delta } W_{21}(O_1) = \eta \times \text{delta } O_1 \times S(Z_2)$$
$$\text{delta } W_{31}(O_1) = \eta \times \text{delta } O_1 \times S(Z_3)$$

$$\text{delta } W_{02}(O_2) = \eta \times \text{delta } O_2 \times 1$$
$$\text{delta } W_{12}(O_2) = \eta \times \text{delta } O_2 \times S(Z_1)$$
$$\text{delta } W_{22}(O_2) = \eta \times \text{delta } O_2 \times S(Z_2)$$
$$\text{delta } W_{32}(O_2) = \eta \times \text{delta } O_2 \times S(Z_3)$$

$$\text{delta } W_{03}(O_3) = \eta \times \text{delta } O_3 \times 1$$
$$\text{delta } W_{13}(O_3) = \eta \times \text{delta } O_3 \times S(Z_1)$$
$$\text{delta } W_{23}(O_3) = \eta \times \text{delta } O_3 \times S(Z_2)$$
$$\text{delta } W_{33}(O_3) = \eta \times \text{delta } O_3 \times S(Z_3)$$

Jutatuch Sattaviriyapun



$$\text{delta } H_1 = (\text{delta } O_1 \times W_{11} + \text{delta } O_2 \times W_{12} + \text{delta } O_3 \times W_{13}) \times S(Z_1) \times (1 - S(Z_1))$$

$$\text{delta } H_2 = (\text{delta } O_1 \times W_{21} + \text{delta } O_2 \times W_{22} + \text{delta } O_3 \times W_{23}) \times S(Z_2) \times (1 - S(Z_2))$$

$$\text{delta } H_3 = (\text{delta } O_1 \times W_{31} + \text{delta } O_2 \times W_{32} + \text{delta } O_3 \times W_{33}) \times S(Z_3) \times (1 - S(Z_3))$$

Δ

$$\text{delta } wh01 (H_1) = \eta \times \text{delta } H_1 \times 1$$
$$\text{delta } wh11 (H_1) = \eta \times \text{delta } H_1 \times x_1$$
$$\text{delta } wh21 (H_1) = \eta \times \text{delta } H_1 \times x_2$$
$$\text{delta } wh31 (H_1) = \eta \times \text{delta } H_1 \times x_3$$

$$\text{delta } wh02 (H_2) = \eta \times \text{delta } H_2 \times 1$$
$$\text{delta } wh12 (H_2) = \eta \times \text{delta } H_2 \times x_1$$
$$\text{delta } wh22 (H_2) = \eta \times \text{delta } H_2 \times x_2$$
$$\text{delta } wh32 (H_2) = \eta \times \text{delta } H_2 \times x_3$$

$$\text{delta } wh03 (H_3) = \eta \times \text{delta } H_3 \times 1$$
$$\text{delta } wh13 (H_3) = \eta \times \text{delta } H_3 \times x_1$$
$$\text{delta } wh23 (H_3) = \eta \times \text{delta } H_3 \times x_2$$
$$\text{delta } wh33 (H_3) = \eta \times \text{delta } H_3 \times x_3$$

columns

x_0 | x_1 | x_2 | x_3 | Target1 | Target2 | Target3 | (H_1) wh01, wh11, wh21, wh31, Z_1, $S(Z_1)$ | (H_2) wh02, wh12, wh22, wh32, Z_2, $S(Z_2)$ | (H_3) wh03, wh13, wh23, wh33, Z_3, $S(Z_3)$

(O_1) W01, W11, W21, W31, Y_1, $S(Y_1)$ | (O_2) W02, W12, W22, W32, Y_2, $S(Y_2)$ | (O_3) W03, W13, W23, W33, Y_3, $S(Y_3)$ | delta O_1 | delta O_2 | delta O_3

delta $W_{01}(O_1)$, delta $W_{11}(O_1)$, delta $W_{21}(O_1)$, delta $W_{31}(O_1)$ | delta $W_{02}(O_2)$, delta $W_{12}(O_2)$, delta $W_{22}(O_2)$, delta $W_{32}(O_2)$

delta $W_{03}(O_3)$, delta $W_{13}(O_3)$, delta $W_{23}(O_3)$, delta $W_{33}(O_3)$ | delta H_1 | delta H_2 | delta H_3

delta wh01 (H_1), delta wh11 (H_1), delta wh21 (H_1), delta wh31 (H_1) | delta wh02 (H_2), delta wh12 (H_2), delta wh22 (H_2), delta wh32 (H_2)

delta wh03 (H_3), delta wh13 (H_3), delta wh23 (H_3), delta wh33 (H_3)

Jutatuch Sathanirugapun