투빅스 16기 정규세션 6주차

Neural Network Basic Assignment

이름: 주지훈

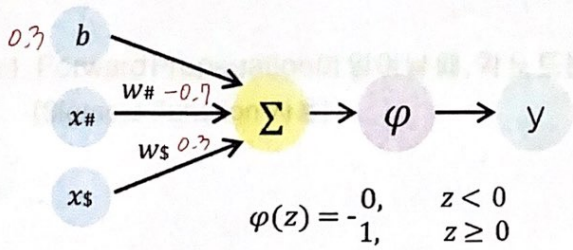
1. Sigmoid Function을 z에 대해 미분하세요.

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

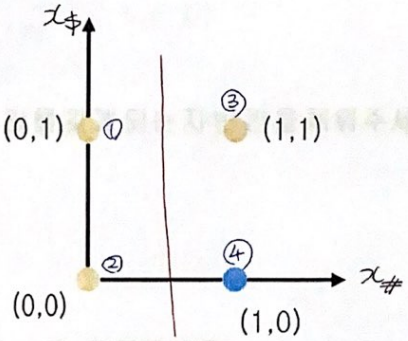
$$\frac{d}{dx}(1+e^{-z})^{-1} = -(1+e^{-z})^{-2} \times (-e^{-z}) = \frac{e^{-z}}{(1+e^{-z})^2} = \frac{1+e^{-z}-1}{(1+e^{-z})^2} = \frac{1+e^{-z}}{(1+e^{-z})^2} - \frac{1}{(1+e^{-z})^2}$$

$$= \frac{1}{1+e^{-z}} - \frac{1}{(1+e^{-z})^2} = \frac{1}{1+e^{-z}}\left(1 - \frac{1}{1+e^{-z}}\right) = \sigma(z) \times (1-\sigma(z))$$

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



$$\varphi(z) = \begin{cases} 0, & z < 0 \\ 1, & z \geq 0 \end{cases}$$



2-1. 🟠, 🔵를 분류하는 임의의 b, w를 선정하고 분류해보세요.

① $0\times(-0.7) + 1\times 0.3 + 0.3 = 0.6 \rightarrow y = 1$

② $0\times(-0.7) + 0\times 0.3 + 0.3 = 0.3 \rightarrow y = 1$

③ $0\times(-0.7) + 1\times 0.3 + 0.3 = -0.1 \rightarrow y = 0$

④ $0\times(-0.7) + 0\times 0.3 + 0.3 = -0.4 \rightarrow y = 0$

$x_\#$	$x_\$$	y
0	1	1
0	0	1
1	1	0
0	0	0

→ 잘못분류됨!

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

$x_0 = 1$, $\eta = 0.05$ 라 설정.

① $b = 0.3 + 0.05(1-1)\times 1 = 0.3$

$W_\# = -0.7 + 0.05(1-1)\times 0 = -0.7$

$W_\$ = 0.3 + 0.05(1-1)\times 1 = 0.3$

$\therefore -0.7\times 0 + 0.3\times 1 + 0.3 = 0.6$

$y = 1$

② $b = 0.3 + 0.05(1-1)\times 1 = 0.3$

$W_\# = -0.7 + 0.05(1-1)\times 0 = -0.7$

$W_\$ = 0.3 + 0.05(1-1)\times 0 = 0.3$

$\therefore -0.7\times 0 + 0.3\times 0 + 0.3 = 0.3$

$y = 1$

③ $b = 0.3 + 0.05(1-0)\times 1 = 0.35$

$W_\# = -0.7 + 0.05(1-0)\times 1 = -0.65$

$W_\$ = 0.3 + 0.05(1-0)\times 1 = 0.35$

$\therefore -0.65\times 1 + 0.35\times 1 + 0.35 = 0.05$

$y = 1$

④ $b = 0.3 + 0.05(0-0)\times 1 = 0.3$

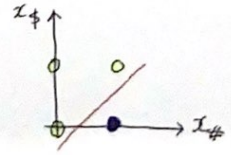
$W_\# = -0.7 + 0.05(0-0)\times 1 = -0.7$

$W_\$ = 0.3 + 0.05(0-0)\times 0 = 0.3$

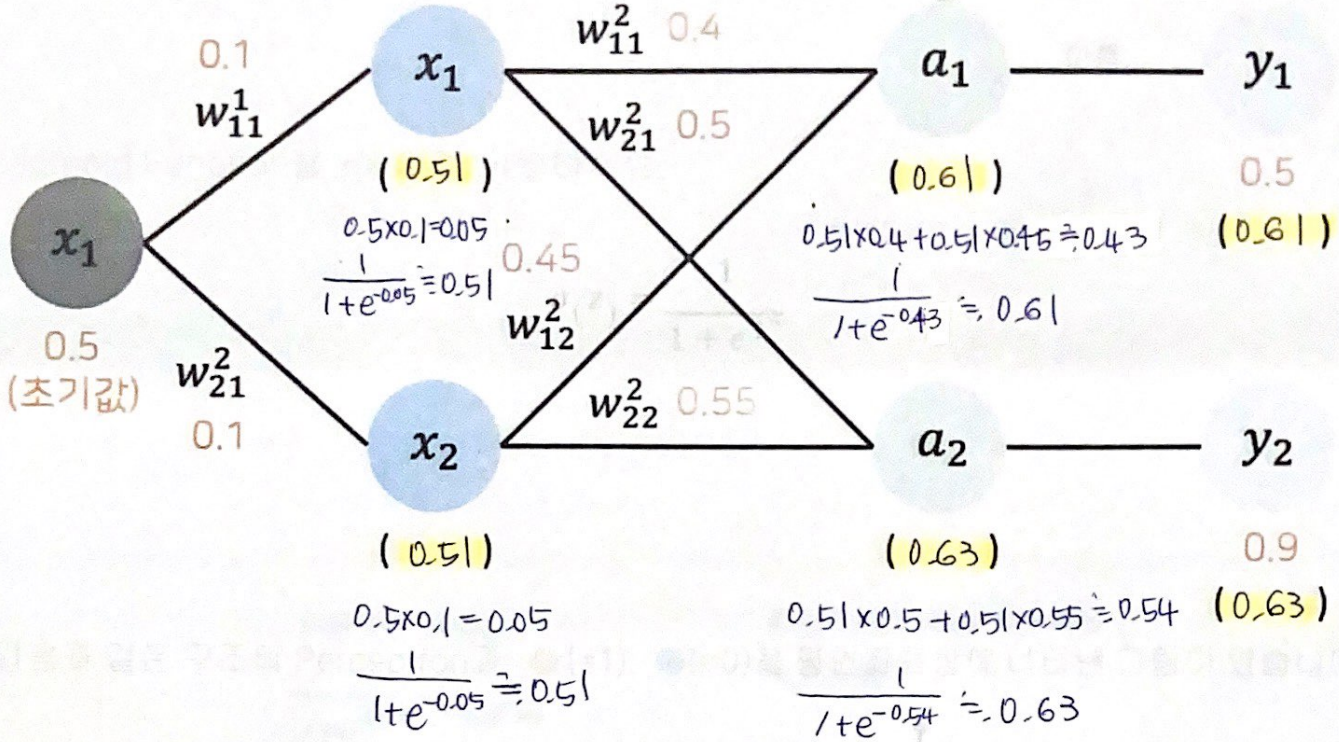
$\therefore -0.7\times 1 + 0.3\times 0 + 0.3 = -0.4$, $y = 0$

$x_\#$	$x_\$$	y
0	1	1
0	0	1
1	1	1
0	0	0

→ 제대로 분류됨!



3. 다음과 같은 구조와 초기값을 가진 Multilayer Perceptron이 있습니다.



$0.5 \times 0.1 = 0.05$, $\frac{1}{1+e^{-0.05}} \fallingdotseq 0.51$

$0.51 \times 0.4 + 0.51 \times 0.45 \fallingdotseq 0.43$, $\frac{1}{1+e^{-0.43}} \fallingdotseq 0.61$

$0.5 \times 0.1 = 0.05$, $\frac{1}{1+e^{-0.05}} \fallingdotseq 0.51$

$0.51 \times 0.5 + 0.51 \times 0.55 \fallingdotseq 0.54$, $\frac{1}{1+e^{-0.54}} \fallingdotseq 0.63$

3-1. Forward Propagation이 일어날 때, 각 노드는 어떤 값을 갖게 되는지 빈 칸을 채워주세요 (Sigmoid Function 사용)

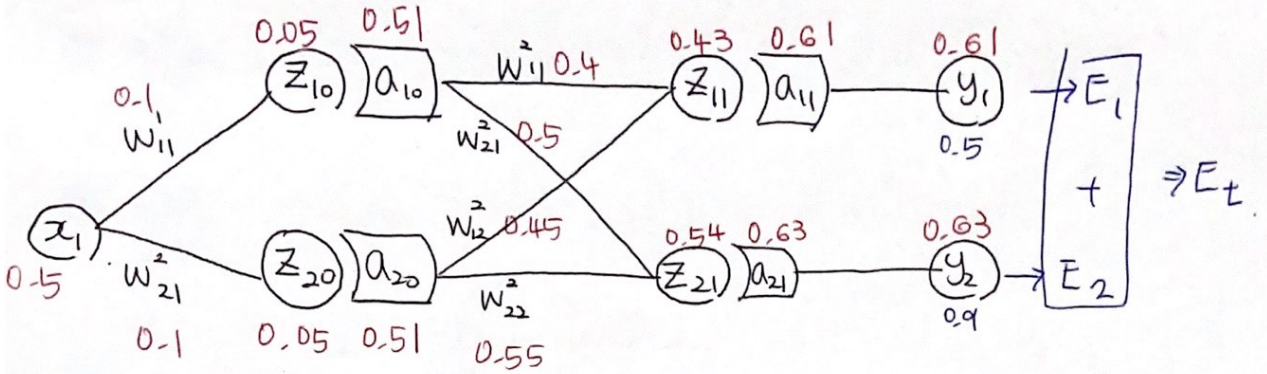
3-2. output layer에 있는 노드들의 Mean Squared Error을 구해주세요.

$\hat{y}_1 = 0.61$, $\hat{y}_2 = 0.63$

$$MSE = \frac{1}{2}\sum_i^2 \frac{1}{2}(y_i - \hat{y}_i)^2 = \frac{1}{2}\left\{\frac{1}{2}(0.5-0.61)^2 + \frac{1}{2}(0.9-0.63)^2\right\}$$

$\fallingdotseq 0.02$

3-3. 3-2에서 구한 답을 토대로, Back Propagation이 일어날 때 가중치 w_{11}^1과 w_{11}^2의 조정된 값을 구해주세요. (learning rate : 0.4)



(1) w_{11}^1

$$\frac{\partial E_t}{\partial w_{11}^1} = \left(\underbrace{\frac{\partial E_1}{\partial a_{10}}}_{①} + \underbrace{\frac{\partial E_2}{\partial a_{10}}}_{②}\right) \cdot \underbrace{\frac{\partial a_{10}}{\partial z_{10}}}_{③} \cdot \underbrace{\frac{\partial z_{10}}{\partial w_{11}^1}}_{④}$$

① $\frac{\partial E_1}{\partial a_{10}} = \frac{\partial E_1}{\partial a_{11}} \cdot \frac{\partial a_{11}}{\partial z_{11}} \cdot \frac{\partial z_{11}}{\partial a_{10}}$

$MSE = \frac{1}{2}\{\underbrace{\frac{1}{2}(y_1 - a_{11})^2}_{E_1} + \underbrace{\frac{1}{2}(y_2 - a_{21})^2}_{E_2}\}$, $\frac{\partial E_1}{\partial a_{11}} = \frac{1}{2} \times 2 \times (y_1 - a_{11}) \times (-1) = 0.11$ (0.5, 0.61)

$\frac{\partial a_{11}}{\partial z_{11}} = \sigma'(z_{11}) = \sigma(z_{11}) \times (1 - \sigma(z_{11})) = 0.61 \times (1 - 0.61) = 0.2379 \fallingdotseq 0.24$

$z_{11} = a_{10} \times w_{11}^2 + a_{20} \times w_{12}^2$, $\frac{\partial z_{11}}{\partial a_{10}} = w_{11}^2 = 0.4$

$\therefore \frac{\partial E_1}{\partial a_{10}} = 0.11 \times 0.24 \times 0.4 = 0.01056 \fallingdotseq \boxed{0.012}$

② $\frac{\partial E_2}{\partial a_{10}} = \frac{\partial E_2}{\partial a_{21}} \cdot \frac{\partial a_{21}}{\partial z_{21}} \cdot \frac{\partial z_{21}}{\partial a_{10}}$

$\frac{\partial E_2}{\partial a_{21}} = \frac{1}{2} \times 2 \times (y_2 - a_{21}) \times (-1) = -(0.9 - 0.63) = -0.27$

$\frac{\partial a_{21}}{\partial z_{21}} = \sigma_{(z_{21})} \times (1 - \sigma_{(z_{21})}) = 0.63 \times (1 - 0.63) = 0.2331 \fallingdotseq 0.23$

$\frac{\partial z_{21}}{\partial a_{10}} = w_{21}^2 = 0.5$

$\therefore \frac{\partial E_2}{\partial a_{10}} = (-0.27) \times 0.23 \times 0.5 = -0.03105 \fallingdotseq \boxed{-0.031}$

③ $\frac{\partial a_{10}}{\partial z_{10}} = \sigma(z_{10}) \times (1-\sigma(z_{10})) = 0.51 \times (1-0.51) = 0.2499 \fallingdotseq 0.25$

④ $\frac{\partial z_{10}}{\partial w_{11}^{1}}$

$z_{10} = x_1 \times w_{11}^{1}$. $\frac{\partial z_{10}}{\partial w_{11}^{1}} = x_1 = 0.5$

$\therefore \frac{\partial E_t}{\partial w_{11}^{1}} = (0.012-0.031) \times 0.25 \times 0.5 = -0.002375 \fallingdotseq -0.0024$

가중치 조정 : $w_i \leftarrow w_i + \eta(y-o)x_i$

$w_{11}^{1} = w_{11}^{1} + 0.4 \times \frac{\partial E_t}{\partial w_{11}^{1}} = 0.1 + 0.4 \times (-0.0024) \fallingdotseq 0.099$

(2) w_{11}^2

$$\frac{\partial E_1}{\partial w_{11}^2} = \underbrace{\frac{\partial E_1}{\partial a_{11}}}_{①} \cdot \underbrace{\frac{\partial a_{11}}{\partial z_{11}}}_{②} \cdot \underbrace{\frac{\partial z_{11}}{\partial w_{11}^2}}_{③}$$

① $\frac{\partial E_1}{\partial a_{11}} = 0.11$ (앞에서 구함)

② $\frac{\partial a_{11}}{\partial z_{11}} = 0.24$

③ $\frac{\partial z_{11}}{\partial w_{11}^2}$

$z_{11} = a_{10} \times w_{11}^2 + a_{20} \times w_{12}^2$ $\quad \frac{\partial z_{11}}{\partial w_{11}^2} = a_{10} = 0.51$

$\therefore \frac{\partial E_1}{\partial w_{11}^2} = 0.11 \times 0.24 \times 0.51 = 0.013464 \fallingdotseq 0.013$

가중치 교정

$w_{11}^2 = w_{11}^2 + 0.4 \times \frac{\partial E_1}{\partial w_{11}^2} = 0.4 + 0.4 \times 0.013 = 0.4052$

답 : $W_{11}^{1} = 0.099$, $W_{11}^{2} = 0.4052$.