투빅스 15기 정규세션 6주차

Neural Network Basic Assignment 1

이름: 김 태 희

1. Sigmoid Function을 z에 대해 미분하세요.

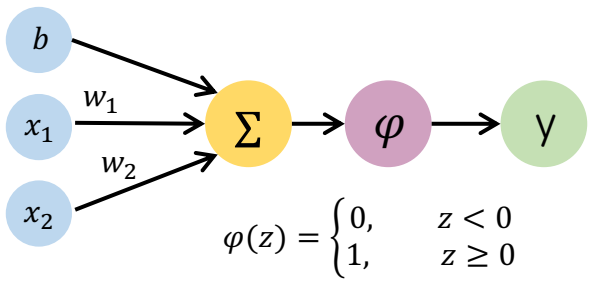
$$\sigma(z) = \frac{1}{1+e^{-z}} = ((1+e^{-z})^{-1})$$

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

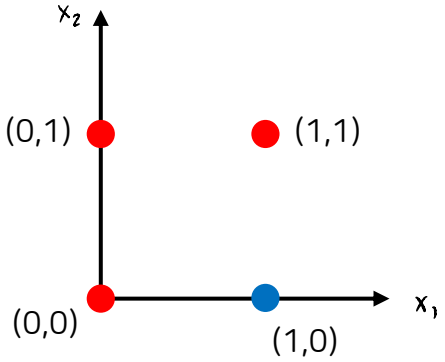
$$= \frac{(-1)(-e^{-z})}{(1+e^{-z})^2} = \frac{1+e^{-z}-1}{(1+e^{-z})^2} = \frac{1}{(1+e^{-z})}\left(1-\frac{1}{1+e^{-z}}\right)$$

$$= \sigma(1-\sigma(z))$$

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.

$w_0 = 1.0$, $w_1 = 1.0$, $w_2 = -1.5$ $b = 1 = x_0$

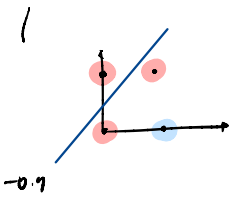
$z = w_0 + w_1x_1 + w_2x_2$

$z_{(0,0)} = w_0 = 1.0$ $\quad \varphi(z_{(0,0)}) = 1$

$z_{(0,1)} = w_0 + w_2x_2 = -0.5$ $\quad \varphi(z_{(0,1)}) = 0$

$z_{(1,0)} = w_0 + w_1x_1 = 2.0$ $\quad \varphi(z_{(1,0)}) = 1$

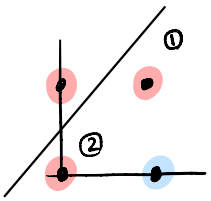
$z_{(1,1)} = w_0 + w_1x_1 + w_2x_2 = 0.5$ $\quad \varphi(z_{(1,1)}) = 1$



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

$\eta = 0.05$ $\quad W_i \leftarrow W_i + \eta(y-0)x_i$

$x_0 = b = 1$



① $w_0 + \eta(y-0)x_0 = w_0 + 0.05(0-1)(1) = w_0 - 0.05$

$w_1 + \eta(y-0)x_1 = w_1 + 0.05(0-1)(1) = w_1 - 0.05$

$w_2 + \eta(y-0)x_2 = w_2 + 0.05(0-1)(1) = w_2 - 0.05$

② $w_0 + 0.05(0-1)(1) = w_0 - 0.05$

$w_1 + 0.05(0-1)(0) = w_1$

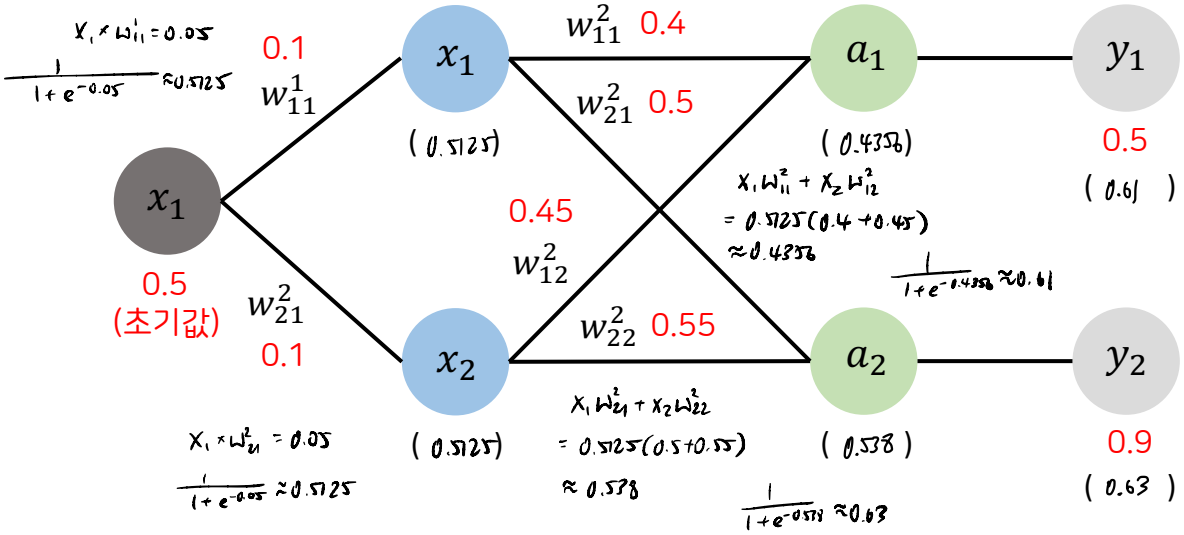
$w_2 + 0.05(0-1)(0) = w_2$

$\therefore w_0 = 1.0 - 0.1 = 0.9$

$w_1 = 1.0 - 0.05 = 0.095$

$w_2 = -1.5 - 0.05 = -1.55$

3. 다음과 같은 구조와 초기값을 가진 Multilayer Perceptron이 있습니다. bias = 0



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.

$$MSE = \frac{1}{n}\sum_{n}\frac{1}{2}(y_i - \hat{y}_i)^2$$
$$= \frac{1}{2}\left\{\frac{1}{2}(0.5-0.648)^2 + \frac{1}{2}(0.9-0.6525)^2\right\}$$
$$= \frac{1}{2}(0.04158)$$
$$\approx \underbrace{0.02079}_{E_{total}}$$

3-3. 3-2에서 구한 답을 토대로, Back Propagation이 일어날 때 가중치 w_{11}^1과 w_{11}^2의 조정된 값을 구해주세요. (학습률 $\eta = 0.5$)

① output → hidden

w_{11}^2, w_{21}^2, w_{12}^2, w_{22}^2

$$\frac{\partial E_{tot}}{\partial w_{11}^1} = \frac{\partial E_{tot}}{\partial w_{12}^2} = -(0.5-0.61)\,0.61\,(1-0.61)\,0.5125 = 0.013411\,6125$$

$$w_{11}^2 = 0.4 - 0.5 \times 0.013411\,6125 = 0.393294$$

$$w_{12}^1 = 0.45 - 0.5 \times 0.013411\,6125 = 0.443294$$

$$\frac{\partial E_{tot}}{\partial w_{21}^2} = \frac{\partial E_{tot}}{\partial w_{22}^2} = -(0.9-0.63)\,0.63\,(1-0.63)\,0.5125 = -0.032255\,2125$$

$$w_{21}^2 = 0.5 - 0.5\,(-0.032255\,2125) = 0.516\,1276$$

$$w_{22}^2 = 0.55 - 0.5\,(-0.032255\,2125) = 0.291\,4936$$

② hidden → input

$$\frac{\partial E_{tot}}{\partial h_1} = -(0.5-0.61)\,0.61\,(1-0.61)\,0.4 = 0.0104696$$

$$\frac{\partial E_{o2}}{\partial h_1} = -(0.9-0.63)\,0.63\,(1-0.63)\,0.5 = -0.0314685$$

$$\frac{\partial E_{tot}}{\partial w_{11}^1} = -0.021 \times 0.5125\,(1-0.5125)\,0.5 = -0.0002623$$

$$w_{11}^1 = 0.1 - 0.5(-0.0002623) = 0.101311$$

$$\frac{\partial E_{tot}}{\partial h_2} = -(0.5-0.61)\,0.61\,(1-0.61)\,0.45 = 0.011796$$

$$\frac{\partial E_{o2}}{\partial h_2} = -(0.9-0.63)\,0.63\,(1-0.63)\,0.55 = -0.0346$$

$$\frac{\partial E_{tot}}{\partial w_{21}^1} = -0.022824 \times 0.5125\,(1-0.5125)\,0.5 = -0.00285$$

$$w_{21}^1 = 0.1 - 0.5\,(-0.00285) = 0.101425$$

수고하셨습니다.

$\therefore w_{11}^1 = 0.101311$, $w_{11}^2 = 0.393254$