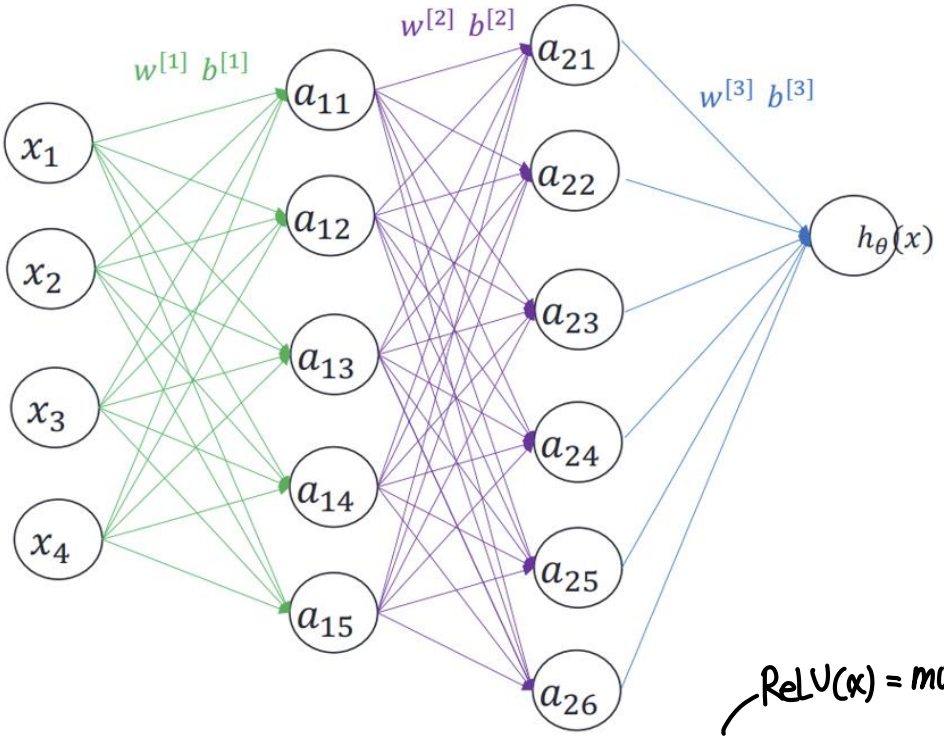
ToBig's 22기 정규세션 4주차

Neural Network 과제

이름: 조하늘



Q1. 이 네트워크를 $w^{[l]}$, $b^{[l]}$, 그리고 활성화함수로 표현해주세요. (ReLU를 활성화함수로 사용하며 마지막 층에서는 사용하지 않음.)

$$a_1 = ReLU(w^{[1]}x + b^{[1]})$$

$$a_2 = ReLU(w^{[2]}a_1 + b^{[2]})$$

$$h(x) = w^{[3]}a_2 + b^{[3]}$$

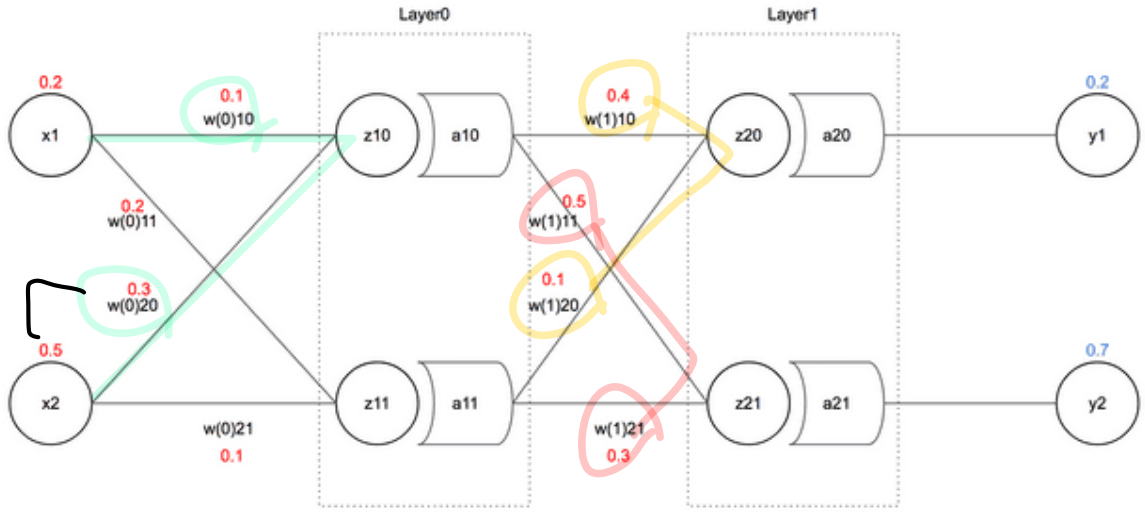
Q2. 이 네트워크를 구성하고 있는 layer 개수와 hidden layer 개수, 그리고 파라미터의 총개수를 각각 구해주세요.

- layer : 3층
- hidden layer : 2층

⇒ Total parameter : 68

hidden \ layer	1	2	3	total
1	20	30	6	56
2	5	6	1	12
				68

다음과 같이 입력과 가중치가 주어진 퍼셉트론이 있을 때, 아래의 물음에 답해주세요. 모든 문제는 풀이과정을 자세하게 적어주세요! (Q3, Q4)



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

Q3. 활성화 함수로 시그모이드(σ)를 사용하고 손실 함수로 평균 제곱 오차를 사용할 때, z, a, 그리고 loss 를 구해주세요.

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

$z_{10} = w_{(0)}^{10} x_1 + w_{(0)}^{20} x_2 = 0.1 \times 0.2 + 0.3 \times 0.5 = 0.17$ (0.02, 0.15)

$z_{11} = w_{(0)}^{11} x_1 + w_{(0)}^{21} x_2 = 0.2 \times 0.2 + 0.1 \times 0.5 = 0.09$ (0.04, 0.05)

$a_{10} = \sigma(z_{10}) = \frac{1}{1+e^{-0.17}} = \frac{1}{1.8437} = 0.542$ ($e^{-0.17} \approx 0.8437$)

$a_{11} = \sigma(z_{11}) = \frac{1}{1+e^{-0.09}} = 0.522$

$z_{20} = w_1^{10} \times a_{10} + w_1^{20} \times a_{11} = 0.4 \times 0.542 + 0.1 \times 0.522 = 0.269$

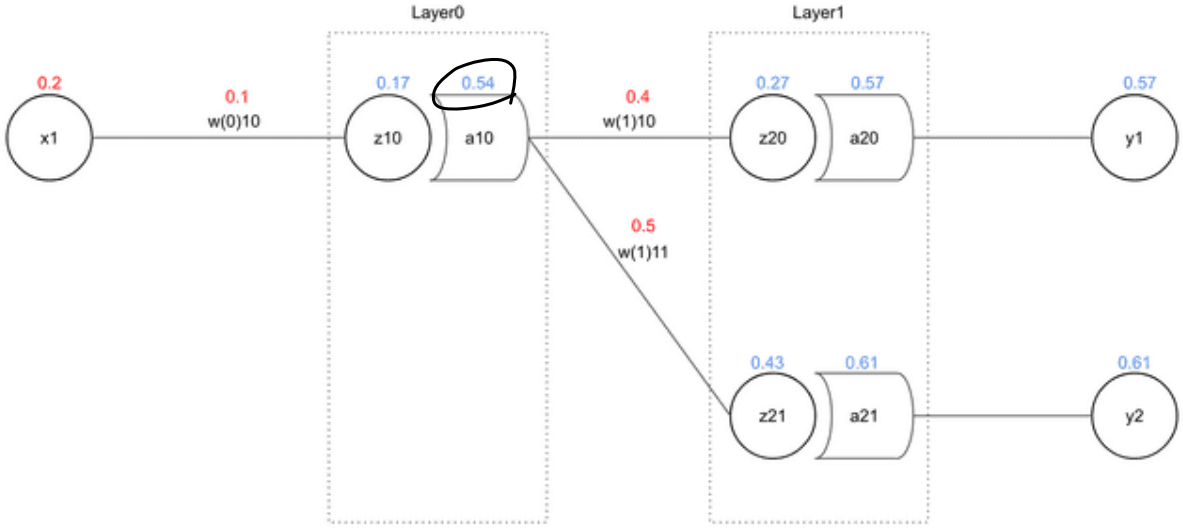
$z_{21} = w_{(1)}^{11} \times a_{10} + w_{(1)}^{21} \times a_{11} = 0.5 \times 0.542 + 0.3 \times 0.522 = 0.428$

$a_{20} = \sigma(z_{20}) = \frac{1}{1+e^{-0.269}} = 0.567$

$a_{21} = \sigma(z_{21}) = \frac{1}{1+e^{-0.428}} = 0.605$

$loss = \frac{1}{2}\left[(0.567 - 0.2)^2 + (0.605 - 0.7)^2\right] = \frac{1}{2}(0.135 + 0.009) = 0.072$

$(0.567-0.2)^2 \approx 0.135$, $(0.605-0.7)^2 \approx 0.009$



Q4. w^1_{10}과 w^0_{10}을 역전파(backpropagation) 기법을 사용하여 갱신하세요

X : Input
A : sigmoid
o : output
Y : Target
E : error

$w_3 = w_{(1)}{}^{10}$
$w_4 = w_{(1)}{}^{11}$
$o_1 = a_{20}$
$o_2 = a_{21}$
$z_3 = z_{20}$
$z_4 = z_{21}$

$\therefore \frac{\partial a_{20}}{\partial z_{20}}$

$* E_{total} = \frac{1}{2}\left[(target_{a_{10}} - Y_1)^2 + (target_{a_{21}} - Y_2)^2\right]$ *target 1, target 2 모두 1/2 가짐

$= \frac{1}{2}\left[(1-0.57)^2 + (1-0.61)^2\right] = 0.169$

$0.43 = 0.185 + 0.152 = 0.337$

$\frac{\partial E_{total}}{\partial a_{20}} = 2 \times \frac{1}{2}(target_{a_{20}} - Y_1)(-1) = -0.43$ || $\frac{\partial E_{total}}{\partial a_{21}} = 2 \times \frac{1}{2}(target_{a_{21}} - Y_2)(-1) = -0.39$

$\frac{\partial Y_1}{\partial z_{20}} = 0.57 \times 0.43 = 0.245$

$\frac{\partial Y_2}{\partial z_{21}} = 0.61 \times 0.39 = 0.238$

$\frac{\partial E_{total}}{\partial w_{(1)}{}^{10}} = -0.43 \times 0.245 \times 0.54 = -0.057$

$\frac{\partial E_{total}}{\partial w_{(1)}{}^{11}} = -0.39 \times 0.238 \times 0.54 = -0.05$

$\frac{\partial z_{10}}{\partial w_{(0)}{}^{10}} = 0.2$

$\frac{\partial a_{10}}{\partial z_{10}} = 0.54 \times 0.46 = 0.248$

$\frac{\partial E_{total}}{\partial z_{20}} = \frac{\partial Loss}{\partial a_{20}} \times \frac{\partial a_{20}}{\partial z_{20}} = -0.43 \times 0.245 = -0.105$

$\frac{\partial E_{total}}{\partial z_{21}} = \frac{\partial Loss}{\partial a_{21}} \times \frac{\partial a_{21}}{\partial z_{21}} = -0.39 \times 0.238 = -0.093$

$\frac{\partial E_{total}}{\partial w_{(0)}{}^{10}} = \{(-0.105 \times 0.4) + (-0.093 \times 0.5)\} \times 0.2 \times 0.248$

$-0.042 - 0.0465 = 0.005$

$= 0.000248$

$\frac{\partial E_{total}}{\partial w_{(1)}{}^{10}} =$

$* \varepsilon = 0.1$ 가정

$\therefore w_{10}{}^{(1)} = w_{10}{}^{(1)} - \varepsilon \times \frac{\partial E_{total}}{\partial w_{(1)}{}^{10}} = 0.4 - 0.1 \times (-0.057) = 0.4057$

$w_{10}{}^{(0)} = w_{10}{}^{(0)} - \varepsilon \times \frac{\partial E_{total}}{\partial w_{(0)}{}^{10}} = 0.1 - 0.1 \times 0.000248 = 0.0999$

0.000248