투빅스 16기 정규세션 6주차

Neural Network Basic Assignment

이름: 김건우

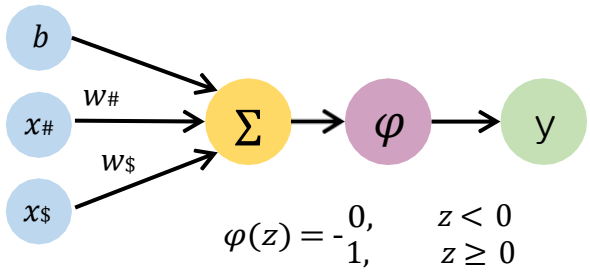
1. Sigmoid Function을 z에 대해 미분하세요.

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

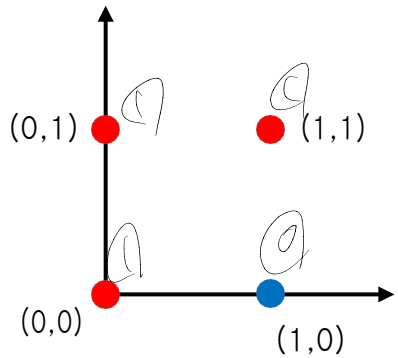
$$\sigma'(z) = \frac{e^{-z}}{(1+e^{-z})^2} = \frac{1+e^{-z}-1}{(1+e^{-z})^2} = \frac{1}{(1+e^{-z})} - \frac{1}{(1+e^{-z})^2} = \frac{1}{1+e^{-z}} \times \left(1 - \frac{1}{(1+e^{-z})}\right)$$

$$\star\ \sigma(z) \times (1-\sigma(z))$$

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.

$w_1 = 0.3$

$w_2 = 0.7$

$b = -0.4$

x_1	x_2	S	y	true
0	0	-0.4	0	1
0	1	0.3	1	1
1	0	-0.1	0	0
1	1	0.6	1	1

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

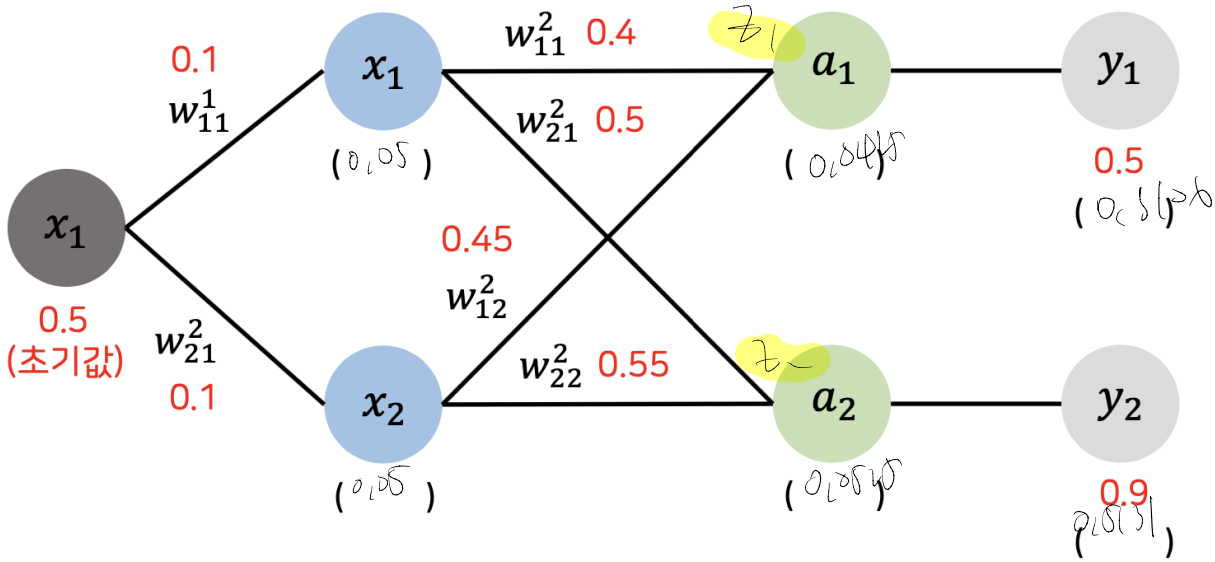
$\eta = 0.1$

$w_1 \leftarrow w_1 + 0.1(0-1) \times 0$ $\quad w_1 = 0.2$

$w_2 \leftarrow w_2 + 0.1(0-1) \times 0$ $\quad w_2 = 0.6$

$b \leftarrow b + 0.1(0-1) \times 1$ $\quad b = -0.5$

3. 다음과 같은 구조와 초기값을 가진 Multilayer Perceptron이 있습니다.



$z_1 = 0.05 \times 0.4 + 0.05 \times 0.45 = 0.0425$

$z_2 = 0.05 \times 0.5 + 0.05 \times 0.55 = 0.0525$

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

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

$a_1 = \sigma(z_1) = 0.5106$

$a_2 = \sigma(z_2) = 0.5131$

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

$MSE = \frac{1}{2}\left((0.3 - 0.5106)^2 + (0.9 - 0.5131)^2\right)$

$= 0.0975$

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)

→ 위에 ...

수고하셨습니다.

w_{11}^1

$$\frac{\partial E}{\partial w_{11}^1} = \frac{\partial E}{\partial x_1}\frac{\partial x_1}{\partial w_{11}^1} = \left(\frac{\partial E_1}{\partial x_1} + \frac{\partial E_2}{\partial x_1}\right)\frac{\partial x_1}{\partial w_{11}^1}$$

$$\therefore \frac{\partial E_1}{\partial x_1} = \frac{\partial E_1}{\partial a_1}\frac{\partial a_1}{\partial z_1}\frac{\partial z_1}{\partial x_1},$$

$$= \frac{1}{2}(0.3106 - 0.5)^2 \times \left(\frac{1}{1+e^{-0.0425}}\right) \times \left(1 - \frac{1}{1+e^{-0.0425}}\right) \times 0.4$$

$$= 5.615 \times 10^{-6} \quad \text{——— (1)}$$

$$\therefore \frac{\partial E_2}{\partial x_1} = \frac{\partial E_2}{\partial a_2}\frac{\partial a_2}{\partial z_2}\frac{\partial z_2}{\partial x_1}$$

$$= \frac{1}{2}(0.5131 - 0.9)^2 \times \left(\frac{1}{1+e^{-0.0525}}\right) \times \left(1 - \frac{1}{1+e^{-0.0525}}\right) \times 0.5$$

$$= 0.0093 \quad \text{——— (2)}$$

$$(1) + (2) \times 0.1 = 0.00093$$

$$w_{11}^1 = 0.1 - 0.4 \times 0.00093 = 0.0996$$

W_{21}^2

$$\frac{\partial E}{\partial W_{21}^2} = \frac{\partial E}{\partial x_2} \frac{\partial x_2}{\partial W_{21}^2} = \left(\frac{\partial E_1}{\partial x_2} + \frac{\partial E_2}{\partial x_2}\right) \frac{\partial x_2}{\partial W_{21}^2}$$

$$\frac{\partial E_1}{\partial x_2} = \frac{\partial E_1}{\partial a_2} \frac{\partial a_2}{\partial z_2} \frac{\partial z_2}{\partial x_2},$$

$$= \frac{1}{2}(0.3106 - 0.5)^2 \times \left(\frac{1}{1+e^{-0.0425}}\right) \times \left(1 - \frac{1}{1+e^{-0.0425}}\right) \times 0.45$$

$$= 6.317 \times 10^{-6} \quad \text{——— (1)}$$

$$\frac{\partial E_2}{\partial x_1} = \frac{\partial E_2}{\partial a_2} \frac{\partial a_2}{\partial z_2} \frac{\partial z_2}{\partial x_1}$$

$$= \frac{1}{2}(0.5131 - 0.9)^2 \times \left(\frac{1}{1+e^{-0.0525}}\right) \times \left(1 - \frac{1}{1+e^{-0.0525}}\right) \times 0.55$$

$$= 0.0103 \quad \text{——— (2)}$$

$$(1) + (2) \times 0.1 = 0.00103$$

$$W_{21}^2 = 0.1 - 0.4 \times 0.00103 = 0.0996$$