

# Neural Network Basic Assignment

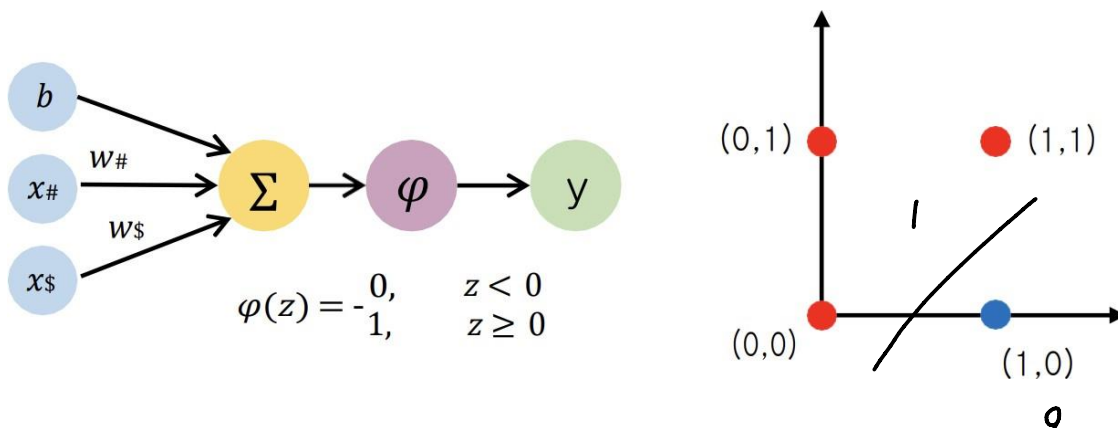
이름: 고영

1. Sigmoid Function을  $z$ 에 대해 미분하세요.

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

$$\begin{aligned} \frac{d}{dz} \sigma(z) &= \frac{d}{dz} (1 + e^{-z})^{-1} = - (1 + e^{-z})^{-2} \cdot (-e^{-z}) \\ &= \frac{e^{-z}}{(1 + e^{-z})^2} \end{aligned}$$

2. 다음과 같은 구조의 Perceptron과 ● (=1), ● (=0)을 평면좌표상에 나타낸 그림이 있습니다.



2-1. ●, ●을 분류하는 임의의  $b, w$ 를 선정하고 분류해보세요.

$$w = \begin{pmatrix} w_{\#} \\ w_{\$} \end{pmatrix} \quad z = b + x_{\#} \cdot w_{\#} + x_{\$} \cdot w_{\$} \quad \longrightarrow \quad \sigma(z)$$

$(0, 0) \rightarrow 3 + 1 \cdot 0 + 2 \cdot 0 = 3$	$\sigma(3) \approx 1$
$(0, 1) \rightarrow 3 + 1 \cdot 0 + 2 \cdot 1 = 5$	$\sigma(5) \approx 1$
$(1, 0) \rightarrow 3 + 1 \cdot 1 + 2 \cdot 0 = 4$	$\sigma(4) \approx 1$
$(1, 1) \rightarrow 3 + 1 \cdot 1 + 2 \cdot 1 = 6$	$\sigma(6) \approx 1$

임의의 정한  $b, w_{\#}, w_{\$}$

$b: 3$   
 $w_{\#}: 2$   
 $w_{\$}: 1$

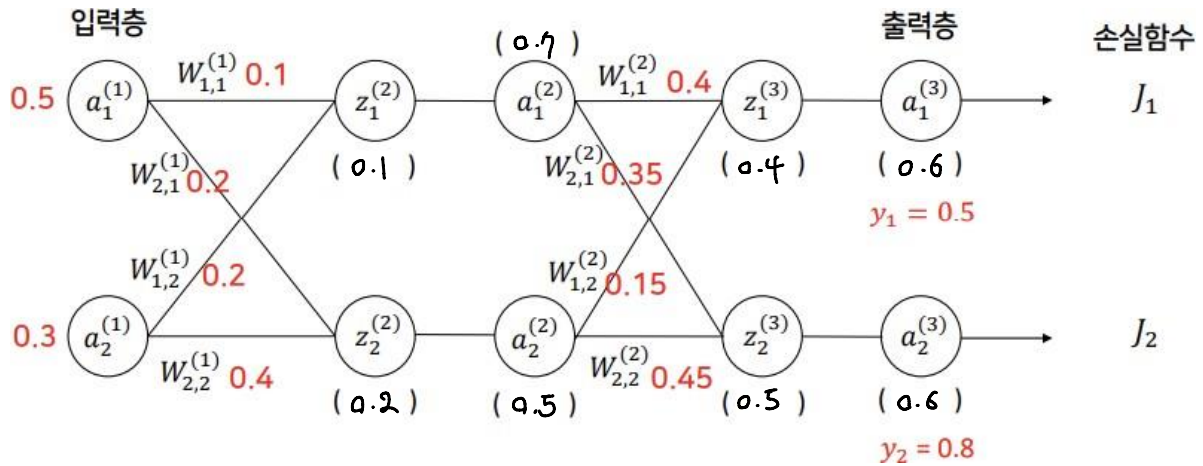
2-2. Perceptron 학습 규칙에 따라 임의의 학습률을 정하고  $b, w$ 를 1회 업데이트 해주세요.

임의의 학습률:  $0.1$

$$\begin{aligned} b &\leftarrow b + 0.1 (0 - 1) \cdot 1 = 3 - 0.1 = 2.9 \\ w_{\#} &\leftarrow w_{\#} + 0.1 (0 - 1) \cdot 1 = 2 - 0.1 = 1.9 \\ w_{\$} &\leftarrow w_{\$} + 0.1 (0 - 1) \cdot 0 = 2 \end{aligned}$$

$x_{\#}$	$x_{\$}$	$o$	$\hat{y}$
0	0	/	/
0	1	/	/
1	0	/	0
1	1	/	/

3. 다음과 같이 입력과 가중치가 주어진 퍼셉트론이 있을 때, 아래의 물음에 답해주세요. 모든 문제는 풀이과정을 자세하게 적어주세요! (3-3까지 있습니다.)



- 3-1. FeedForward가 일어날 때, 각 노드가 갖는 값을 빈칸에 써주세요. 단, 활성화함수는 sigmoid 함수입니다. (모든 계산의 결과는 소수점 셋째자리에서 반올림하여 둘째자리까지만 써주세요.)  $\phi(z) = \frac{1}{1+e^{-z}}$

$$z_1^{(2)} = w_{1,1}^{(1)} \cdot a_1^{(1)} + w_{1,2}^{(1)} \cdot a_2^{(1)} = 0.1 \times 0.5 + 0.2 \times 0.3 = 0.1$$

$$z_2^{(2)} = w_{2,1}^{(1)} \cdot a_1^{(1)} + w_{2,2}^{(1)} \cdot a_2^{(1)} = 0.2 \times 0.5 + 0.4 \times 0.3 = 0.2$$

$$a_1^{(2)} = (1 + e^{-0.1})^{-1} = 0.7 \quad a_2^{(2)} = (1 + e^{-0.2})^{-1} = 0.5$$

$$z_1^{(3)} = w_{1,1}^{(2)} \cdot a_1^{(2)} + w_{1,2}^{(2)} \cdot a_2^{(2)} = 0.4 \times 0.7 + 0.15 \times 0.5 = 0.4$$

$$z_2^{(3)} = w_{2,1}^{(2)} \cdot a_1^{(2)} + w_{2,2}^{(2)} \cdot a_2^{(2)} = 0.35 \times 0.7 + 0.45 \times 0.5 = 0.5$$

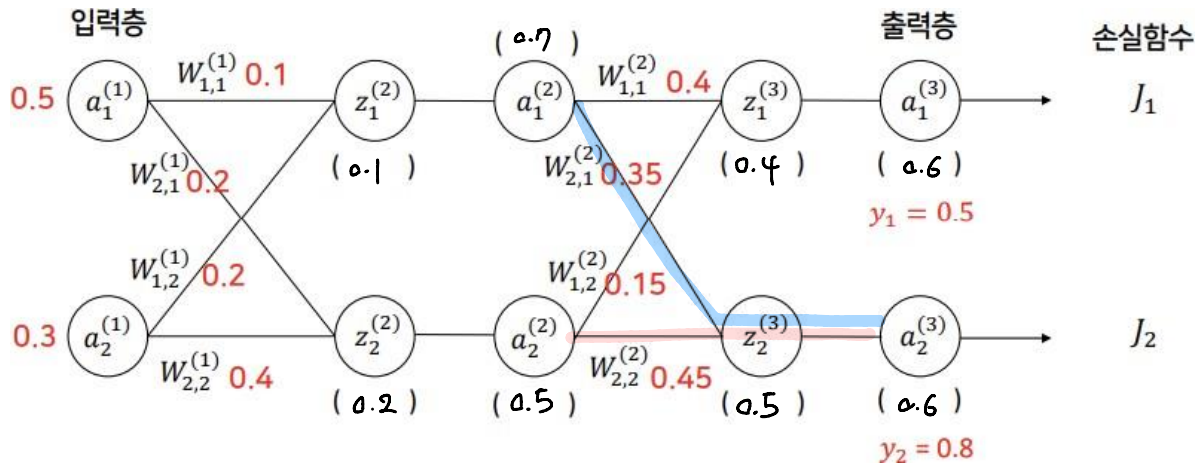
$$a_1^{(3)} = (1 + e^{-0.4})^{-1} = 0.6, \quad a_2^{(3)} = (1 + e^{-0.5})^{-1} = 0.6$$

- 3-2. 3-1에서 구한 값을 이용하여 손실함수  $J_1$ 과  $J_2$ 의 값을 구해주세요. ( $J_1$ 과  $J_2$ 는 반올림하지 말고 써주세요.)

$$MSE = \frac{1}{2N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

$$J_1 = \frac{1}{2} (a_1^{(3)} - y_1)^2 = \frac{1}{2} (0.6 - 0.5)^2 = 0.005$$

$$J_2 = \frac{1}{2} (a_2^{(3)} - y_2)^2 = \frac{1}{2} (0.6 - 0.8)^2 = 0.02$$



- 3-3. 위에서 구한 값을 토대로, BackPropagation이 일어날 때  $w_{2,2}^{(2)}$ 와  $w_{2,1}^{(1)}$ 의 조정된 값을 구해주세요.  
단, learning rate는 0.1입니다. (계산 과정에서 소수점 넷째자리에서 반올림하여 셋째자리까지만 써주시고, 마지막 결과인  $w_{2,1}^{(1)}$ 과  $w_{2,2}^{(2)}$ 의 값만 반올림하지 말고 써주세요.)

$$\begin{aligned}
 w_{2,2}^{(2)} &= w_{2,2}^{(2)} - \frac{\partial J_{total}}{\partial w_{2,2}^{(2)}} = w_{2,2}^{(2)} - \frac{\partial J_{total}}{\partial z_2^{(3)}} \times \frac{\partial z_2^{(3)}}{\partial w_{2,2}^{(2)}} \\
 &= w_{2,2}^{(2)} - \frac{\partial J_2}{\partial z_2^{(3)}} \times \frac{\partial z_2^{(3)}}{\partial w_{2,2}^{(2)}} \quad (z_2^{(3)} | J_{total} \approx J_2) \\
 &= w_{2,2}^{(2)} - \left( a_2^{(3)} \times (1 - a_2^{(3)}) \right) \times a_2^{(2)} \\
 &= 0.45 - (0.6 + (1 - 0.6)) \times 0.5 \\
 &= 0.33
 \end{aligned}$$

$$\begin{aligned}
 w_{2,1}^{(2)} &= w_{2,1}^{(2)} - \frac{\partial J_{total}}{\partial w_{2,1}^{(2)}} = w_{2,1}^{(2)} - \frac{\partial J_{total}}{\partial z_2^{(3)}} \times \frac{\partial z_2^{(3)}}{\partial w_{2,1}^{(2)}} \\
 &= w_{2,1}^{(2)} - \frac{\partial J_2}{\partial z_2^{(3)}} \times \frac{\partial z_2^{(3)}}{\partial w_{2,1}^{(2)}} \quad (z_2^{(3)} | J_{total} \approx J_2) \\
 &= w_{2,1}^{(2)} - \left( a_2^{(3)} \times (1 - a_2^{(3)}) \right) \times a_1^{(2)} \\
 &= 0.35 - (0.6 + (1 - 0.6)) \times 0.7 \\
 &= 0.182
 \end{aligned}$$