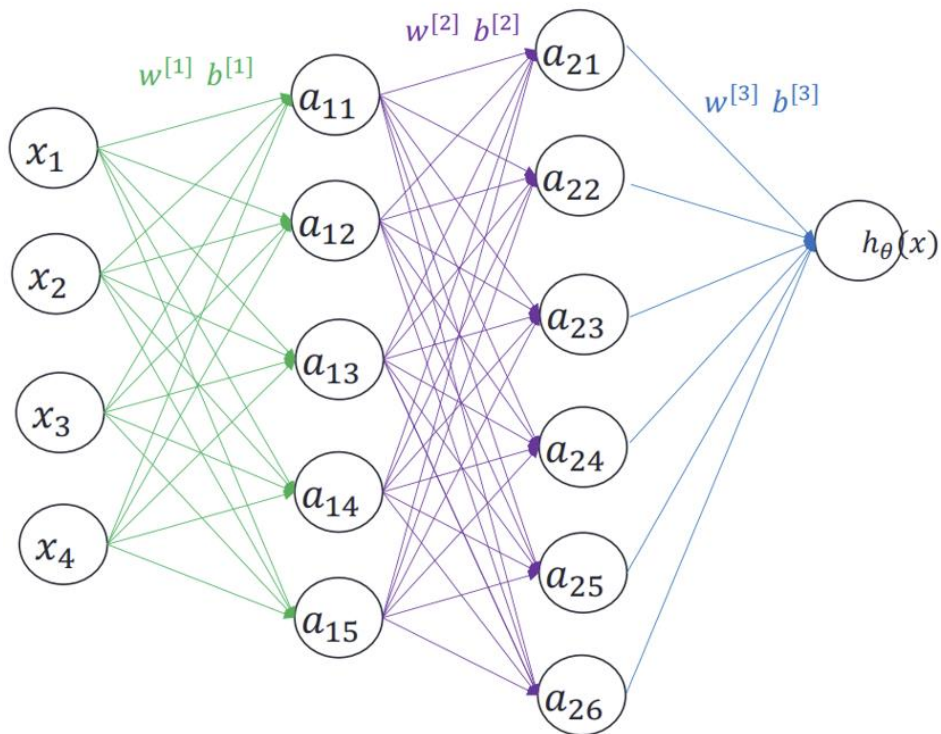


ToBig's 22기 정규세션 4주차

Neural Network 과제

이름: 양가현



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

$$\cdot z^{[1]} = W^{[1]}x + b^{[1]} \quad (W^{[1]}: 5 \times 4, b^{[1]}: 5 \times 1)$$

$$\cdot A^{[1]} = \text{ReLU}(z^{[1]})$$

$$\cdot A^{[2]} = \text{ReLU}(z^{[2]}) \quad (W^{[2]}: 6 \times 5, b^{[2]}: 6 \times 1)$$

$$\cdot z^{[3]} = W^{[3]}A^{[2]} + b^{[3]} \quad (W^{[3]}: 1 \times 6, b^{[3]}: 1 \times 1) \Rightarrow h_{\theta}(x) = z^{[3]}$$

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

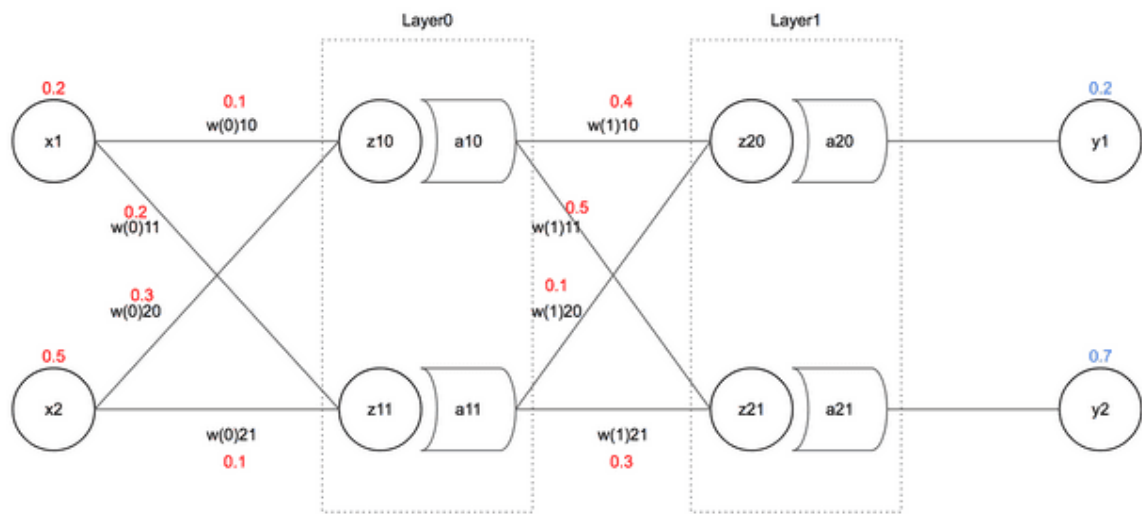
· layer 개수: (입력층 제외) 3개

· hidden layer 개수: 2개

$$\cdot \text{파라미터 총 개수: } \{(5 \times 4) + (5 \times 1)\} + \{(6 \times 5) + (6 \times 1)\} + \{(1 \times 6) + (1 \times 1)\}$$

$$= 25 + 36 + 7 = 68 \text{ (개)}$$

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.

$$\Delta(x) = \frac{1}{1+e^{-x}}, \quad MSE = \frac{1}{n} \sum (y - \hat{y})^2$$

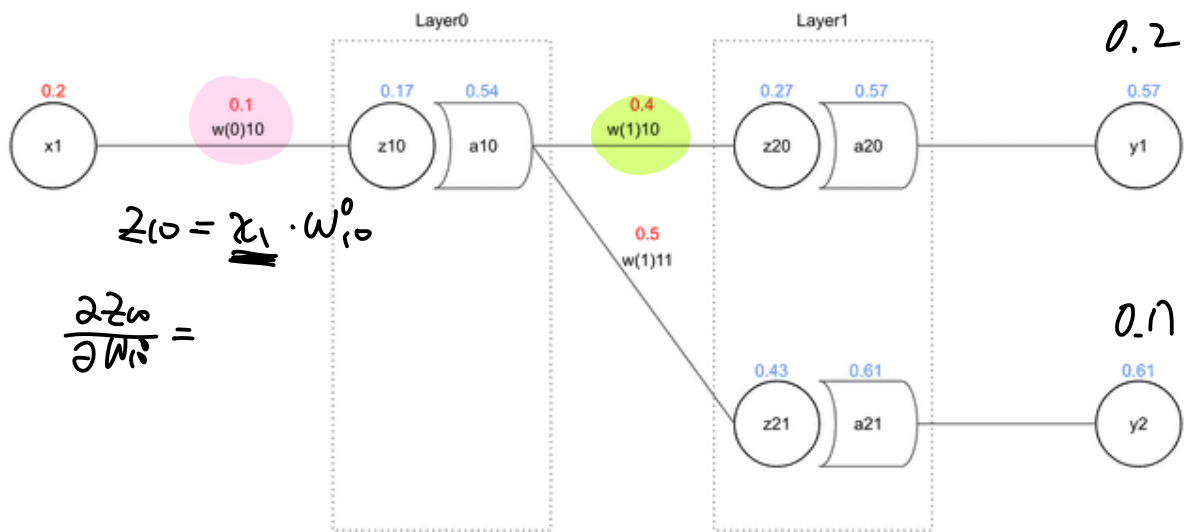
$$\begin{cases} \cdot z_{10} = 0.1x_1 + 0.3x_2 = 0.1 \times 0.2 + 0.3 \times 0.5 = 0.02 + 0.15 = 0.17 \\ \cdot z_{11} = 0.2x_1 + 0.1x_2 = 0.2 \times 0.2 + 0.1 \times 0.5 = 0.04 + 0.05 = 0.09 \end{cases}$$

$$\begin{cases} \cdot a_{10} = \Delta(z_{10}) = \frac{1}{1+e^{-0.17}} \approx 0.542 \\ \cdot a_{11} = \Delta(z_{11}) = \frac{1}{1+e^{-0.09}} \approx 0.522 \end{cases}$$

$$\begin{cases} \cdot z_{20} = 0.4a_{10} + 0.1a_{11} = 0.269 \\ \cdot z_{21} = 0.5a_{10} + 0.3a_{11} \approx 0.428 \end{cases}$$

$$\begin{cases} \cdot a_{20} = \Delta(z_{20}) \approx 0.567 \\ \cdot a_{21} = \Delta(z_{21}) \approx 0.605 \end{cases}$$

$$\cdot MSE = \frac{1}{2} \{ (0.2 - 0.567)^2 + (0.7 - 0.605)^2 \} \approx 0.072$$



$$f(x) = \frac{1}{1+e^{-x}}, \quad f'(x) = f(x) \cdot (1-f(x))$$

Q4. w^1_{10} 과 w^0_{10} 을 역전파(backpropagation) 기법을 사용하여 갱신하세요 ($\eta = 0.1$)

$$\cdot \text{Loss} = \frac{1}{2} \{ (t_1 - a_{20})^2 + (t_2 - a_{21})^2 \}$$

$$\cdot \frac{\partial \text{Loss}}{\partial a_{20}} = (t_1 - a_{20}) \times (-1) = (0.2 - 0.57) \times (-1) = 0.37$$

$$\cdot \frac{\partial a_{20}}{\partial z_{20}} = a_{20} (1 - a_{20}) = 0.57 \times (1 - 0.57) \approx 0.25$$

$$\cdot \frac{\partial z_{20}}{\partial w^1_{10}} = a_{10} = 0.54$$

$$\frac{\partial \text{Loss}}{\partial w^1_{10}} = \frac{\partial \text{Loss}}{\partial a_{20}} \times \frac{\partial a_{20}}{\partial z_{20}} \times \frac{\partial z_{20}}{\partial w^1_{10}} = 0.37 \times 0.25 \times 0.54 = 0.049$$

$$\cdot w^{1+}_{10} = w^1_{10} - (0.1)(0.049) = 0.395$$

$$\cdot \frac{\partial \text{Loss}_1}{\partial a_{10}} = \frac{\partial \text{Loss}_1}{\partial a_{20}} \times \frac{\partial a_{20}}{\partial z_{20}} \times \frac{\partial z_{20}}{\partial a_{10}} = -(t_1 - a_{20}) \times a_{20} (1 - a_{20}) \times w^1_{10} = 0.036$$

$$\cdot \frac{\partial \text{Loss}_2}{\partial a_{10}} = -0.0107$$

$$\cdot \frac{\partial \text{Loss}}{\partial w^0_{10}} = \left(\frac{\partial \text{Loss}_1}{\partial a_{10}} + \frac{\partial \text{Loss}_2}{\partial a_{10}} \right) \cdot \frac{\partial a_{10}}{\partial z_{10}} \cdot \frac{\partial z_{10}}{\partial w^0_{10}} = 0.0253 \times 0.2484 \times 0.2 = 0.00126$$

$$\cdot w^{0+}_{10} = w^0_{10} - 0.1 \times 0.00126 = 0.1 - 0.1 \times 0.00126 = 0.09974$$