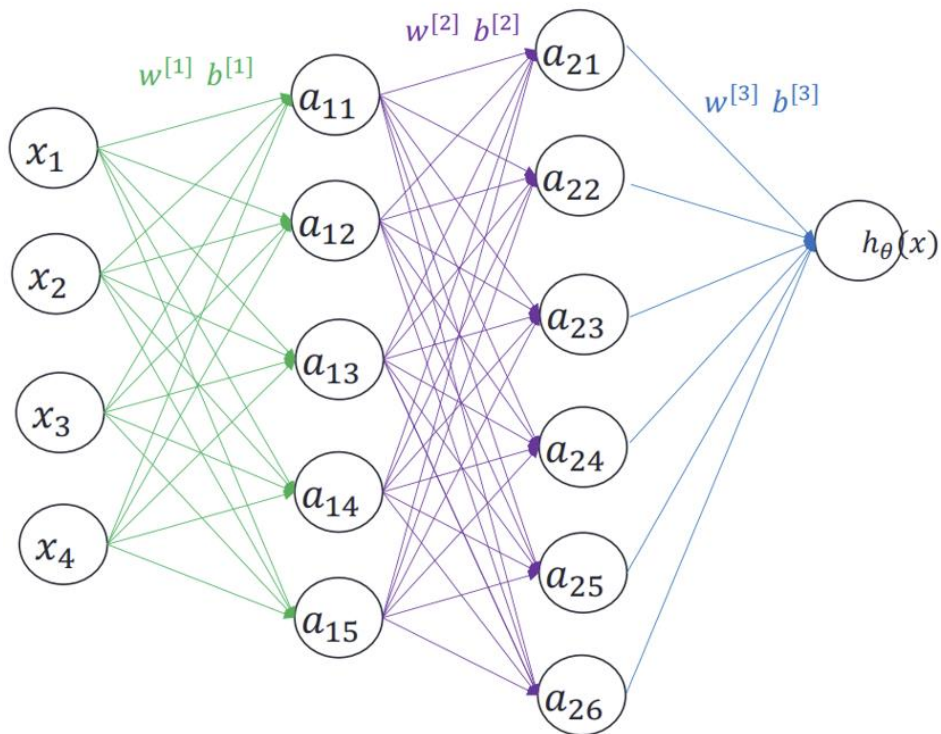


ToBig's 22기 정규세션 4주차

Neural Network 과제

이름: 김희망



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

$$\begin{aligned} X &= [x_1, x_2, x_3, x_4]^T \\ A1 &= [a_{11}, a_{12}, a_{13}, a_{14}, a_{15}]^T \\ A2 &= [a_{21}, a_{22}, a_{23}, a_{24}, a_{25}, a_{26}]^T \end{aligned}$$

위 처럼 노드를 벡터로 표시하자. 그러면

$$\begin{aligned} A1 &= \text{ReLU}(w^{[1]}b^{[1]} * X) \\ A2 &= \text{ReLU}(w^{[2]}b^{[2]} * A1) \\ h_{\theta}(x) &= w^{[3]}b^{[3]} * A2 \end{aligned}$$

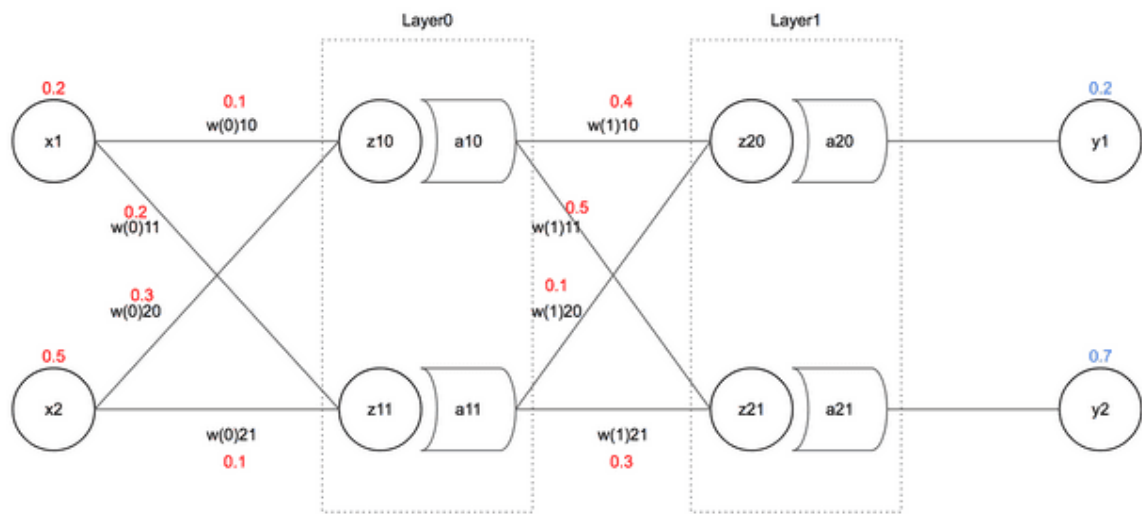
Q2. 이 네트워크를 구성하고 있는 layer 개수와 hidden layer 개수, 그리고 파라미터의 총개수를 각각 구해주세요.

input layer: 1개
hidden layer: 2개
output layer: 1개

total layer: 4개

파라미터 총 개수는 노드의 개수를 이용하여 구할 수 있다.
weight의 총 개수: $(4 \times 5) + (5 \times 6) + (6 \times 1) = 56$
bias의 총 개수: $5 + 6 + 1 = 12$
총 parameter의 개수: 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.

$$z_{10} = x_1 \cdot w(0)_{10} + x_2 \cdot w(0)_{20}$$

$$z_{20} = x_1 \cdot w(0)_{11} + x_2 \cdot w(0)_{21}$$

$$a_{10} = \sigma(z_{10})$$

$$a_{11} = \sigma(z_{11})$$

$$z_{20} = a_{10} \cdot w(1)_{10} + a_{11} \cdot w(1)_{20}$$

$$z_{21} = a_{10} \cdot w(1)_{11} + a_{11} \cdot w(1)_{21}$$

$$a_{20} = \sigma(z_{20})$$

$$a_{21} = \sigma(z_{21})$$

$$loss = \frac{1}{2} \cdot (y_1 - a_{20})^2 + \frac{1}{2} \cdot (y_2 - a_{21})^2$$

$$z_{10} = 0.17$$

$$z_{20} = 0.09$$

$$a_{10} = 0.54240$$

$$a_{11} = 0.52248$$

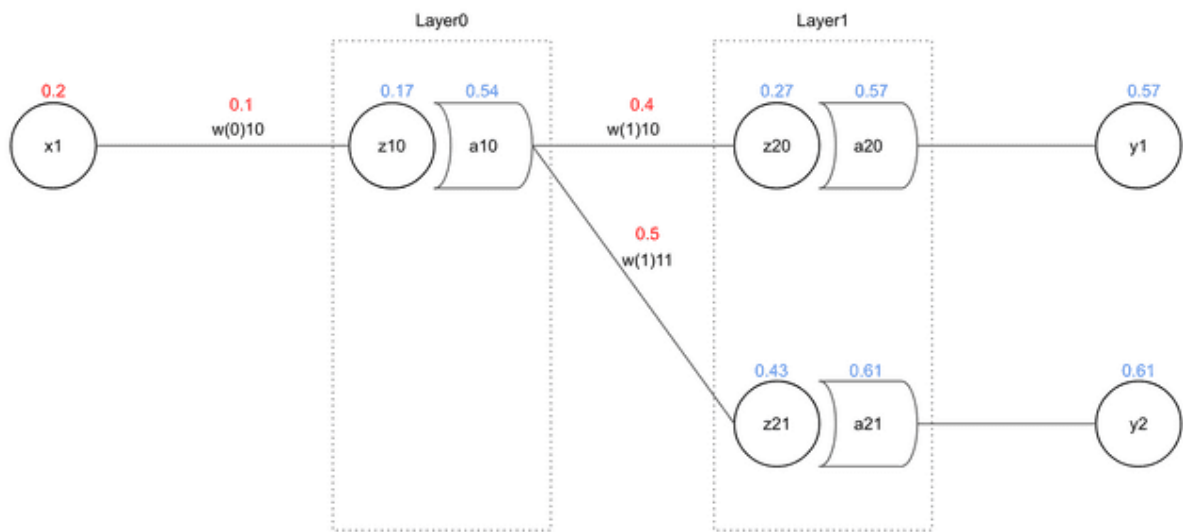
$$z_{20} = 0.269208$$

$$z_{21} = 0.427944$$

$$a_{20} = 0.56690$$

$$a_{21} = 0.60538$$

$$loss = 0.0717842772$$



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

$$\partial L / \partial w(1)_{10} = \partial L / \partial y_1 * \partial y_1 / \partial a_{20} * \partial a_{20} / \partial z_{20} * \partial z_{20} / \partial w(1)_{10}$$

$$\begin{aligned} \partial L / \partial w(0)_{10} &= \partial L_1 / \partial w(0)_{10} + \partial L_2 / \partial w(0)_{10} \\ &= \partial L / \partial y_1 * \partial y_1 / \partial a_{20} * \partial a_{20} / \partial z_{20} * \partial z_{20} / \partial z_{10} * \partial a_{10} / \partial z_{10} * \partial z_{10} / \partial w(0)_{10} \\ &\quad + \partial L / \partial y_2 * \partial y_2 / \partial a_{21} * \partial a_{21} / \partial z_{21} * \partial z_{21} / \partial z_{10} * \partial a_{10} / \partial z_{10} * \partial z_{10} / \partial w(0)_{10} \end{aligned}$$

이 때, $\sigma'(z) = \sigma(z) \times (1 - \sigma(z))$ 이므로,

$$\partial L / \partial w(1)_{10} = (y_1 - a_{20}) \times a_{20} \times (1 - a_{20}) \times a_{10} = -0.04897$$

$$\begin{aligned} \partial L / \partial w(0)_{10} &= [(y_1 - a_{20}) \times a_{20} \times (1 - a_{20}) \times w(1)_{10} + (y_2 - a_{21}) \times a_{21} \times (1 - a_{21}) \times w(1)_{11}] \times a_{10} \times (1 - a_{10}) \times x_1 \\ &= -0.0012707 \end{aligned}$$

결과적으로

$$\begin{aligned} w(1)_{10} &= w(1)_{10} + \eta * 0.04897 \\ w(0)_{10} &= w(0)_{10} + \eta * 0.0012707 \end{aligned}$$