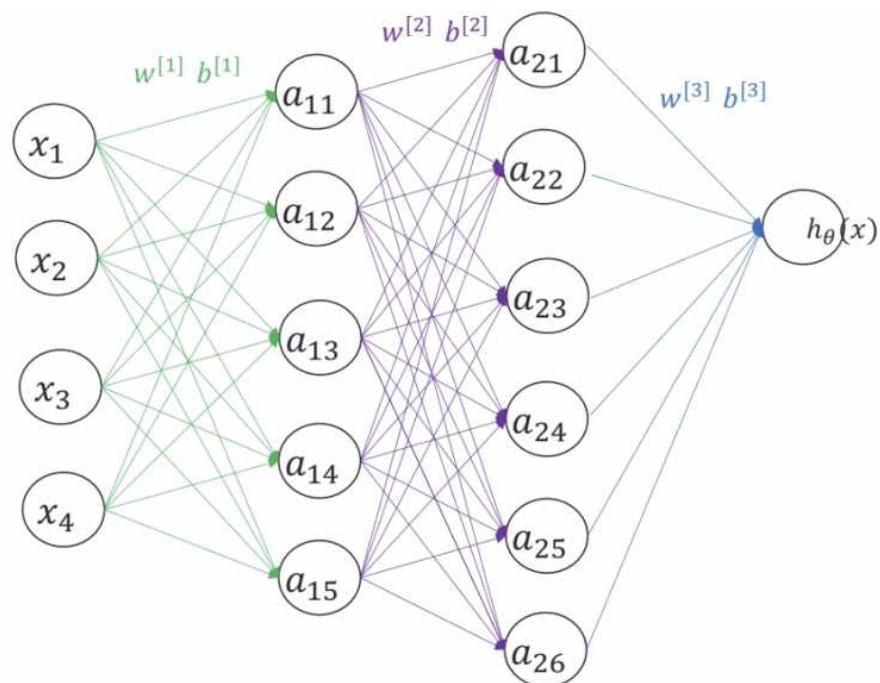


ToBig's 22기 정규세션 4주차 Neural Network 과제

22기 송수현



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

A1.

$$z[l] = w[l] \cdot a[l-1] + b[l]$$

$$a[l] = \text{ReLU}(z[l]) \text{ (마지막 층을 제외한 모든 층)}$$

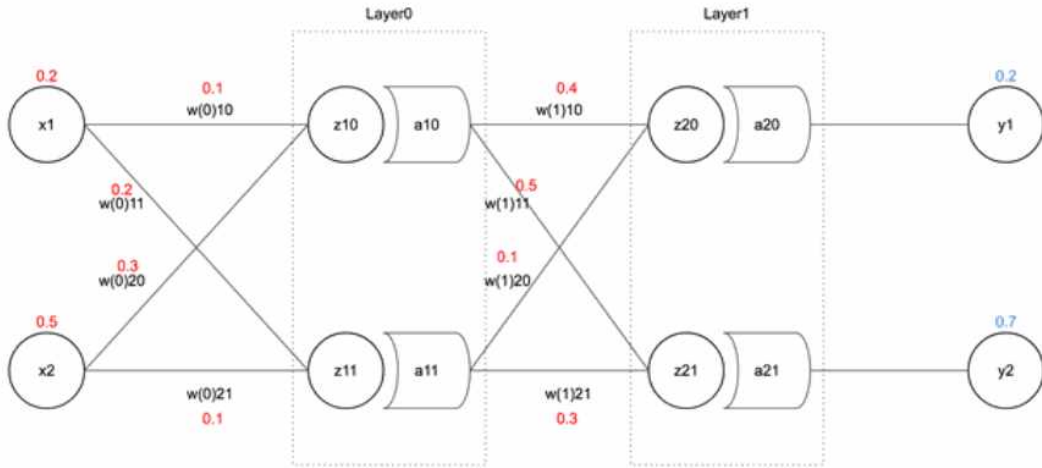
마지막 층의 출력은 $a[L] = z[L]$ (활성화 함수 없음)

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

A2.

- Layer 개수: 4개 (입력층 1개, 은닉층 2개, 출력층 1개)
- Hidden Layer 개수 : 2개
- 파라미터의 총개수 : 54개

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.

A3.

1. Layer 0의 뉴런 계산:

첫 번째 노드 z_{10} 계산:

입력 $x_1 = 0.2$ 와 $x_2 = 0.5$ 가 주어졌으며, 각 가중치는 다음과 같다:

- $w(0)_{10} = 0.1$
- $w(0)_{11} = 0.2$

z_{10} 는 다음과 같이 계산된다:

$$z_{10} = (0.1 \times 0.2) + (0.2 \times 0.5) = 0.02 + 0.1 = 0.12$$

두 번째 노드 z_{11} 계산:

각 가중치는 다음과 같다:

- $w(0)_{20} = 0.2$
- $w(0)_{21} = 0.3$

z_{11} 는 다음과 같이 계산된다:

$$z_{11} = (0.2 \times 0.2) + (0.3 \times 0.5) = 0.04 + 0.15 = 0.19$$

2. Layer 0의 활성화 함수 계산:

시그모이드 활성화 함수는 다음과 같이 계산된다:

$$a_{10} = \sigma(z_{10}) = \frac{1}{1 + e^{-0.12}} \approx 0.52996$$

$$a_{11} = \sigma(z_{11}) = \frac{1}{1 + e^{-0.19}} \approx 0.54738$$

3. Layer 1의 뉴런 계산:

첫 번째 노드 z_{20} 계산:

각 가중치는 다음과 같다:

- $w(1)_{10} = 0.4$
- $w(1)_{11} = 0.5$

z_{20} 은 다음과 같이 계산된다:

$$z_{20} = (0.4 \times a_{10}) + (0.5 \times a_{11}) = (0.4 \times 0.52996) + (0.5 \times 0.54738) = 0.211984 + 0.27369 = 0.485674$$

두 번째 노드 z_{21} 계산:

각 가중치는 다음과 같다:

- $w(1)_{20} = 0.4$
- $w(1)_{21} = 0.5$

$$z_{21} = (0.4 \times a_{10}) + (0.5 \times a_{11}) = (0.4 \times 0.52996) + (0.5 \times 0.54738) = 0.211984 + 0.27369 = 0.485674$$

4. Layer 1의 활성화 함수 계산:

$$a_{20} = \sigma(z_{20}) = \frac{1}{1 + e^{-0.485674}} \approx 0.61912$$

$$a_{21} = \sigma(z_{21}) = \frac{1}{1 + e^{-0.485674}} \approx 0.61912$$

5. 손실 함수 $loss$ 계산:

평균 제곱 오차(MSE) 손실 함수는 다음과 같이 계산된다:

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

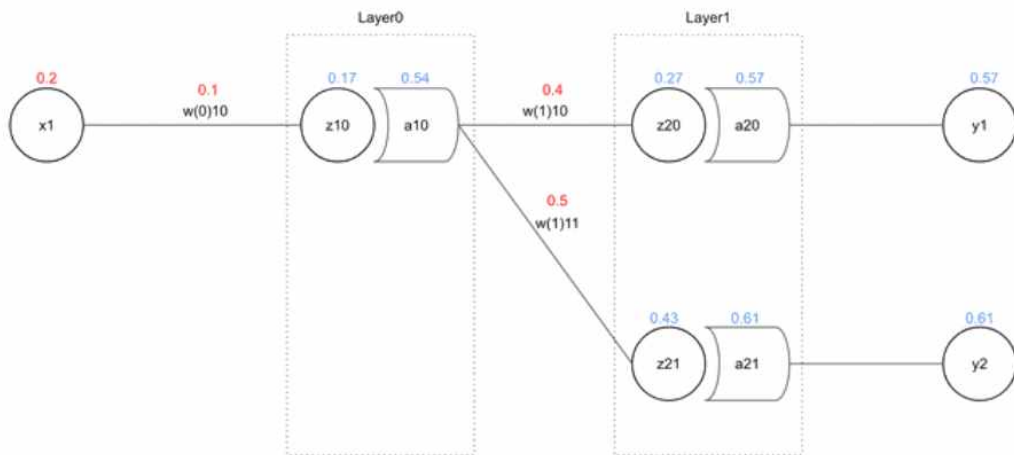
여기서 $y_1 = 0.2$, $y_2 = 0.7$ 이다.

$$loss = \frac{1}{2} [(0.2 - 0.61912)^2 + (0.7 - 0.61912)^2]$$

$$loss = \frac{1}{2} [(-0.41912)^2 + (0.08088)^2]$$

$$loss = \frac{1}{2} [0.17567 + 0.00654] = \frac{1}{2} \times 0.18221 \approx 0.091105$$

따라서, 손실 $loss$ 는 약 0.091105이다.



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

A4.

역전파 과정

순전파 (Forward Propagation)

- 주어진 신경망에서 입력 $x_1 = 0.2$ 와 가중치 $w(0)_{10} = 0.1$ 를 사용하여 Layer 0의 노드 z_{10} 값을 계산한다.
- 이미지에서 이미 $z_{10} = 0.17$ 로 주어져 있으며, 활성화 함수 $ReLU$ 를 거쳐 $a_{10} = 0.54$ 가 계산된다.
- Layer1에서의 출력 값도 이미지에서 제공된 대로 $z_{20} = 0.27$, $a_{20} = 0.57$, $z_{21} = 0.43$, $a_{21} = 0.61$ 이다.

오차 계산 (Error Calculation)

- 출력층에서의 오차는 δ_{20} 와 δ_{21} 로 구할 수 있다.
- 출력층 노드 $y_1 = 0.57$, $y_2 = 0.61$ 에 대해, 오차는 다음과 같다:

$$\delta_{20} = a_{20} - y_1 = 0.57 - 0.57 = 0$$

$$\delta_{21} = a_{21} - y_2 = 0.61 - 0.61 = 0$$

- 위의 값들이 0이 아닌 경우 일반적인 경우로 계산을 진행해야 한다.

은닉층에서의 오차 계산 (Hidden Layer Error Calculation)

- 은닉층에서의 오차는 다음과 같이 계산된다:

$$\delta_{10} = (\delta_{20} \times w_{10}^1 + \delta_{21} \times w_{11}^1) \times \sigma'(z_{10})$$

여기서 $\sigma'(z_{10})$ 는 $ReLU$ 의 도함수로, $z_{10} > 0$ 인 경우에는 1이 된다.

- 주어진 문제에서 $w_{10}^1 = 0.4$ 이고 $w_{11}^1 = 0.5$ 이다.

가중치 갱신 (Weight Update)

- 각 가중치를 갱신한다:

$$w_{10}^{1(new)} = w_{10}^{1(old)} - \eta \times \delta_{20} \times a_{10}$$

$$w_{10}^{0(new)} = w_{10}^{0(old)} - \eta \times \delta_{10} \times x_1$$

여기서 η 는 학습률이다.