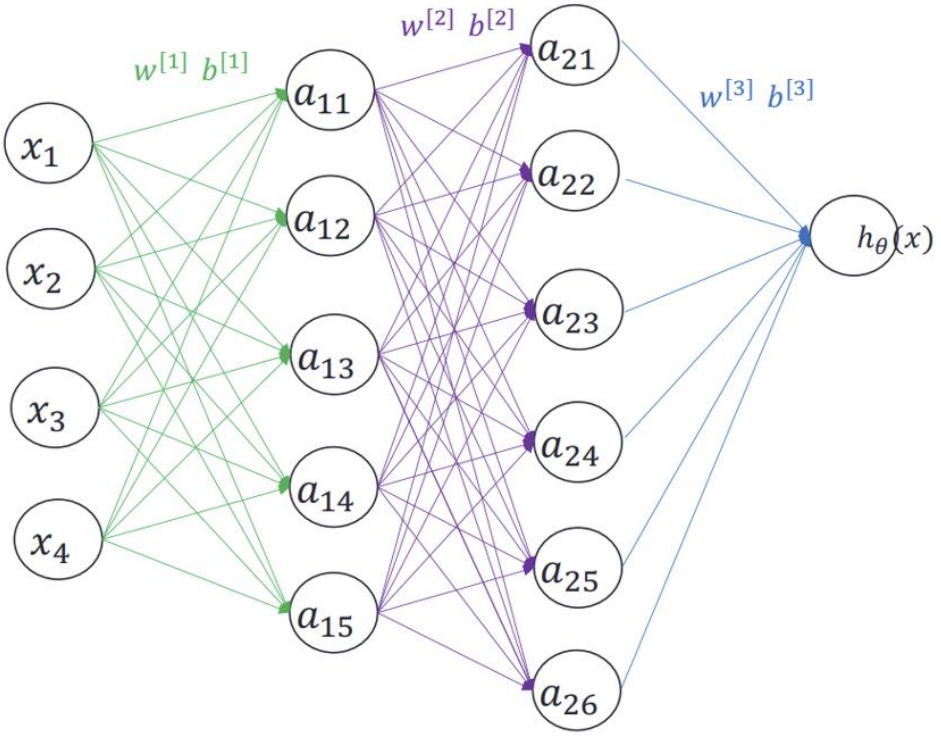
ToBig's 22기 정규세션 4주차

Neural Network 과제

이름: 진민탄



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

$z^{[1]} = W^{[1]} \cdot x + b^{[1]}$, $a^{[1]} = ReLU(z^{[1]})$

$z^{[2]} = W^{[2]} \cdot a^{[1]} + b^{[2]}$, $a^{[2]} = ReLU(z^{[2]})$

$h_\theta(x) = W^{[3]} \cdot a^{[2]} + b^{[3]}$

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

첫번째 레이어

W : 4×5 = 20

b : 5

두번째 레이어

W : 5×6 = 30

b : 6

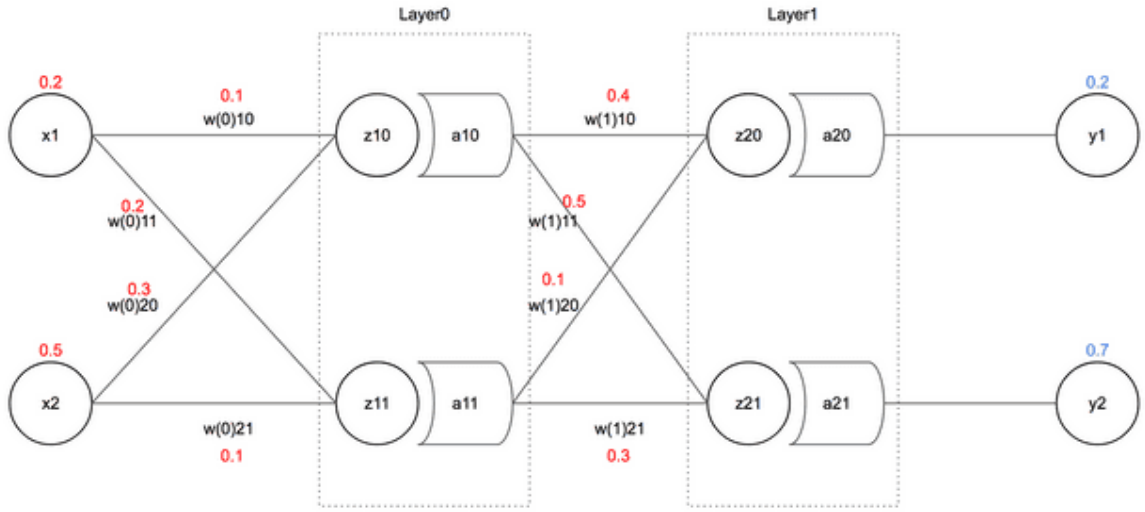
세번째 레이어

W : 6×1 = 6

b : 1

총 파라미터 수 : 68

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



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

$z_{10} = w_{10}x_1 + w_{20}x_2 = 0.2\times0.2 + 0.1\times0.5 = 0.09$

$z_{11} = w_{11}x_1 + w_{21}x_2 = 0.1\times0.2 + 0.2\times0.5 = 0.12$

$a_{10} = \frac{1}{1+e^{-0.09}} \approx 0.5225$

$a_{11} = \frac{1}{1+e^{-0.12}} \approx 0.5299$

$z_{20} = w_{10}a_{10} + w_{20}a_{11} = 0.4\times0.5225 + 0.1\times0.5299 \approx 0.26199$

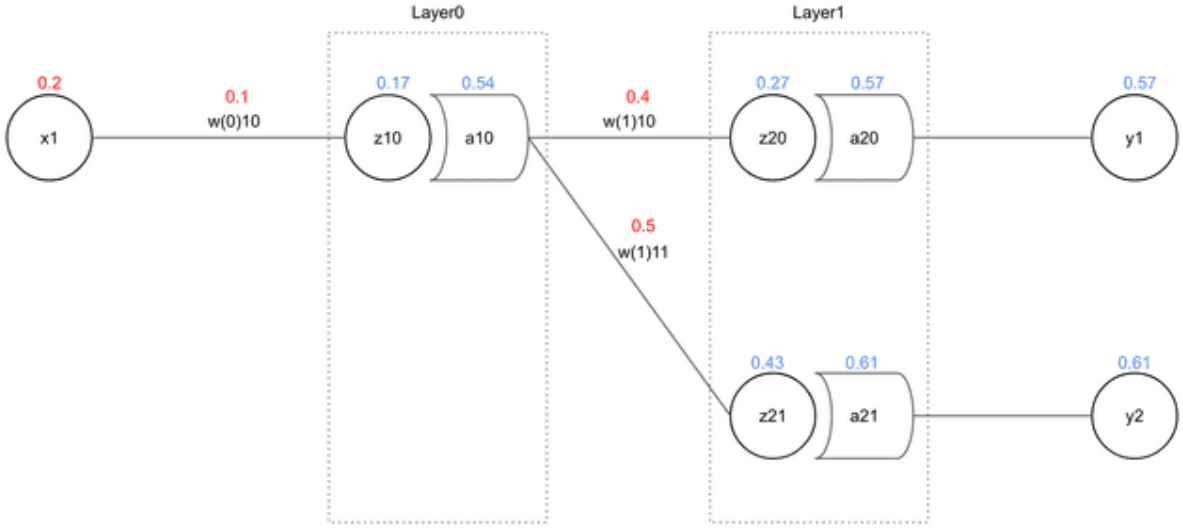
$z_{21} = w_{11}a_{10} + w_{21}a_{11} = 0.5\times0.5225 + 0.3\times0.5299 \approx 0.42022$

$a_{20} = \frac{1}{1+e^{-0.26199}} \approx 0.5651$

$a_{21} = \frac{1}{1+e^{-0.42022}} \approx 0.6035$

$MSE = \frac{1}{2}[(a_{20}-y_1)^2 + (a_{21}-y_2)^2] = \frac{1}{2}[(0.5651-0.2)^2 + (0.6035-0.7)^2]$

≈ 0.07128



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

$y_{1\,target} = 1$, $y_{2\,target} = 0$, 학습률 $\varepsilon = 0.1$, 손실 함수는 MSE라 한다면

$$Loss = \frac{1}{2}\left((y_{1target} - y_1)^2 + (y_{2target} - y_2)^2\right) \approx 0.2785$$

$$\frac{\partial Loss}{\partial y_1} = y_1 - y_{1target} = -0.43 \qquad \frac{\partial Loss}{\partial y_2} = y_2 - y_{2target} = 0.61$$

$$\frac{\partial y_1}{\partial z_{20}} = a_{20}(1 - a_{20}) \approx 0.2444 \qquad \frac{\partial y_2}{\partial z_{21}} = a_{21}(1 - a_{21}) \approx 0.2379$$

$$\frac{\partial Loss}{\partial w(1)10} = \frac{\partial Loss}{\partial y_1} \times \frac{\partial y_1}{\partial z_{20}} \times \frac{\partial z_{20}}{\partial w(1)10} = -0.43 \times 0.2444 \times 0.54 \approx -0.0569$$

$$\frac{\partial Loss}{\partial w(1)11} = \frac{\partial Loss}{\partial y_2} \times \frac{\partial y_2}{\partial z_{21}} \times \frac{\partial z_{21}}{\partial w(1)11} = 0.61 \times 0.2379 \times 0.54 \approx 0.0781$$

$$\frac{\partial z_{10}}{\partial w(0)10} = x_1 = 0.2, \qquad \frac{\partial a_{10}}{\partial z_{10}} = a_{10}(1 - a_{10}) = 0.54 \times (1 - 0.54) \approx 0.2484$$

$$\frac{\partial Loss}{\partial w(0)10} = \left(\frac{\partial Loss}{\partial z_{20}} \times w(1)10 \times \frac{\partial z_{10}}{\partial w(0)10} + \frac{\partial Loss}{\partial z_{21}} \times w(1)11 \times \frac{\partial z_{10}}{\partial w(0)10}\right) \times \frac{\partial a_{10}}{\partial z_{10}}$$

$$\frac{\partial Loss}{\partial z_{20}} = \frac{\partial Loss}{\partial y_1} \times \frac{\partial y_1}{\partial z_{20}} = -0.43 \times 0.2449 \approx -0.1053$$

$$\frac{\partial Loss}{\partial z_{21}} = \frac{\partial Loss}{\partial y_2} \times \frac{\partial y_2}{\partial z_{21}} = 0.61 \times 0.2379 \approx 0.1451$$

$$\frac{\partial Loss}{\partial w(0)10} = (-0.1053 \times 0.4 + 0.1451 \times 0.5) \times 0.2 \times 0.2484 \approx 0.0014$$

$$w(1)10(new) = w(1)10 - \varepsilon \times \frac{\partial Loss}{\partial w(1)10} = 0.4 - 0.1 \times (-0.0584) = 0.40584$$

$$w(0)10(new) = w(0)10 - \varepsilon \times \frac{\partial Loss}{\partial w(0)10} = 0.0d486$$