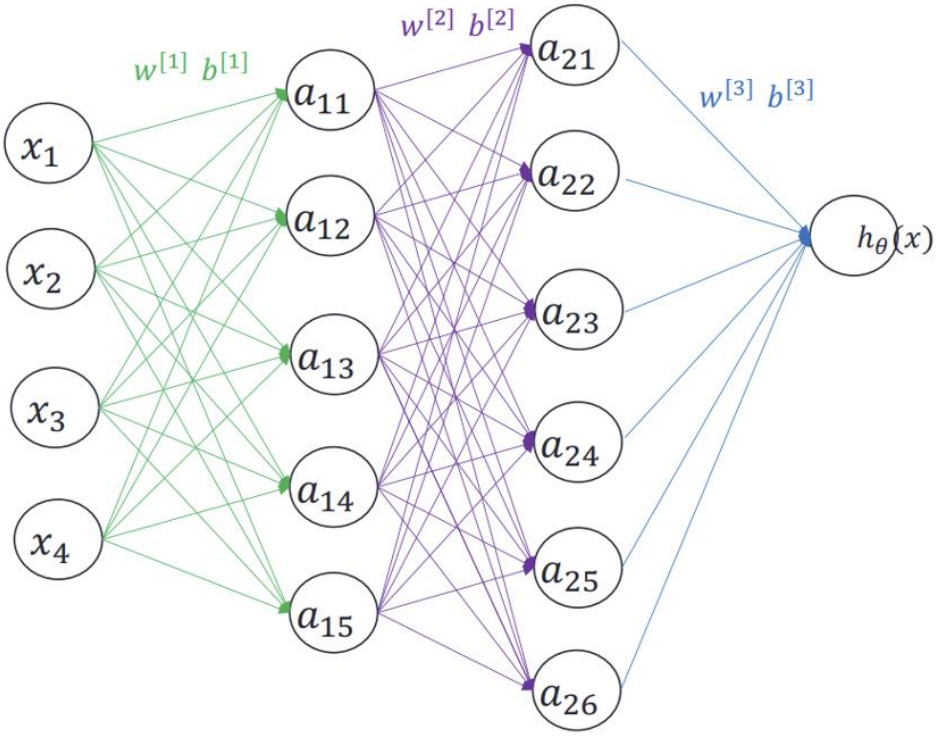
ToBig's 22기 정규세션 4주차

Neural Network 과제

이름: 한다은



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

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad h_\theta(x) = W^{[3]}\Big(ReLU\Big(W^{[2]}\Big(ReLU\Big(W^{[1]} \cdot X + b^{[1]}\Big)\Big) + b^{[2]}\Big)\Big) + b^{[3]}$$

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

1| layer 개수: 4개 (input layer 1개, hidden layer 2개, output layer 1개)

2| hidden layer 개수 : 2개

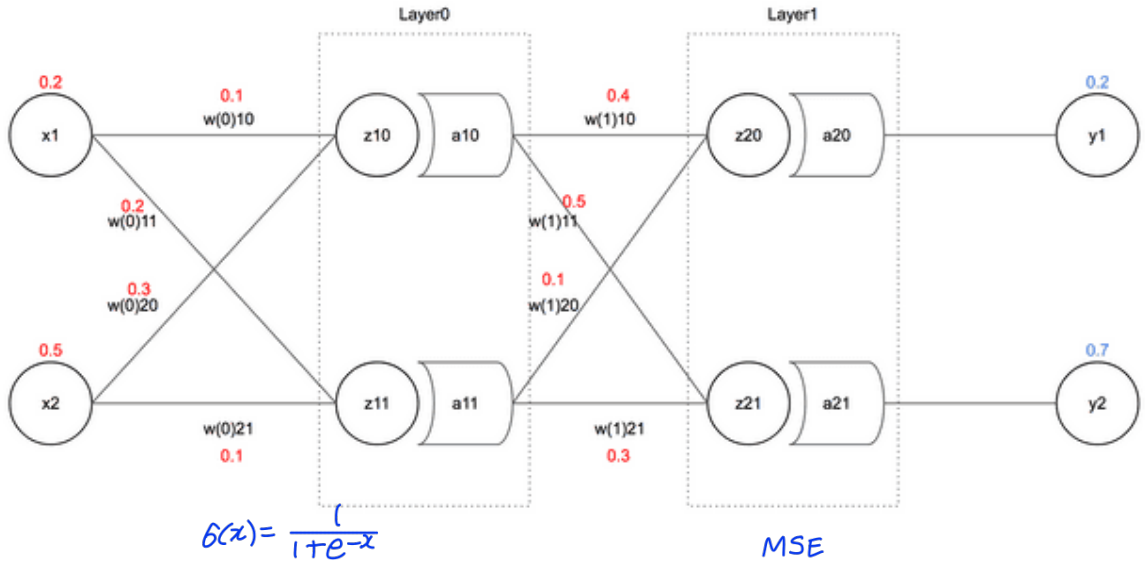
3| 파라미터의 총개수

$(4\times5+5)+(5\times6+6)+(6\times1+1)$

$= 25+36+7 = 68$

3번 답 : 68개

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.

[propagation]

$z_{10} = 0.1 \cdot (x_1) + (0.3)(x_2) = (0.1)\cdot(0.2) + (0.3)\cdot(0.5) = \boxed{0.17}$

$z_{11} = 0.2 \cdot (x_1) + (0.1)(x_2) = (0.2)\cdot(0.2) + (0.1)\cdot(0.5) = \boxed{0.09}$

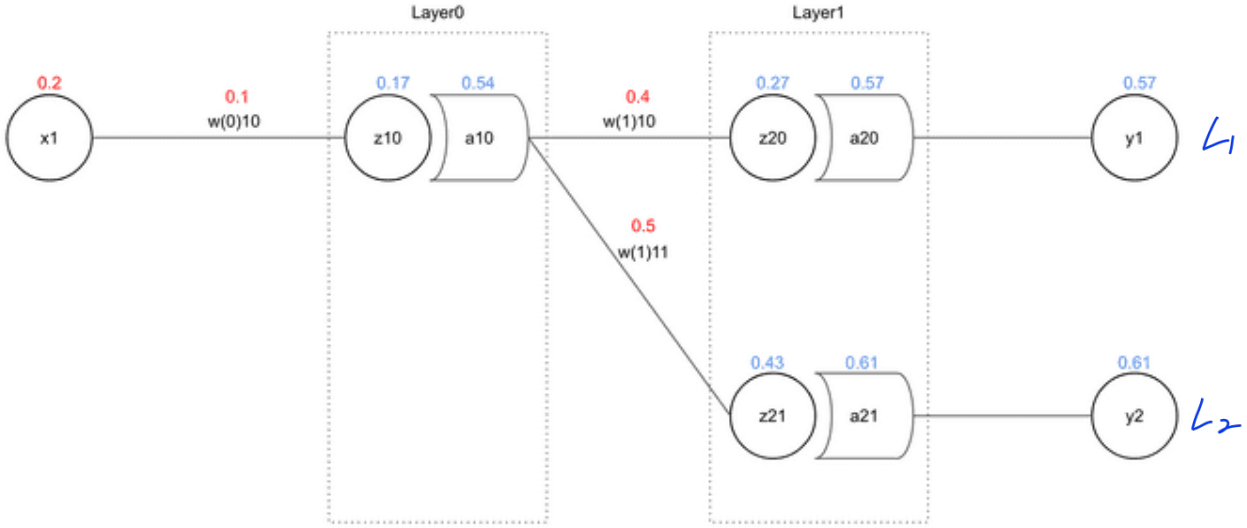
$a_{10} = \frac{1}{1+e^{-0.17}} = \boxed{0.5424}$, $a_{11} = \frac{1}{1+e^{-0.09}} = \boxed{0.5225}$

$z_{20} = (0.4)\cdot(a_{10}) + (0.1)(a_{11}) = \boxed{0.2692}$

$z_{21} = (0.5)\cdot(a_{10}) + (0.3)(a_{11}) = \boxed{0.428}$

$a_{20} = \frac{1}{1+e^{-0.2692}} = \boxed{0.5669}$, $a_{21} = \frac{1}{1+e^{-0.428}} = \boxed{0.6054}$

$Loss = \frac{1}{2}((0.5669-0.2)^2 + (0.6054-0.7)^2) = \boxed{0.0718}$



시그모이드 함수 미분 $\sigma(x)(1-\sigma(x))$

(학습률 $\alpha = 0.1$로 설정)

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

$y_1 = 0.2$, $y_2 = 0.7$ ← Q3에서 , $L = \frac{1}{2}((a_{20}-y_1)^2 + (a_{21}-y_2)^2)$ $a_{20}=0.57$, $a_{21}=0.61$

$$\frac{\partial L}{\partial w'_{10}} = \frac{\partial L}{\partial a_{20}} \cdot \frac{\partial a_{20}}{\partial z_{20}} \cdot \frac{\partial z_{20}}{\partial w'_{10}}$$

$$\frac{\partial L}{\partial a_{20}} = \frac{1}{2}\left(2(a_{20}-y_1)\right) = a_{20}-y_1 = 0.57-0.2 = 0.37$$

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

$$\frac{\partial z_{20}}{\partial w'_{10}} = a_{10} = 0.54$$

$$\therefore \frac{\partial L}{\partial w'_{10}} = 0.37 \times 0.25 \times 0.54 = 0.049$$

↓ 이 값을 이용하여 w'_{10} 업데이트

w'_{10}의 업데이트식

$w'_{10} \leftarrow w'_{10} - \alpha \cdot \frac{\partial L}{\partial w'_{10}}$

$$w'_{10} = 0.4 - (0.1 \times 0.049) = 0.3951$$ → 답①

전체 에러 L라 하면

$L = L_1 + L_2$

$$\frac{\partial L}{\partial w^0_{10}} = \left(\frac{\partial L}{\partial a_{20}} \cdot \frac{\partial a_{20}}{\partial z_{20}} \cdot \frac{\partial z_{20}}{\partial a_{10}} \cdot \frac{\partial a_{10}}{\partial z_{10}} \cdot \frac{\partial z_{10}}{\partial w^0_{10}}\right) + \left(\frac{\partial L}{\partial a_{21}} \cdot \frac{\partial a_{21}}{\partial z_{21}} \cdot \frac{\partial z_{21}}{\partial a_{10}} \cdot \frac{\partial a_{10}}{\partial z_{10}} \cdot \frac{\partial z_{10}}{\partial w^0_{10}}\right)$$

$\frac{\partial L}{\partial a_{20}} = 0.37$, $\frac{\partial a_{20}}{\partial z_{20}} = 0.25$

$z_{20} = w'_{10} \cdot a_{10}$ 이므로 $\frac{\partial z_{20}}{\partial a_{10}} = w'_{10} = 0.4$

$\frac{\partial a_{10}}{\partial z_{10}} = \sigma(z_{10})(1-\sigma(z_{10}))$

$= 0.54(1-0.54) = 0.2484$

$\frac{\partial z_{10}}{\partial w^0_{10}} = 0.2$

$\frac{\partial L}{\partial a_{21}} = \frac{1}{2}\left(2(a_{21}-y_2)\right) = a_{21}-y_2 = 0.61-0.7$

$= -0.09$

$\frac{\partial a_{21}}{\partial z_{21}} = 0.61(1-0.61) = 0.2379$

$\frac{\partial z_{21}}{\partial a_{10}} = 0.5$

따라서

$$\frac{\partial L}{\partial w_{10}^{0}} = (0.31 \times 0.25 \times 0.4 \times 0.2484 \times 0.2)$$

$$+ (-0.09 \times 0.2314 \times 0.5 \times 0.2484 \times 0.2)$$

$$= 0.0013$$

$$w_{10}^{0} \leftarrow w_{10}^{0} - \alpha \cdot \frac{\partial L}{\partial w_{10}^{0}}, \quad (\alpha = 0.1)$$

$$w_{10}^{0} = 0.1 - 0.1 \times (0.0013) = 0.09,$$ $\boxed{w_{10}^{0} = 0.09\text{로 갱신}}$ ← 답