x_1 W1 0.15 $z_1 | h_1$ w5 0.4

W2=0.2

W3 0.25

x_2 W4 0.3 $z_2 | h_2$ W6 0.45

학습률 : 0.5
실제값 : 0.3

1. 순전파.

° 입력값 : $x_1 = 0.1$, $x_2 = 0.2$.

° 초기 가중치 : $w_1 = 0.15$, $w_2 = 0.2$, $w_3 = 0.25$,
         $w_4 = 0.3$, $w_5 = 0.4$, $w_6 = 0.45$,

은닉층계산.

1. 은닉층의 첫 번째 뉴런 h_1:

$$z_1 = (w_1 \cdot x_1) + (w_3 \cdot x_2) = \{(0.15) \cdot (0.1)\} + \{(0.25) \cdot (0.2)\}$$
$$= 0.015 + 0.05 = 0.065.$$

$$h_1 = sigmoid(z_1) = \frac{1}{1 + e^{-0.065}} \approx 0.5162.$$

2. 은닉층의 두번째 뉴런 h_2:

$$z_2 = (w_2 \cdot x_1) + (w_4 \cdot x_2) = \{(0.2) \cdot (0.1)\} + \{(0.3) \cdot (0.2)\}$$
$$= 0.02 + 0.06 = 0.08.$$

$$h_2 = sigmoid(z_2) = \frac{1}{1 + e^{-0.08}} \approx 0.5199$$

출력층 계산

출력 뉴런 y:

$$y_{input} = \{(w_5 \cdot h_1)\} + \{(w_6 \cdot h_2)\} = \{(0.4)(0.5162)\} + \{(0.45) \cdot (0.5199)\}$$
$$\approx 0.2065 + 0.2339 = 0.4404.$$

$$y = sigmoid(y_{input}) = \frac{1}{1 + e^{-0.4404}} \approx 0.6084.$$

2. 오차계산.

실제값이 0.3이므로, 오차는 다음과 같다.

$$오차 \approx \frac{1}{2}\{(0.3) - (0.6084)\}^2 \approx 0.0475.$$

3. 역전파.

출력층의 가중치 W_5, W_6 업데이트
출력층에서의 2차 기울기 구하기.

1. $\delta_y = (0.3 - 0.6084) \cdot y \cdot (1-y) = -0.3084 \cdot 0.6084 \cdot (1 - 0.6084) \approx -0.0737.$

출력층 오차

2. W_5에 대한 기울기.

$\Delta W_5 = \delta_y \cdot h_1 = (-0.0737) \cdot (0.5162) \approx -0.0380.$

3. W_6에 대한 기울기

$\Delta W_6 = \delta_y \cdot h_2 = (-0.0737) \cdot (0.5199) \approx -0.0382.$

4. 업데이트된 W_5와 W_6:

$W_5 = W_5 + (학습률 \cdot \Delta W_5) = 0.4 +\{(0.5 \times -0.0380)\} = (0.4) - (0.0190) = 0.3810.$

$W_6 = W_6 + (학습률 \cdot \Delta W_6) = 0.45 +\{(0.5) \cdot (-0.0382)\} = 0.45 - (0.0191) = 0.4309.$

은닉층의 가중치 W_1, W_2, W_3, W_4 업데이트.

1. 은닉층의 첫 번째 뉴런 h_1에 대한 오차:

$\delta_{h_1} = \delta_y \cdot W_5 \cdot h_1 \cdot (1- h_1) = (-0.0737) \cdot (0.4) \cdot (0.5162) \cdot (1- 0.5162) \approx -0.0073.$

2. 은닉층의 두 번째 뉴런 h_2에 대한 오차:

$\delta_{h_2} = \delta_y \cdot W_6 \cdot h_2 \cdot (1- h_2) = (-0.0737) \cdot (0.45) \cdot (0.5199) \cdot (1- 0.5199) \approx 0.0081.$

3. W_1에 대한 기울기.

$\Delta W_1 = \delta_{h_1} \cdot X_1 = (-0.0073) \cdot (0.1) = -0.0007.$

4. W_2에 대한 기울기.

$\Delta W_2 = \delta_{h_2} \cdot X_1 = (-0.0081) \cdot (0.1) = -0.0008.$

5. W_3에 대한 기울기.

$\Delta W_3 = \delta_{h_1} \cdot X_2 = (-0.0073) \cdot (0.2) = -0.0015.$

6. W_4에 대한 기울기.

$\Delta W_4 = \delta_{h_2} \cdot X_2 = (-0.0081) \cdot (0.2) = (-0.0016)$

7. 업데이트된 W_1, W_2, W_3, W_4:

$W_1 = W_1 + (학습률 \cdot \Delta W_1) = 0.15 +\{(0.5) \cdot (-0.0007)\} = 0.1496.$

$W_2 = W_2 + (학습률 \cdot \Delta W_2) = 0.2 +\{(0.5) \cdot (-0.0008)\} = 0.1996$

$W_3 = W_3 + (학습률 \cdot \Delta W_3) = 0.25 +\{(0.5) \cdot (-0.0015)\} = 0.2493.$

$W_4 = W_4 + (학습률 \cdot \Delta W_4) = 0.3 + \{(0.5)(-0.0016)\} = 0.2992.$

최종결과.

○ 업데이트 된 가중치:

$W_1 = 0.1496$
$W_2 = 0.1996$
$W_3 = 0.2493$
$W_4 = 0.2992$
$W_5 = 0.3810$
$W_6 = 0.4309.$

출력 뉴런.

$y_{input} = (W_5 \cdot h_1) + (W_6 \cdot h_2)$
$= \{(0.4) \cdot 0.5162\}$
$+ \{(0.45) \cdot (0.5199)\}$
$\approx 0.2065 + 0.2339 = 0.4404$

$y = sigmoid(y_{input})$
$= \dfrac{1}{1+ e^{-0.4404}} \approx 0.6084.$