

17기 정규세션

ToBig's 16기 주지훈

Neural Network Basic

Contents

Unit 01 | Perceptron

Unit 02 | Backpropagation

01 | Perceptron

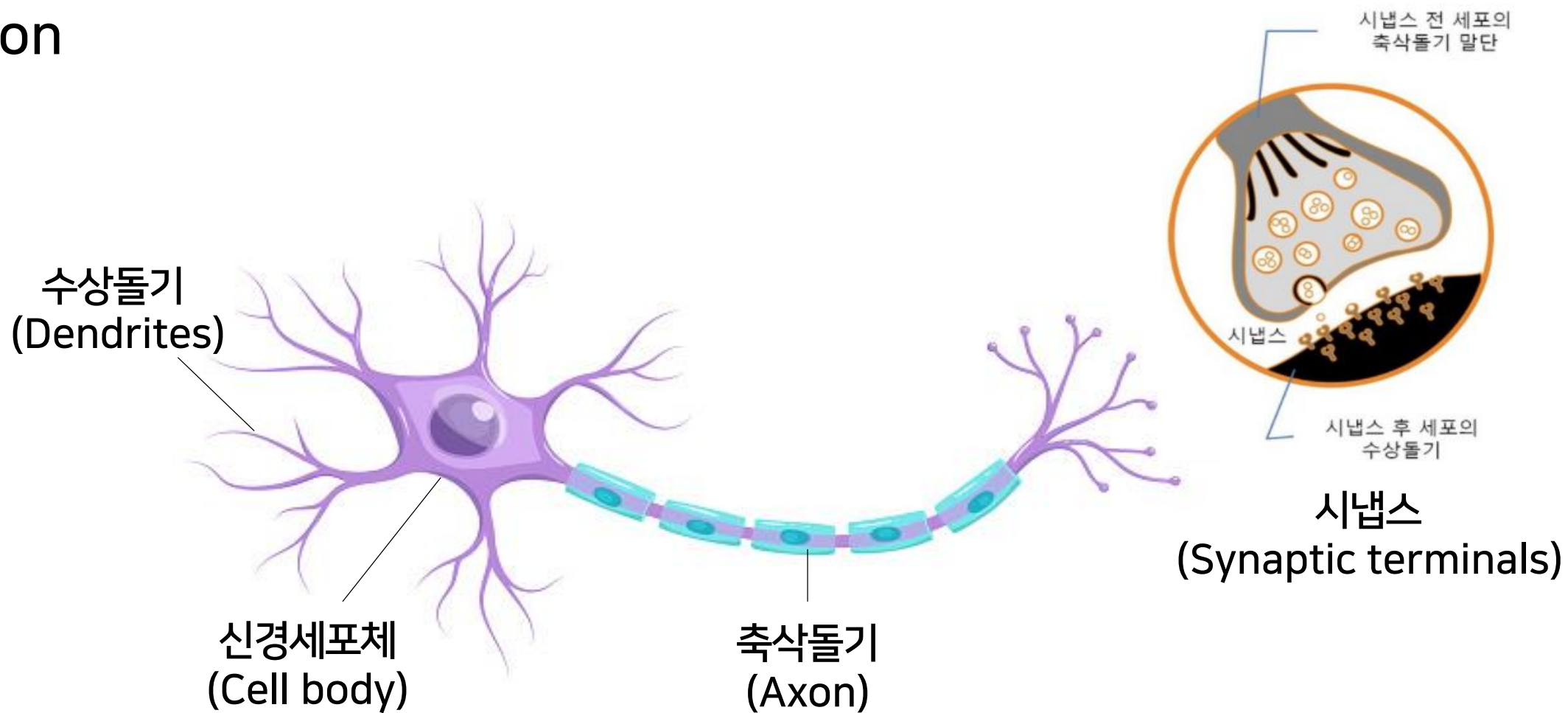
Unit 01 | Perceptron

인공신경망이란?

“ 수학적 논리학이 아닌 **인간의 두뇌를 모방**하여
수많은 간단한 처리기들(뉴런)의 네트워크를 통해
문제를 해결하는 기계학습 모델 ”

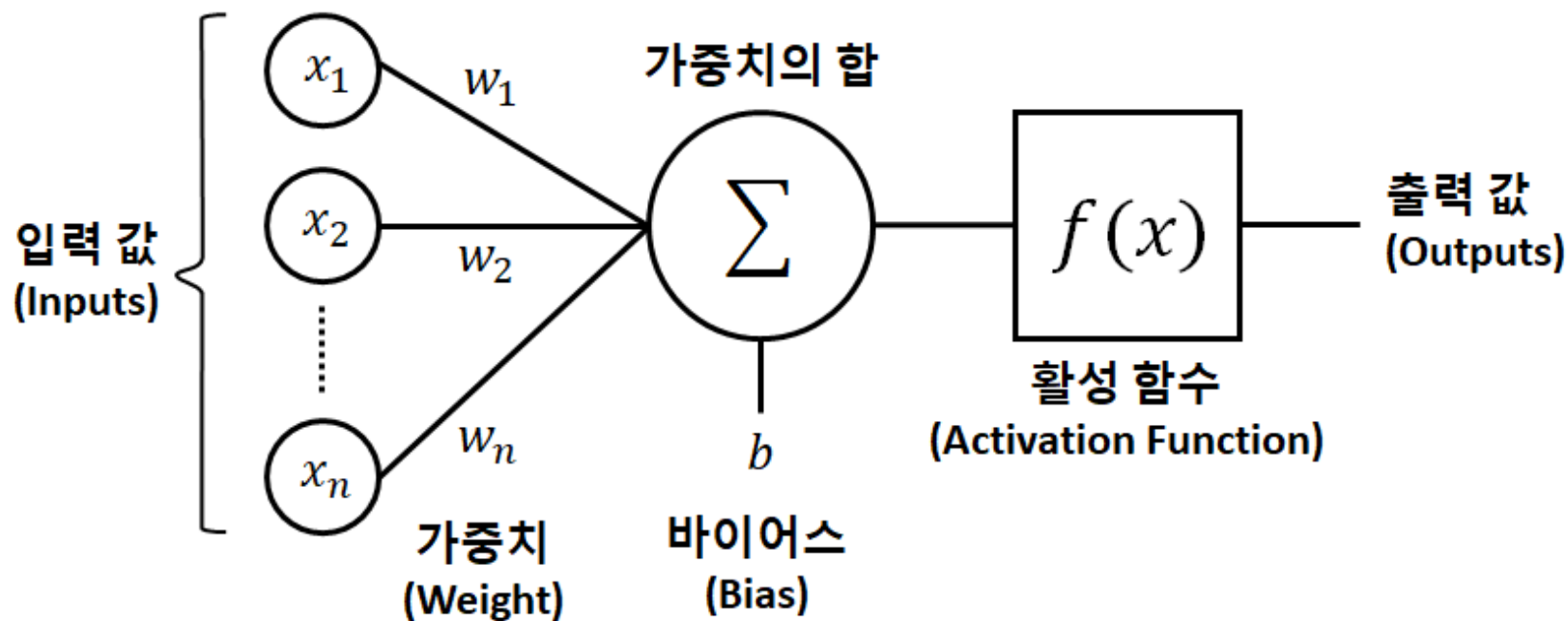
Unit 01 | Perceptron

Neuron



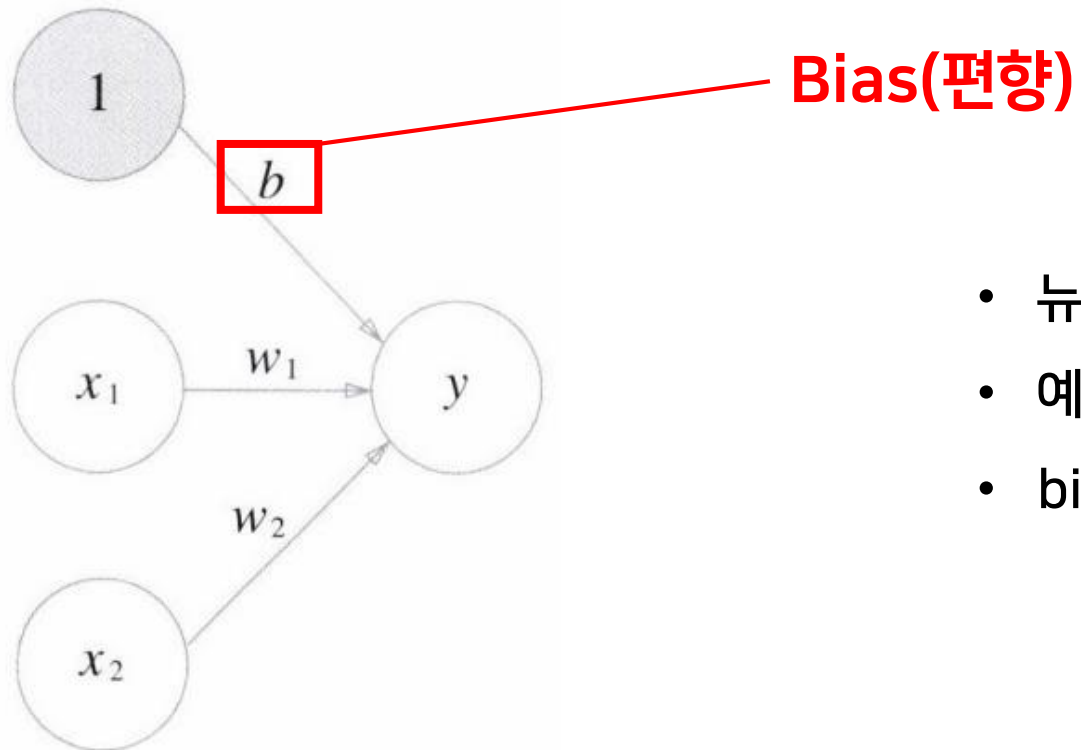
Unit 01 | Perceptron

Perceptron



Unit 01 | Perceptron

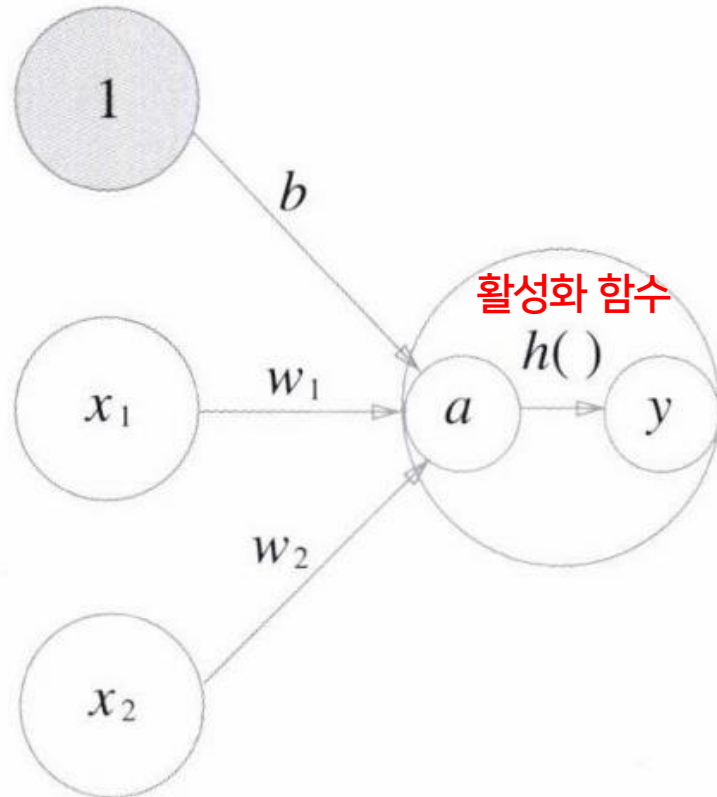
Perceptron



- 뉴런이 얼마나 쉽게 활성화 되느냐를 제어
- 예상하지 못한 잡음(노이즈)를 제거하기 위한 것
- bias가 작으면 입력에 더 의존하게 됨

Unit 01 | Perceptron

활성화 함수



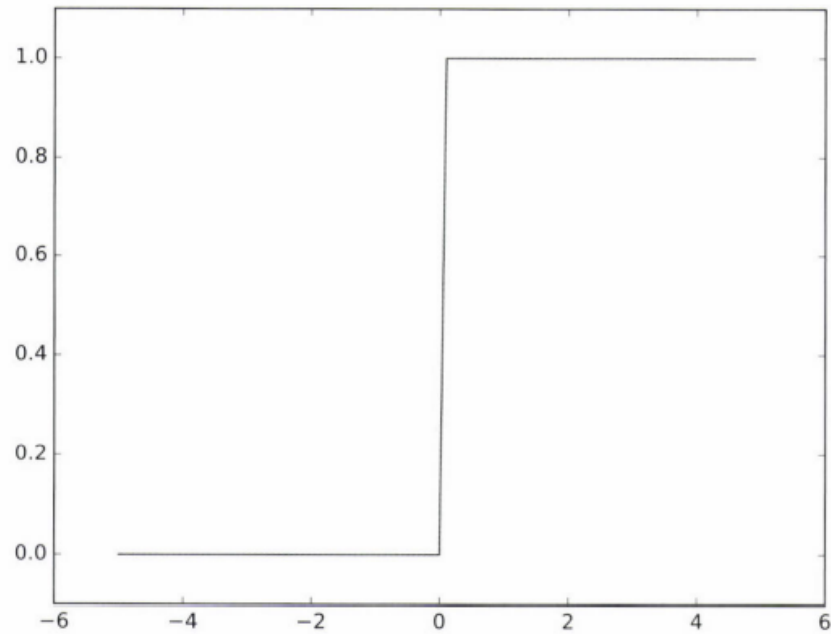
$$a = b + w_1x_1 + w_2x_2$$

$$y = h(a)$$

$$h(a) = \begin{cases} 0 & (a \leq 0) \\ 1 & (a > 0) \end{cases}$$

Unit 01 | Perceptron

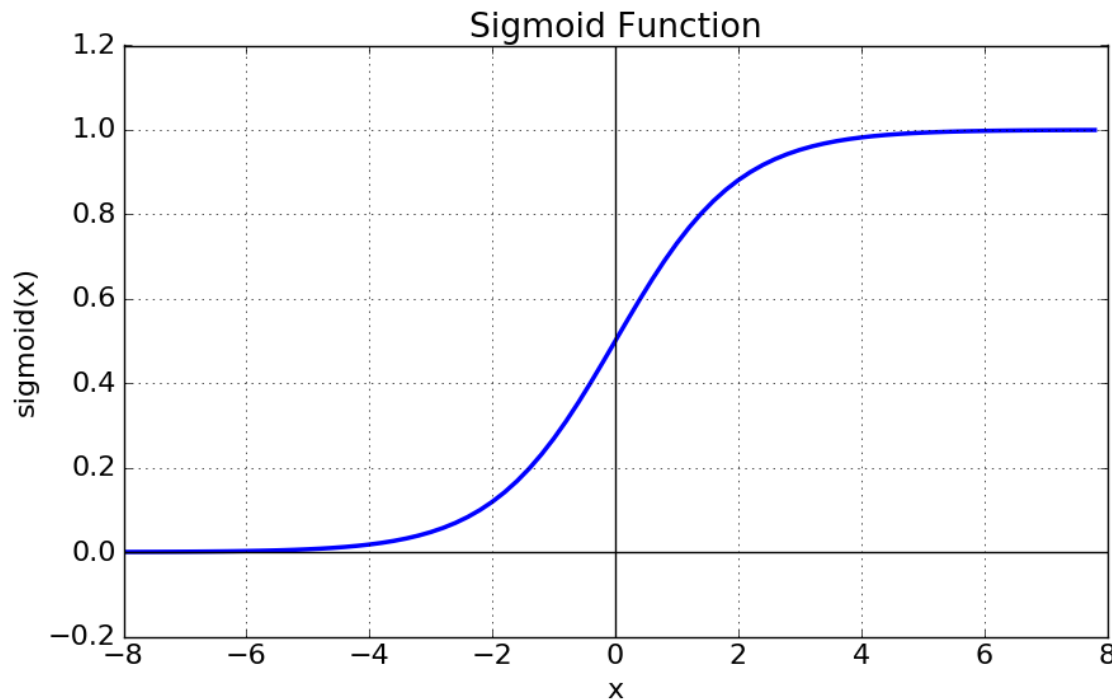
활성화 함수 - Step function



- 출력이 0 또는 1
- 활성화할지 말지 여부만 반환

Unit 01 | Perceptron

활성화 함수 - Sigmoid

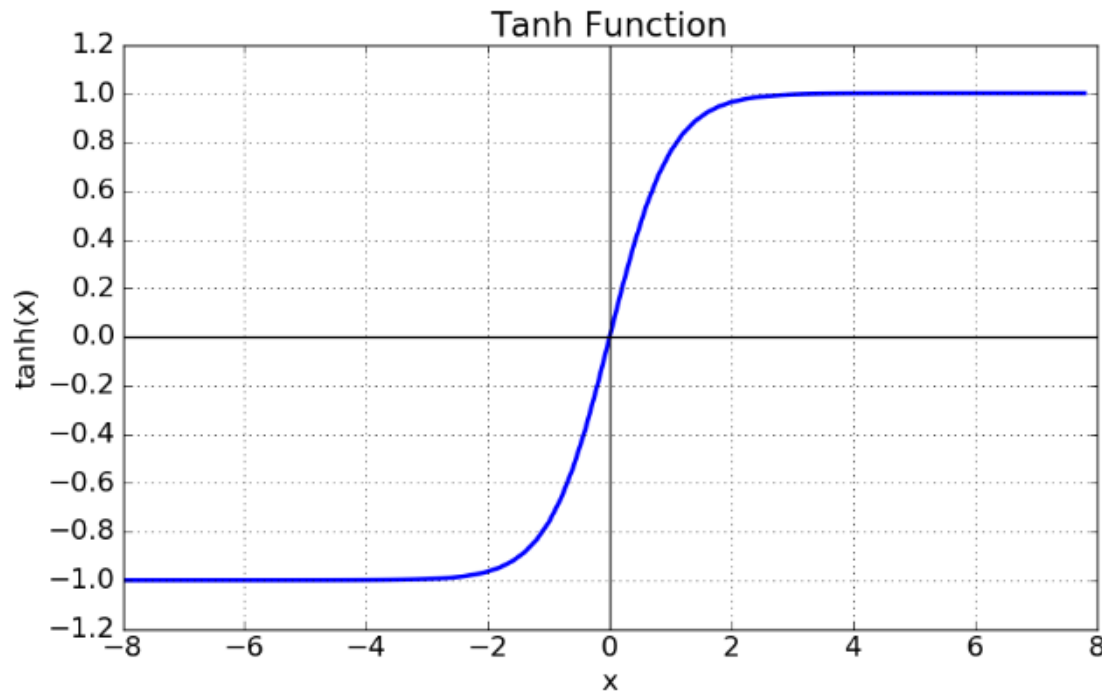


$$h(x) = \frac{1}{1 + \exp(-x)}$$

- 0에서 1 사이의 값 출력
- 활성화 여부가 아닌, 활성화 정도를 반환
- 1에 가까울수록 많이 활성화됐다는 뜻

Unit 01 | Perceptron

활성화 함수 - Tanh

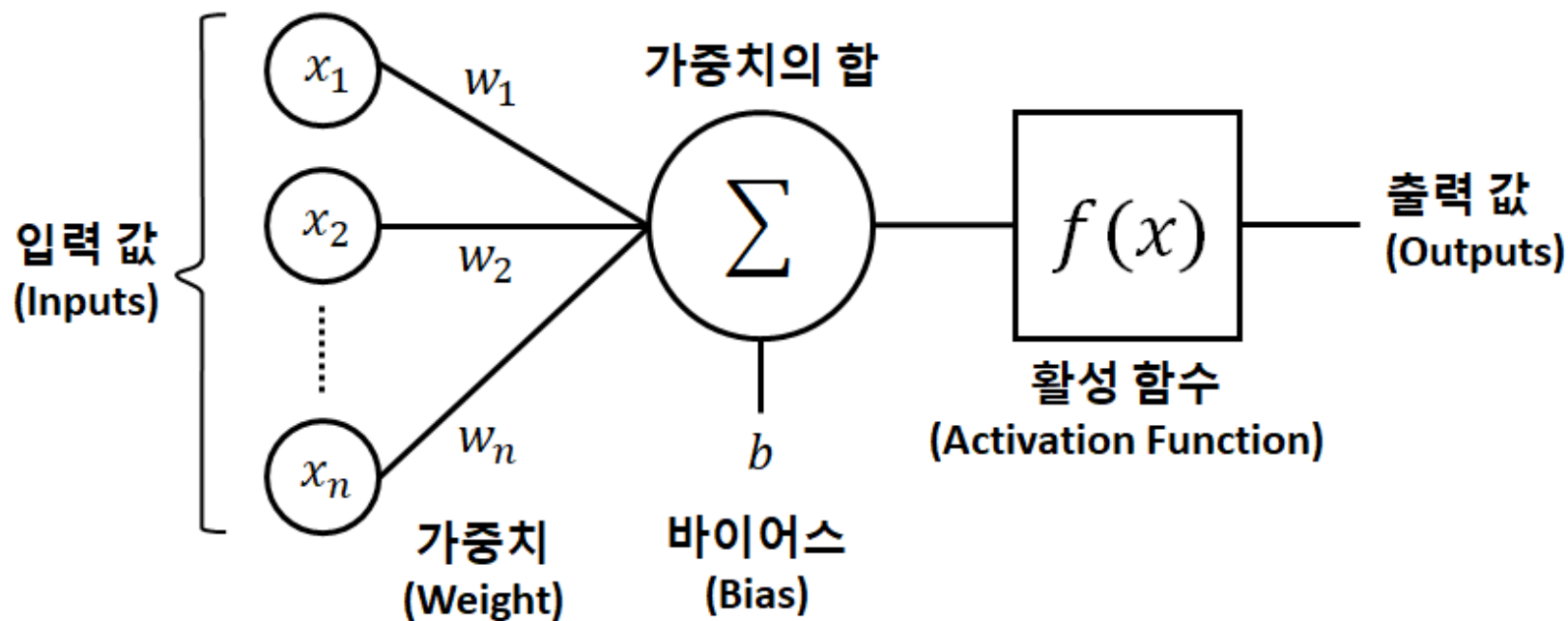


$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- -1에서 1 사이의 값 출력
- 항상 양수만 나오는 sigmoid의 문제를 해결

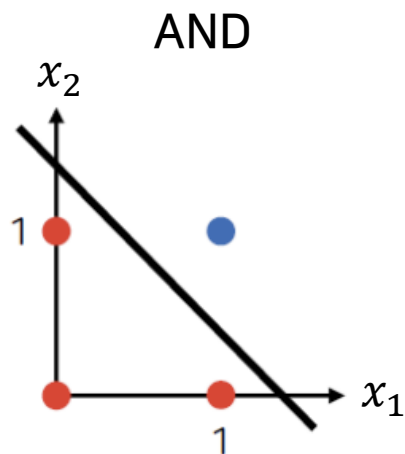
Unit 01 | Perceptron

Perceptron



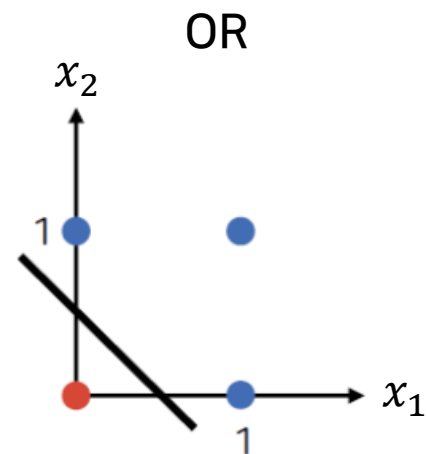
Unit 01 | Perceptron

Perceptron 연산



$$w_0 = 1.0, w_1 = 1.0, b = -1.5$$

x_1	x_2	S	y
0	0	-1.5	0
0	1	-0.5	0
1	0	-0.5	0
1	1	0.5	1

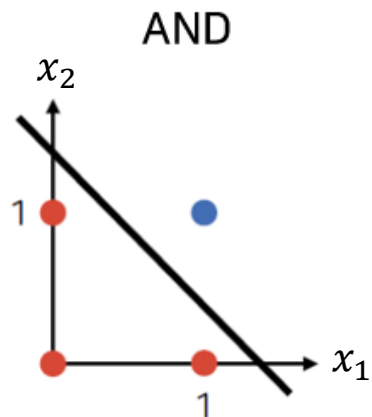


$$w_0 = -0.5, w_1 = 1.0, b = 1.0$$

x_1	x_2	S	y
0	0	-0.5	0
0	1	0.5	1
1	0	0.5	1
1	1	1.5	1

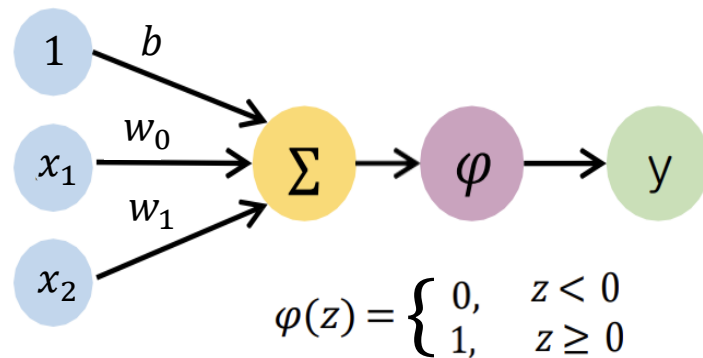
Unit 01 | Perceptron

Perceptron 연산



$$w_0 = 1.0, w_1 = 1.0, b = -1.5$$

x_1	x_2	S	y
0	0	-1.5	0
0	1	-0.5	0
1	0	-0.5	0
1	1	0.5	1

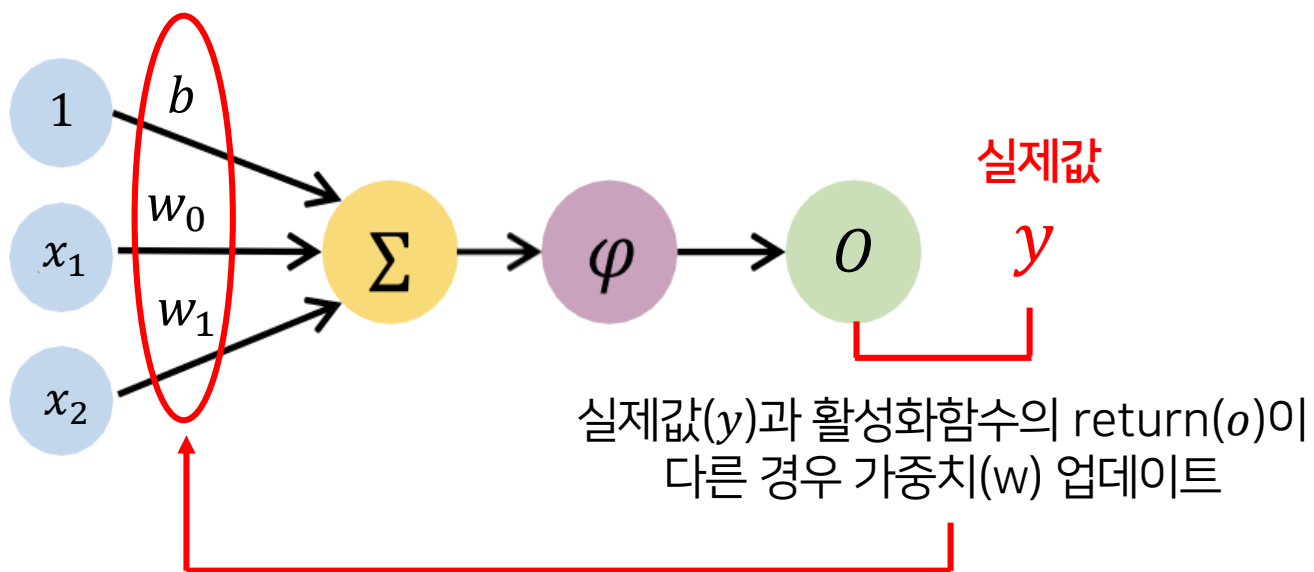


$$\varphi(w_1x_1 + w_2x_2 + w_3) = y$$

- ① $(1.0 \times 0) + (1.0 \times 0) + (-1.5) = -1.5$
 $\varphi((1.0 \times 0) + (1.0 \times 0) + (-1.5)) = 0$
- ② $(1.0 \times 0) + (1.0 \times 1) + (-1.5) = -0.5$
 $\varphi((1.0 \times 0) + (1.0 \times 1) + (-1.5)) = 0$
- ③ $(1.0 \times 1) + (1.0 \times 0) + (-1.5) = -0.5$
 $\varphi((1.0 \times 1) + (1.0 \times 0) + (-1.5)) = 0$
- ④ $(1.0 \times 1) + (1.0 \times 1) + (-1.5) = 0.5$
 $\varphi((1.0 \times 1) + (1.0 \times 1) + (-1.5)) = 1$

Unit 01 | Perceptron

Perceptron 학습



가중치 조정 식

$$w_i \leftarrow w_i + \eta(y - o)x_i$$

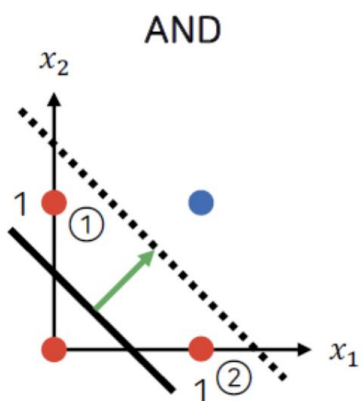


학습률(learning rate)

너무 작으면 학습 속도가 매우 느리고
너무 크면 오차의 최솟값을 지나칠 수 있음

Unit 01 | Perceptron

Perceptron 학습



$$w_1 = 0.55, w_2 = 0.55, b = -0.65$$

x_1	x_2	o	y
0	0	0	0
0	1	1	0
1	0	1	0
1	1	1	1

①

②

$$w_i \leftarrow w_i + \eta(y - o)x_i \quad \eta = 0.05$$

$$\begin{aligned} b &\leftarrow b + 0.05(0 - 1) \times 1 \\ \textcircled{1} \quad w_1 &\leftarrow w_1 + 0.05(0 - 1) \times 0 \\ w_2 &\leftarrow w_2 + 0.05(0 - 1) \times 1 \end{aligned}$$

$$\begin{aligned} b &\leftarrow -0.65 + 0.05(0 - 1) \times 1 = -0.7 \\ w_1 &\leftarrow 0.55 + 0.05(0 - 1) \times 0 = 0.55 \\ w_2 &\leftarrow 0.55 + 0.05(0 - 1) \times 1 = 0.5 \end{aligned}$$

$$\begin{aligned} b &\leftarrow b + 0.05(0 - 1) \times 1 \\ \textcircled{2} \quad w_1 &\leftarrow w_1 + 0.05(0 - 1) \times 1 \\ w_2 &\leftarrow w_2 + 0.05(0 - 1) \times 0 \end{aligned}$$

$$\begin{aligned} b &\leftarrow -0.65 + 0.05(0 - 1) \times 1 = -0.7 \\ w_1 &\leftarrow 0.55 + 0.05(0 - 1) \times 1 = 0.5 \\ w_2 &\leftarrow 0.55 + 0.05(0 - 1) \times 0 = 0.55 \end{aligned}$$

Unit 01 | Perceptron

Perceptron의 한계

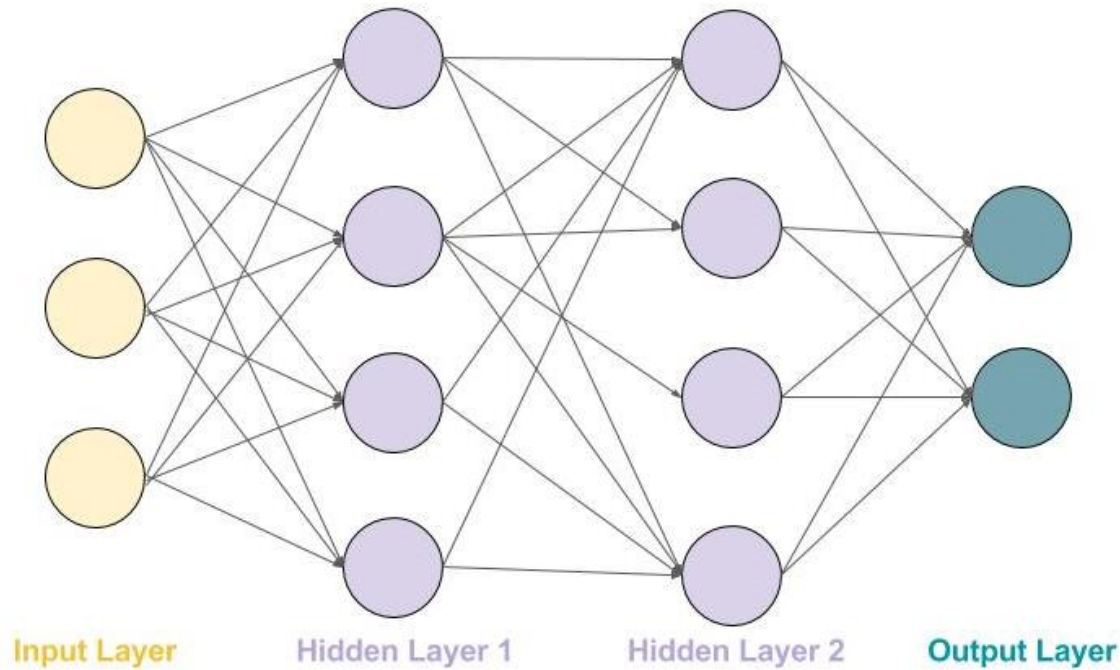


1969년 MIT AI랩 창립자 Minsky & Papert
“현재의 퍼셉트론으로는 XOR 문제를 해결할 수 없다.”

02 | Backpropagation

Unit 02 | Backpropagation

Multi Layer Perceptron



MLP를 학습시키는 방법
: 오류 역전파

Unit 02 | Backpropagation

역전파(Backpropagation)

순전파(Feedforward) 알고리즘에서 발생한 오차를 줄이기 위해
새로운 가중치를 업데이트하고, 새로운 가중치로 다시 학습하는 과정

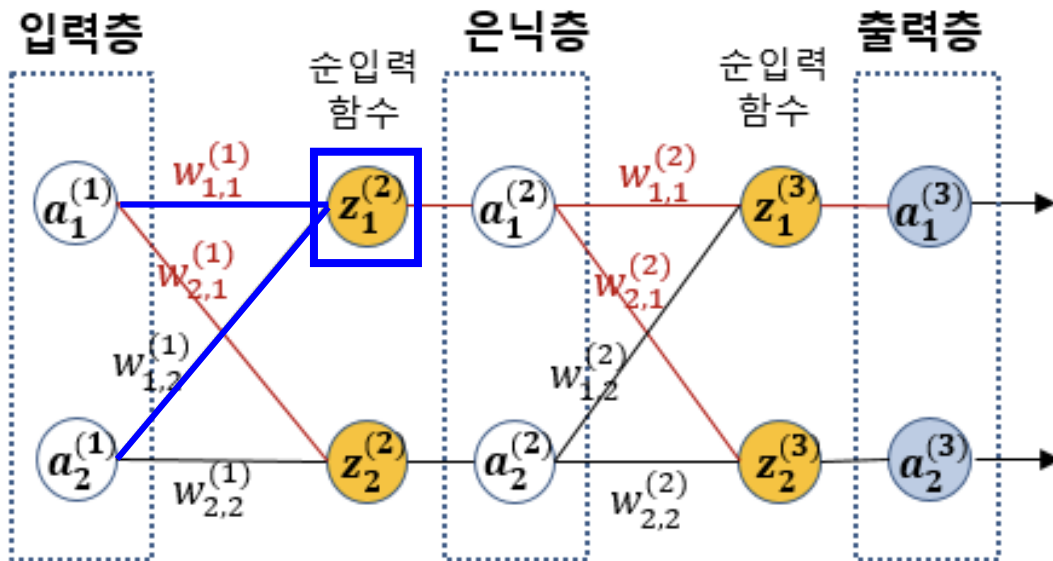
Unit 02 | Backpropagation

역전파(Backpropagation)

순전파(Feedforward) 알고리즘에서 발생한 오차를 줄이기 위해
새로운 가중치를 업데이트하고, 새로운 가중치로 다시 학습하는 과정

Unit 02 | Backpropagation

순전파(Feedforward)



$$\phi(z) = \frac{1}{1 + e^{-z}}$$

$$z_1^{(2)} = w_{1,1}^{(1)} a_1^{(1)} + w_{1,2}^{(1)} a_2^{(1)}$$

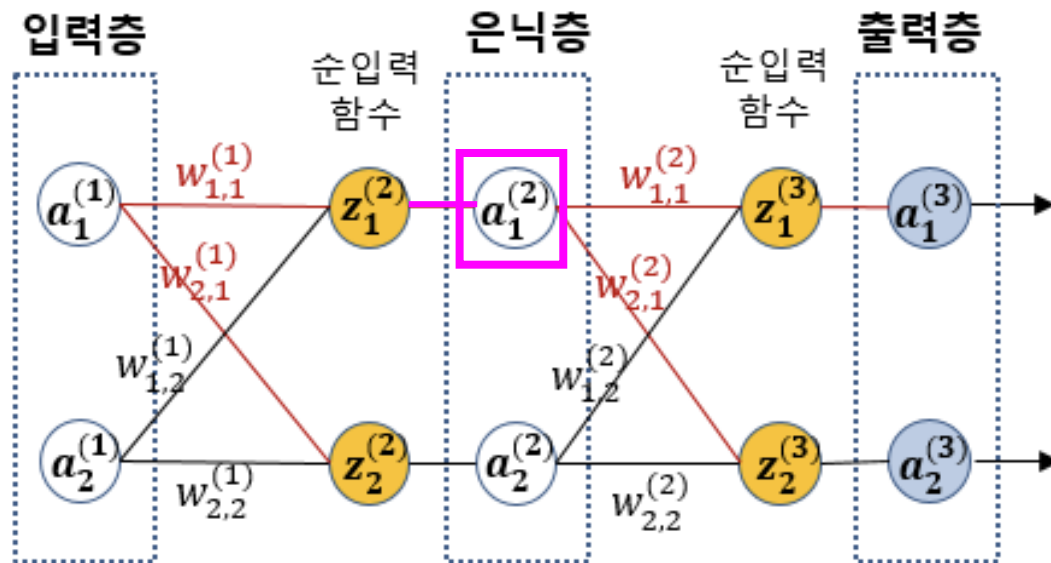
$$a_1^{(2)} = \phi(z_1^{(2)})$$

$$z_1^{(3)} = w_{1,1}^{(2)} a_1^{(2)} + w_{1,2}^{(2)} a_2^{(2)}$$

$$a_1^{(3)} = \phi(z_1^{(3)})$$

Unit 02 | Backpropagation

순전파(Feedforward)



$$\phi(z) = \frac{1}{1 + e^{-z}}$$

$$z_1^{(2)} = w_{1,1}^{(1)} a_1^{(1)} + w_{1,2}^{(1)} a_2^{(1)}$$

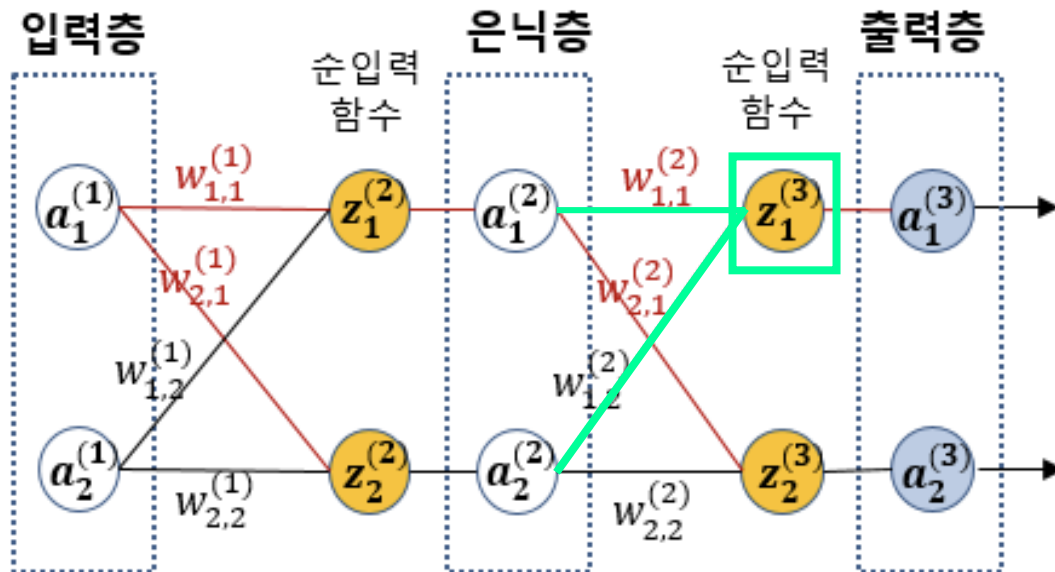
$$a_1^{(2)} = \phi(z_1^{(2)})$$

$$z_1^{(3)} = w_{1,1}^{(2)} a_1^{(2)} + w_{1,2}^{(2)} a_2^{(2)}$$

$$a_1^{(3)} = \phi(z_1^{(3)})$$

Unit 02 | Backpropagation

순전파(Feedforward)



$$\phi(z) = \frac{1}{1 + e^{-z}}$$

$$z_1^{(2)} = w_{1,1}^{(1)} a_1^{(1)} + w_{1,2}^{(1)} a_2^{(1)}$$

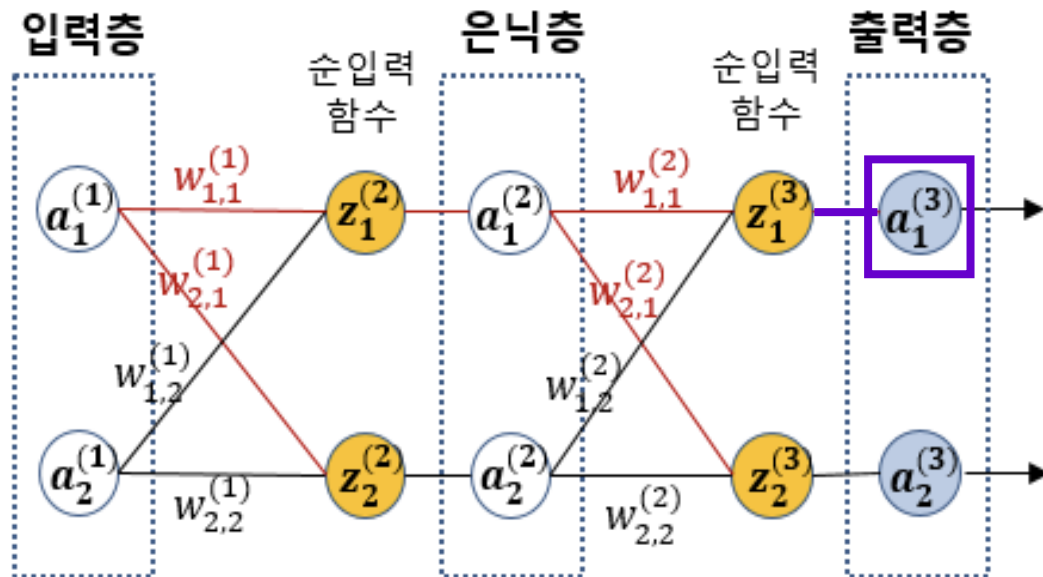
$$a_1^{(2)} = \phi(z_1^{(2)})$$

$$z_1^{(3)} = w_{1,1}^{(2)} a_1^{(2)} + w_{1,2}^{(2)} a_2^{(2)}$$

$$a_1^{(3)} = \phi(z_1^{(3)})$$

Unit 02 | Backpropagation

순전파(Feedforward)



$$\phi(z) = \frac{1}{1 + e^{-z}}$$

$$z_1^{(2)} = w_{1,1}^{(1)} a_1^{(1)} + w_{1,2}^{(1)} a_2^{(1)}$$

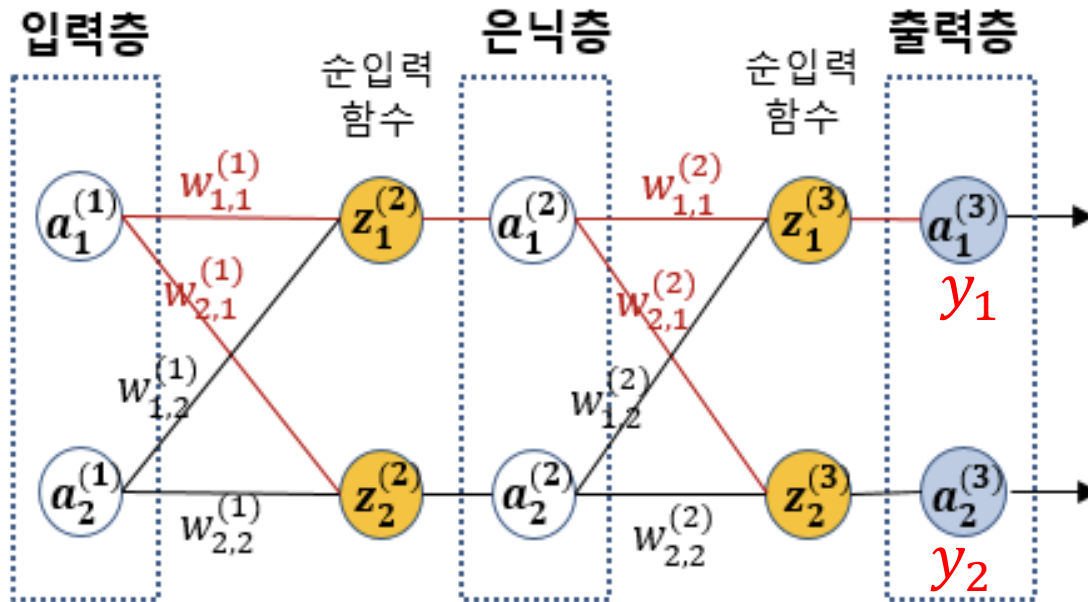
$$a_1^{(2)} = \phi(z_1^{(2)})$$

$$z_1^{(3)} = w_{1,1}^{(2)} a_1^{(2)} + w_{1,2}^{(2)} a_2^{(2)}$$

$$a_1^{(3)} = \phi(z_1^{(3)})$$

Unit 02 | Backpropagation

손실함수(Cost Function)



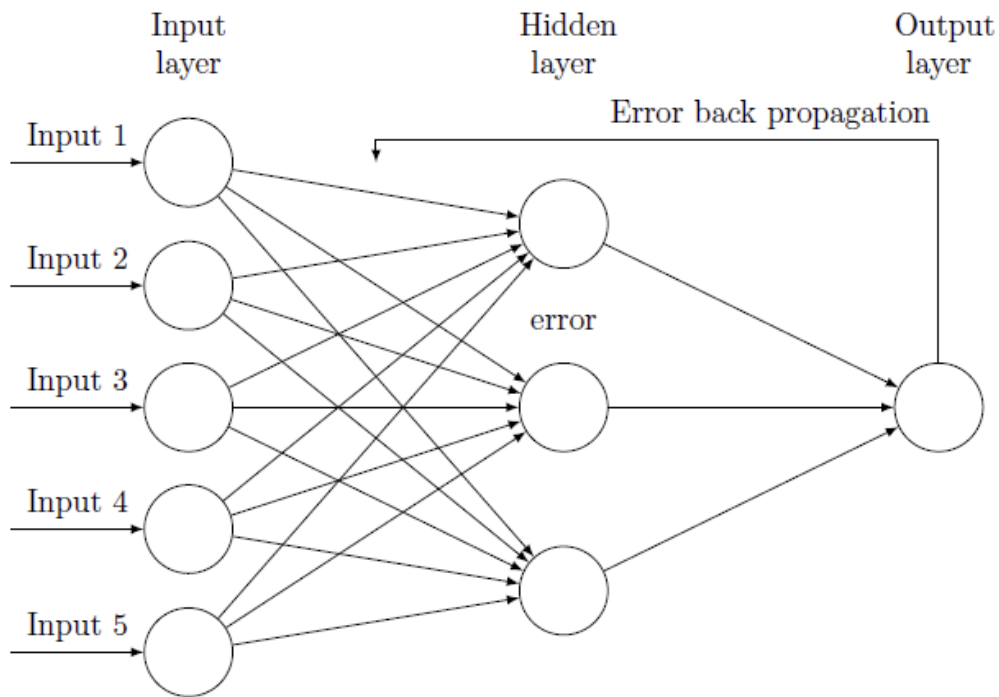
$$\text{MSE} = \frac{1}{2N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

$$J_1 = \frac{1}{2} (a_1^{(3)} - y_1)^2$$

$$J_2 = \frac{1}{2} (a_2^{(3)} - y_2)^2$$

Unit 02 | Backpropagation

역전파(Backpropagation)



- Input과 output 값을 알고 있는 상태에서 신경망을 학습시키는 방법
- 출력부터 반대 방향으로 순차적으로 편미분을 수행해 가면서 weight와 bias 값을 갱신시킴

$$w_j = w_j - \eta \frac{\partial J_{total}}{\partial w_j}$$

가중치 업데이트 식

Unit 02 | Backpropagation

편미분

다변수함수의 특정 변수를 제외한 나머지 변수를 상수로 생각하여 미분

$$z = f(x, y) = x^2 + xy + y^2$$

$$\frac{\partial z}{\partial x} = 2x + y, \quad \frac{\partial z}{\partial y} = 2y + x$$

$$\Delta f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (2x + y, 2y + x)$$

Unit 02 | Backpropagation

Chain Rule

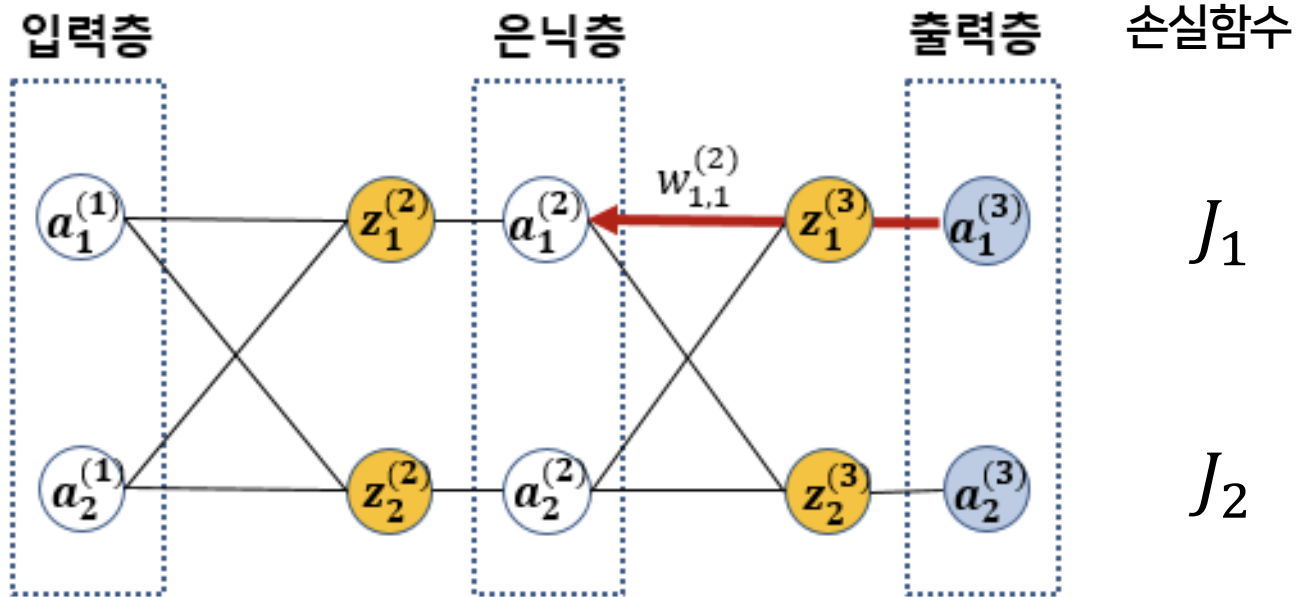
연쇄 법칙, 합성 함수를 미분할 때의 계산 공식

$$f(g(x))' = f'(g(x))g'(x)$$

$$y = f(x), u = g(x) \text{일 때, } \frac{\partial y}{\partial x} = \frac{\partial y}{\partial u} \cdot \frac{\partial u}{\partial x} \text{ 성립}$$

Unit 02 | Backpropagation

역전파(Backpropagation)



$$w_j = w_j - \eta \frac{\partial J_{total}}{\partial w_j}$$

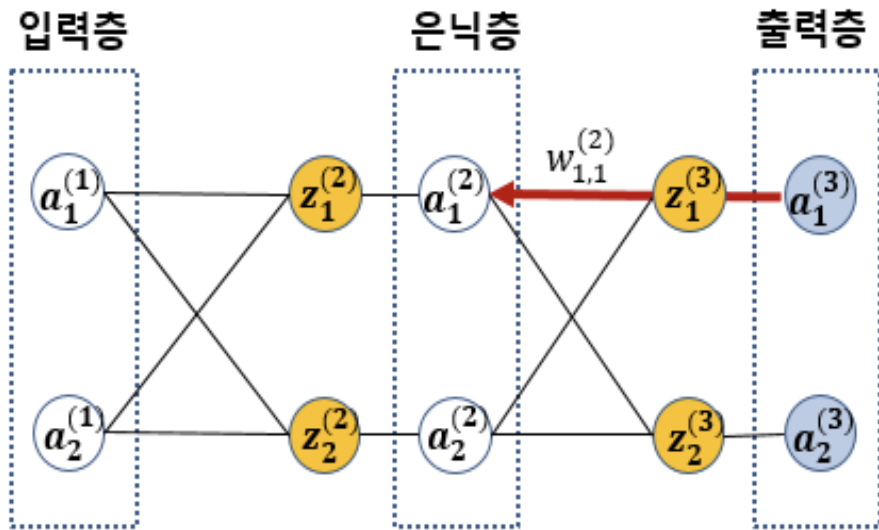
$$w_{1,1}^{(2)} = w_{1,1}^{(2)} - \frac{\partial J_{total}}{\partial w_{1,1}^{(2)}}$$

역전파의 출발노드인 $a_1^{(3)}$ 의 J_{total} 은 J_1

$$\frac{\partial J_{total}}{\partial w_{1,1}^{(2)}} = \frac{\partial J_1}{\partial a_1^{(3)}} \times \frac{\partial a_1^{(3)}}{\partial z_1^{(3)}} \times \frac{\partial z_1^{(3)}}{\partial w_{1,1}^{(2)}}$$

Unit 02 | Backpropagation

역전파(Backpropagation)



손실함수

$$J_1$$

$$J_2$$

$$\frac{\partial J_{total}}{\partial w_{1,1}^{(2)}} = \frac{\partial J_1}{\partial a_1^{(3)}} \times \frac{\partial a_1^{(3)}}{\partial z_1^{(3)}} \times \frac{\partial z_1^{(3)}}{\partial w_{1,1}^{(2)}}$$

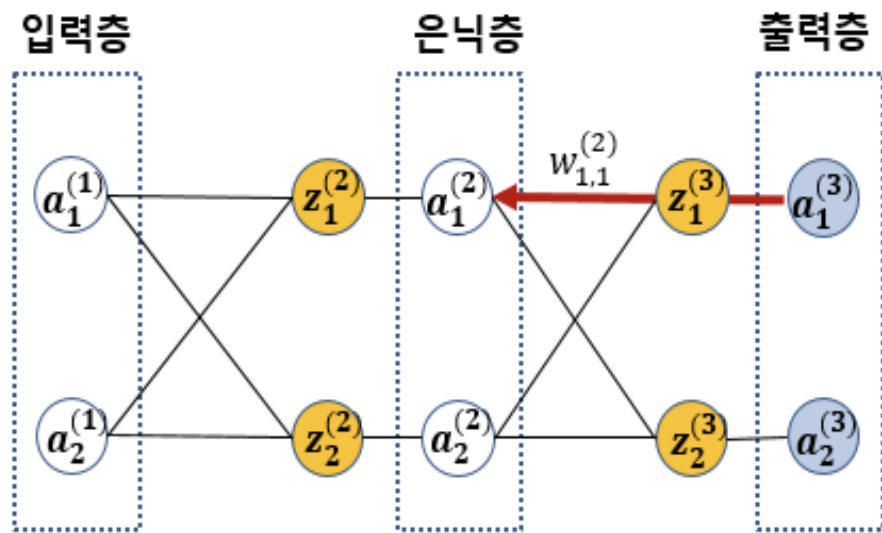
① ② ③

참고: $J_1 = \frac{1}{2} (a_1^{(3)} - y_1)^2$

$$\textcircled{1} \quad \frac{\partial J_1}{\partial a_1^{(3)}} = \frac{1}{2} \frac{\partial}{\partial a_1^{(3)}} (a_1^{(3)} - y_1)^2 = (a_1^{(3)} - y_1)$$

Unit 02 | Backpropagation

역전파(Backpropagation)

 J_1 J_2

$$\frac{\partial J_{total}}{\partial w_{1,1}^{(2)}} = \frac{\partial J_1}{\partial a_1^{(3)}} \times \frac{\partial a_1^{(3)}}{\partial z_1^{(3)}} \times \frac{\partial z_1^{(3)}}{\partial w_{1,1}^{(2)}}$$

①

②

③

참고: $a_1^{(3)} = \phi(z_1^{(3)})$ $\sigma'(x) = \frac{\partial}{\partial x} \frac{1}{1+e^{-x}}$

$$= \frac{e^{-x}}{(1+e^{-x})^2}$$

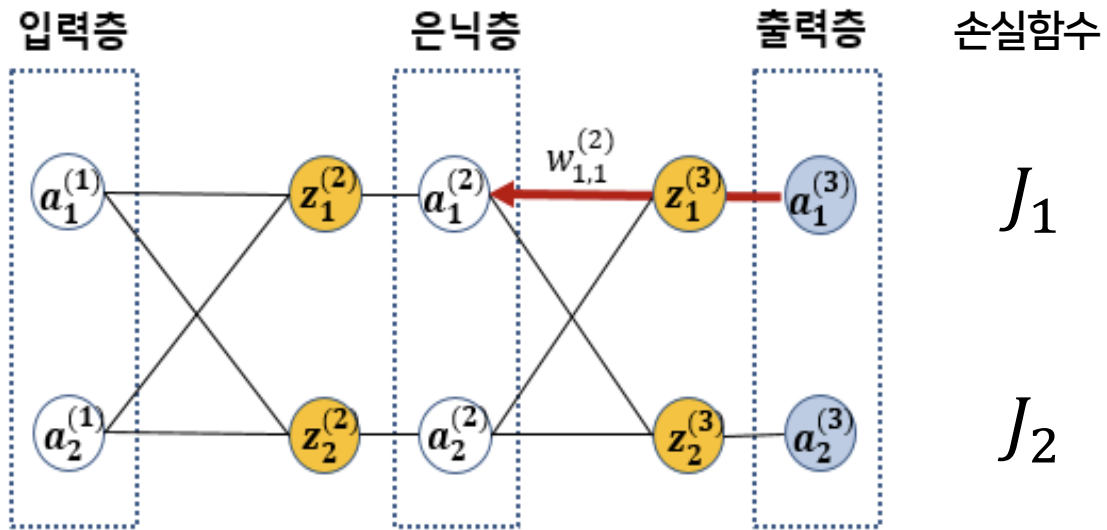
$$= \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}}$$

$$= \sigma(x)(1 - \sigma(x))$$

② $\frac{\partial a_1^{(3)}}{\partial z_1^{(3)}} = \phi(z_1^{(3)}) (1 - \phi(z_1^{(3)})) = a_1^{(3)} (1 - a_1^{(3)})$

Unit 02 | Backpropagation

역전파(Backpropagation)



$$\frac{\partial J_{total}}{\partial w_{1,1}^{(2)}} = \underbrace{\frac{\partial J_1}{\partial a_1^{(3)}}}_{\textcircled{1}} \times \underbrace{\frac{\partial a_1^{(3)}}{\partial z_1^{(3)}}}_{\textcircled{2}} \times \underbrace{\frac{\partial z_1^{(3)}}{\partial w_{1,1}^{(2)}}}_{\textcircled{3}}$$

참고: $z_1^{(3)} = w_{1,1}^{(2)} a_1^{(2)} + w_{1,2}^{(2)} a_2^{(2)}$

$\textcircled{3} \quad \frac{\partial z_1^{(3)}}{\partial w_{1,1}^{(2)}} = a_1^{(2)}$

Unit 02 | Backpropagation

역전파(Backpropagation)

$$\frac{\partial J_{total}}{\partial w_{1,1}^{(2)}} = \frac{\partial J_1}{\partial a_1^{(3)}} \times \frac{\partial a_1^{(3)}}{\partial z_1^{(3)}} \times \frac{\partial z_1^{(3)}}{\partial w_{1,1}^{(2)}} = \overset{\textcircled{1}}{(a_1^{(3)} - y_1)} \times \overset{\textcircled{2}}{a_1^{(3)}(1 - a_1^{(3)})} \times \overset{\textcircled{3}}{a_1^{(2)}}$$

$$\delta_1^{(3)} = \frac{\partial J_1}{\partial z_1^{(3)}} = (a_1^{(3)} - y_1) \times a_1^{(3)}(1 - a_1^{(3)})$$

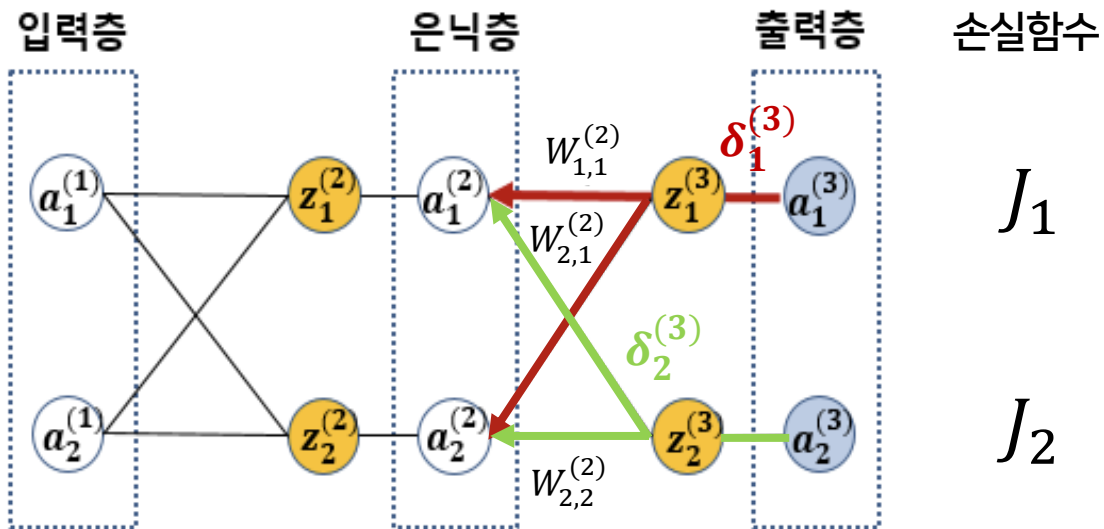
$$w_{1,1}^{(2)} = w_{1,1}^{(2)} - \frac{\partial J_{total}}{\partial w_{1,1}^{(2)}} = w_{1,1}^{(2)} - \delta_1^{(3)} a_1^{(2)}$$

Unit 02 | Backpropagation

역전파(Backpropagation)

$$\delta_1^{(3)} = \frac{\partial J_1}{\partial z_1^{(3)}} = (a_1^{(3)} - y_1) \times a_1^{(3)}(1 - a_1^{(3)})$$

$$w_{1,1}^{(2)} = w_{1,1}^{(2)} - \frac{\partial J_{total}}{\partial w_{1,1}^{(2)}} = w_{1,1}^{(2)} - \delta_1^{(3)} a_1^{(2)}$$



같은 방식으로

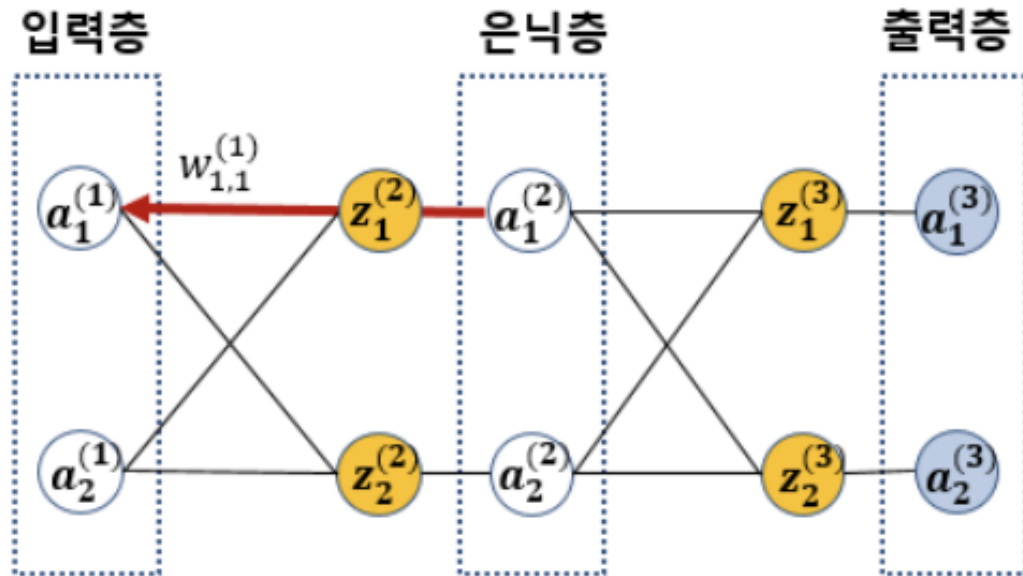
$$\delta_2^{(3)} = \frac{\partial J_2}{\partial z_2^{(3)}} = (a_2^{(3)} - y_2) \times a_2^{(3)}(1 - a_2^{(3)})$$

$$w_{2,1}^{(2)} = w_{2,1}^{(2)} - \delta_2^{(3)} a_1^{(2)}$$

$$w_{2,2}^{(2)} = w_{2,2}^{(2)} - \delta_2^{(3)} a_2^{(2)}$$

Unit 02 | Backpropagation

역전파(Backpropagation)

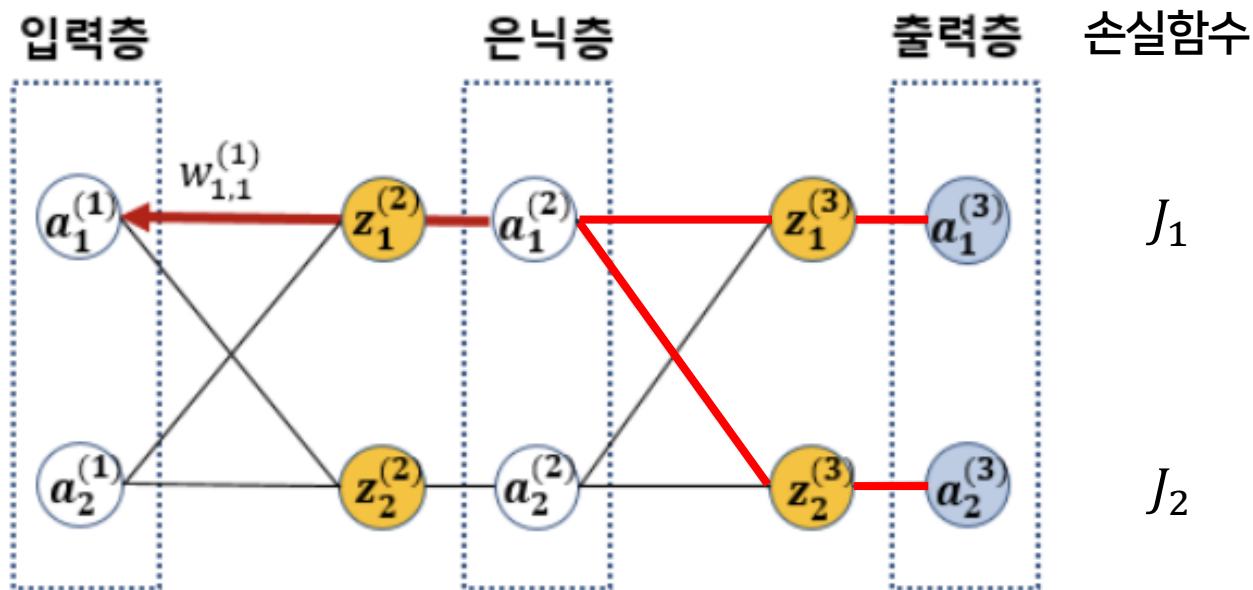


$$w_j = w_j - \eta \frac{\partial J_{total}}{\partial w_j}$$

$$\frac{\partial J_{total}}{\partial w_{1,1}^{(1)}} = \frac{\partial J_{total}}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial w_{1,1}^{(1)}}$$

Unit 02 | Backpropagation

역전파(Backpropagation)



$$w_j = w_j - \eta \frac{\partial J_{total}}{\partial w_j}$$

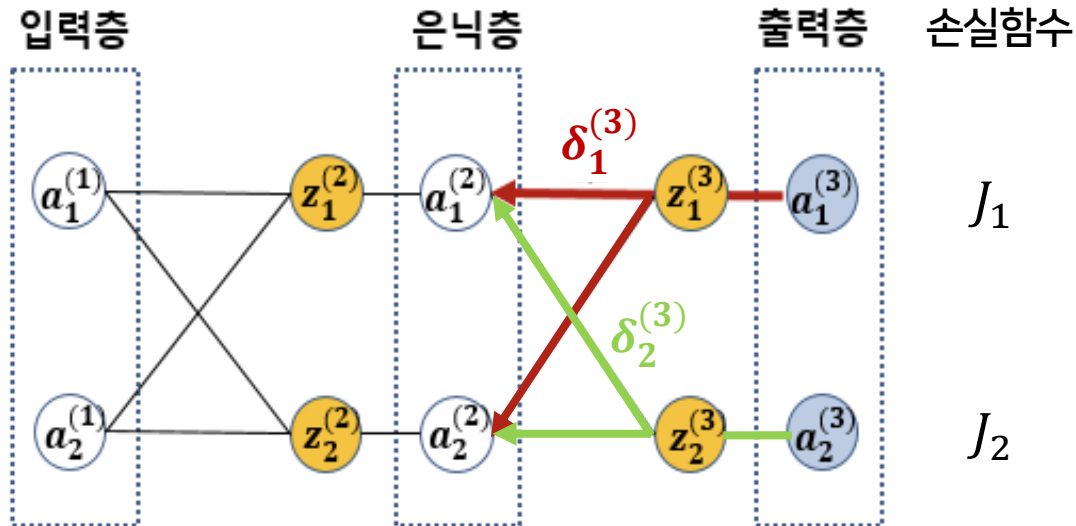
$$\frac{\partial J_{total}}{\partial w_{1,1}^{(1)}} = \frac{\partial J_{total}}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial w_{1,1}^{(1)}}$$

$a_1^{(2)}$ 의 J_{total} 은 $J_1 + J_2$

$$\frac{\partial J_{total}}{\partial a_1^{(2)}} = \frac{\partial J_1}{\partial a_1^{(2)}} + \frac{\partial J_2}{\partial a_1^{(2)}} = \frac{\partial J_1}{\partial z_1^{(3)}} \times \frac{\partial z_1^{(3)}}{\partial a_1^{(2)}} + \frac{\partial J_2}{\partial z_2^{(3)}} \times \frac{\partial z_2^{(3)}}{\partial a_1^{(2)}}$$

Unit 02 | Backpropagation

역전파(Backpropagation)



$$\frac{\partial J_{total}}{\partial w_{1,1}^{(1)}} = \frac{\partial J_{total}}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial w_{1,1}^{(1)}}$$

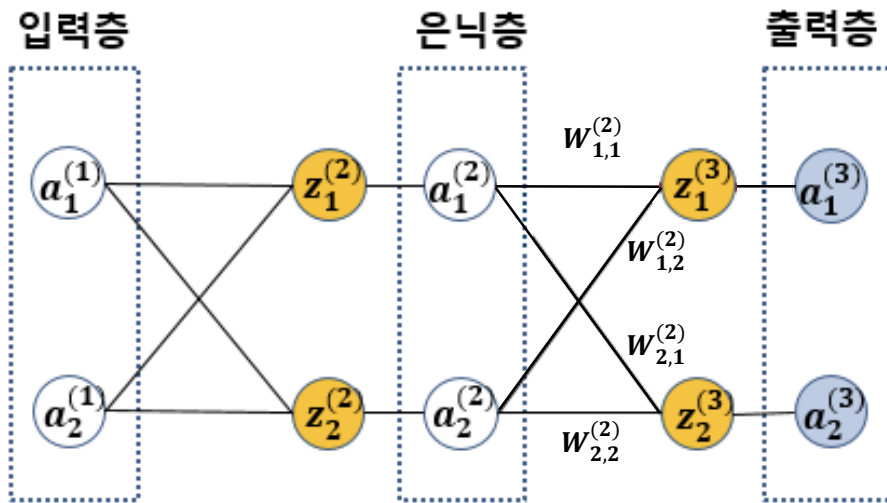
$$\delta_1^{(3)} = \frac{\partial J_1}{\partial z_1^{(3)}} = (a_1^{(3)} - y_1) \times a_1^{(3)} (1 - a_1^{(3)})$$

$$\delta_2^{(3)} = \frac{\partial J_2}{\partial z_2^{(3)}} = (a_2^{(3)} - y_2) \times a_2^{(3)} (1 - a_2^{(3)})$$

$$\begin{aligned} \frac{\partial J_{total}}{\partial a_1^{(2)}} &= \frac{\partial J_1}{\partial a_1^{(2)}} + \frac{\partial J_2}{\partial a_1^{(2)}} = \frac{\partial J_1}{\partial z_1^{(3)}} \times \frac{\partial z_1^{(3)}}{\partial a_1^{(2)}} + \frac{\partial J_2}{\partial z_2^{(3)}} \times \frac{\partial z_2^{(3)}}{\partial a_1^{(2)}} \\ &= \delta_1^{(3)} w_{1,1}^{(2)} + \delta_2^{(3)} w_{2,1}^{(2)} \end{aligned}$$

Unit 02 | Backpropagation

역전파(Backpropagation)



$$z_1^{(3)} = w_{1,1}^{(2)} a_1^{(2)} + w_{1,2}^{(2)} a_2^{(2)}$$

$$z_2^{(3)} = w_{2,1}^{(2)} a_1^{(2)} + w_{2,2}^{(2)} a_2^{(2)}$$

$$\frac{\partial J_{total}}{\partial w_{1,1}^{(1)}} = \frac{\partial J_{total}}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial w_{1,1}^{(1)}}$$

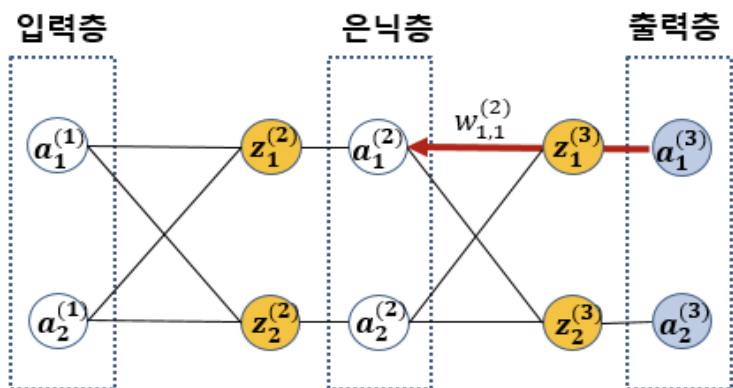
$$\delta_1^{(3)} = \frac{\partial J_1}{\partial z_1^{(3)}} = (a_1^{(3)} - y_1) \times a_1^{(3)} (1 - a_1^{(3)})$$

$$\delta_2^{(3)} = \frac{\partial J_2}{\partial z_2^{(3)}} = (a_2^{(3)} - y_2) \times a_2^{(3)} (1 - a_2^{(3)})$$

$$\begin{aligned} \frac{\partial J_{total}}{\partial a_1^{(2)}} &= \frac{\partial J_1}{\partial a_1^{(2)}} + \frac{\partial J_2}{\partial a_1^{(2)}} = \frac{\partial J_1}{\partial z_1^{(3)}} \times \frac{\partial z_1^{(3)}}{\partial a_1^{(2)}} + \frac{\partial J_2}{\partial z_2^{(3)}} \times \frac{\partial z_2^{(3)}}{\partial a_1^{(2)}} \\ &= \delta_1^{(3)} w_{1,1}^{(2)} + \delta_2^{(3)} w_{2,1}^{(2)} \end{aligned}$$

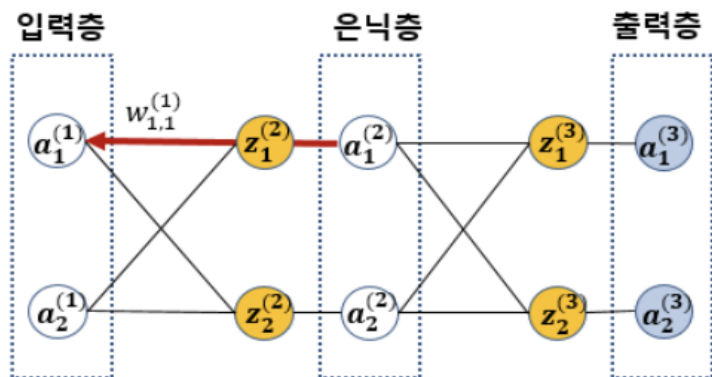
Unit 02 | Backpropagation

역전파(Backpropagation)



$$\frac{\partial J_{total}}{\partial w_{1,1}^{(2)}} = \frac{\partial J_1}{\partial a_1^{(3)}} \times \frac{\partial a_1^{(3)}}{\partial z_1^{(3)}} \times \frac{\partial z_1^{(3)}}{\partial w_{1,1}^{(2)}}$$

$$\frac{\partial J_{total}}{\partial w_{1,1}^{(2)}} = \underbrace{(a_1^{(3)} - y_1)} \times \underbrace{a_1^{(3)} (1 - a_1^{(3)})} \times \underbrace{a_1^{(2)}}$$

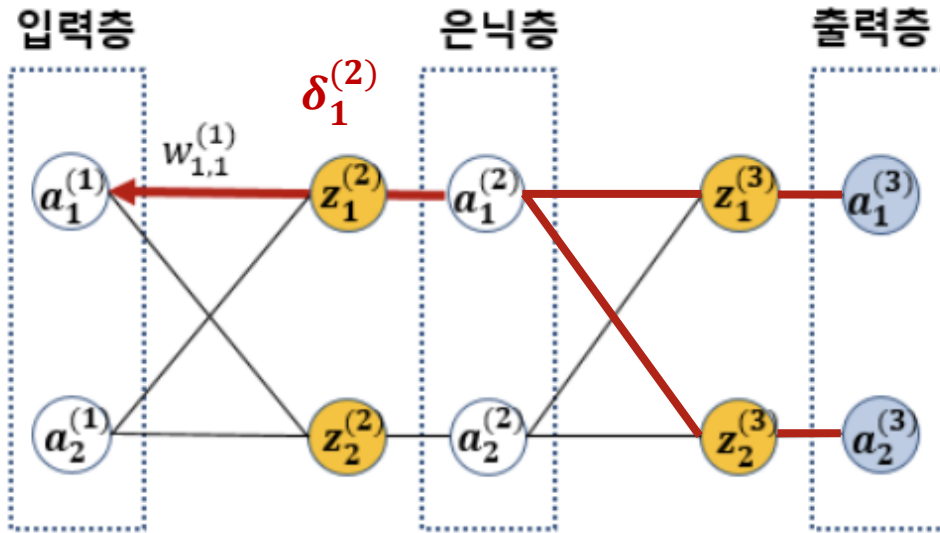


$$\frac{\partial J_{total}}{\partial w_{1,1}^{(1)}} = \frac{\partial J_{total}}{\partial a_1^{(2)}} \times \frac{\partial a_1^{(2)}}{\partial z_1^{(2)}} \times \frac{\partial z_1^{(2)}}{\partial w_{1,1}^{(1)}}$$

$$\frac{\partial J_{total}}{\partial w_{1,1}^{(1)}} = \underbrace{(\delta_1^{(3)} w_{1,1}^{(2)} + \delta_2^{(3)} w_{2,1}^{(2)})} \times \underbrace{a_1^{(2)} (1 - a_1^{(2)})} \times \underbrace{a_1^{(1)}}$$

Unit 02 | Backpropagation

역전파(Backpropagation)



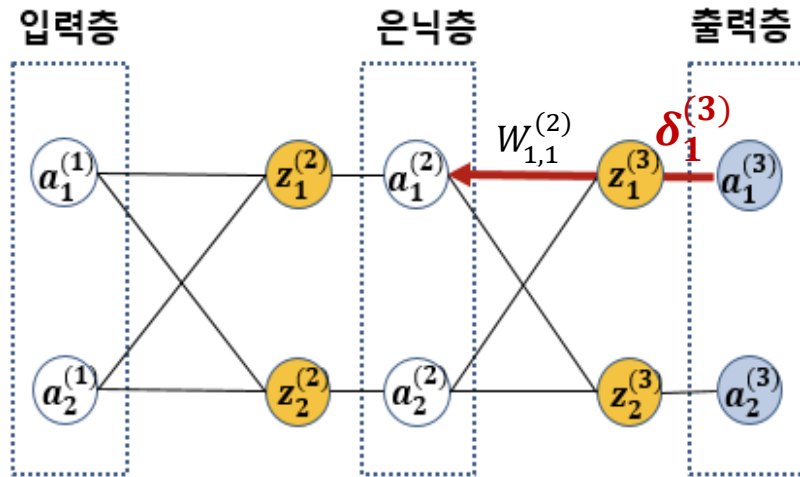
$$\frac{\partial J_{total}}{\partial w_{1,1}^{(1)}} = \left(\delta_1^{(3)} w_{1,1}^{(2)} + \delta_2^{(3)} w_{2,1}^{(2)} \right) \times a_1^{(2)} (1 - a_1^{(2)}) \times a_1^{(1)}$$

$$\delta_1^{(2)} = (\delta_1^{(3)} w_{1,1}^{(2)} + \delta_2^{(3)} w_{2,1}^{(2)}) \times a_1^{(2)} (1 - a_1^{(2)})$$

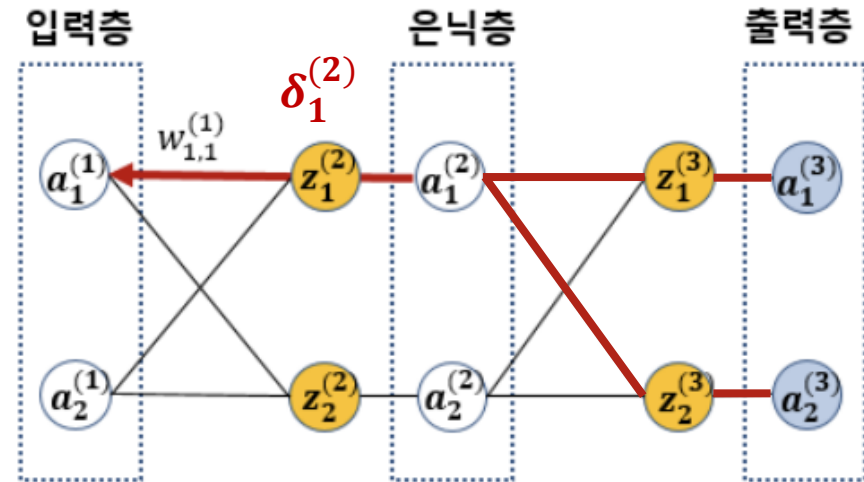
$$w_{1,1}^{(1)} = w_{1,1}^{(1)} - \frac{\partial J_{total}}{\partial w_{1,1}^{(1)}} = \boxed{w_{1,1}^{(1)} - \delta_1^{(2)} a_1^{(1)}}$$

Unit 02 | Backpropagation

역전파(Backpropagation)



$$w_{1,1}^{(2)} = w_{1,1}^{(2)} - \delta_1^{(3)} a_1^{(2)}$$



$$w_{1,1}^{(1)} = w_{1,1}^{(1)} - \delta_1^{(2)} a_1^{(1)}$$

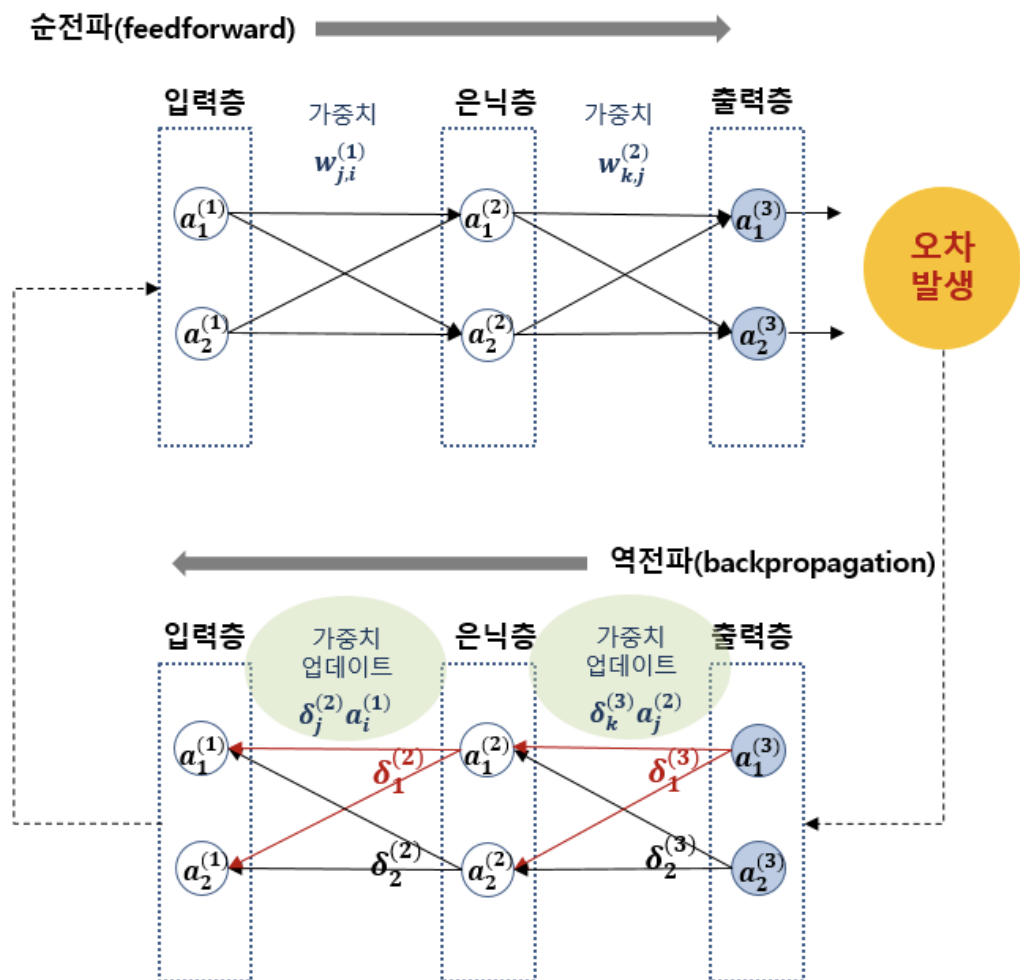
Unit 02 | Backpropagation

역전파(Backpropagation)

$$w_{j,i}^{(l)} = w_{j,i}^{(l)} - \delta_j^{(l+1)} a_i^{(l)}$$

$$\delta_j^{(3)} = (a_j^{(3)} - y_j) \times a_j^{(3)} (1 - a_j^{(3)})$$

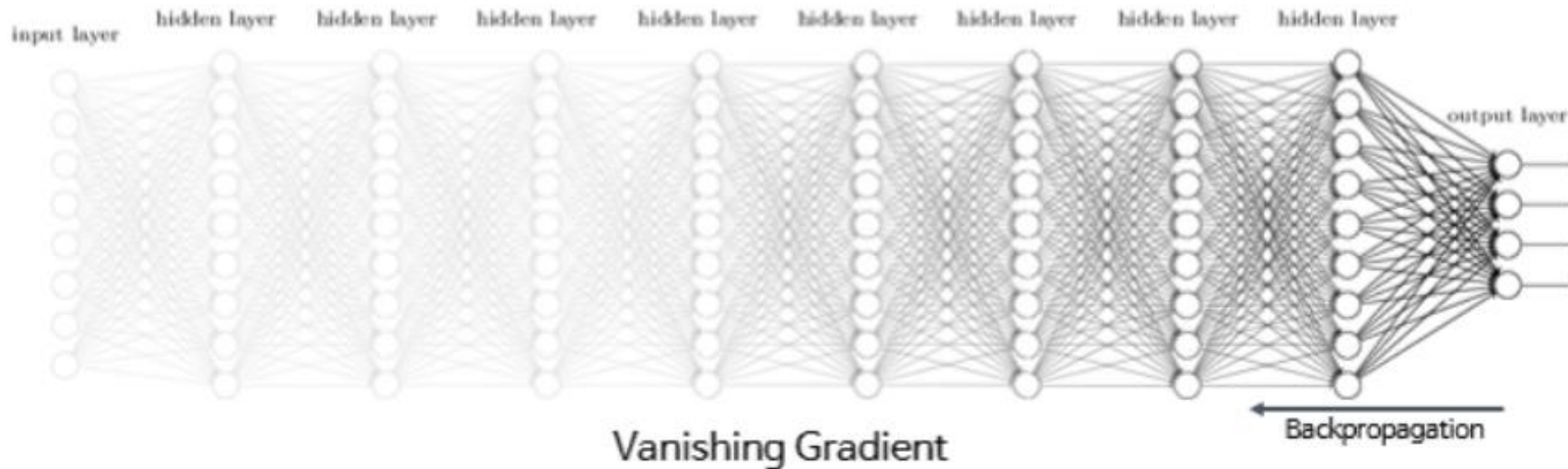
$$\delta_j^{(2)} = (\delta_1^{(3)} w_{1,j}^{(2)} + \delta_2^{(3)} w_{2,j}^{(2)}) \times a_j^{(2)} (1 - a_j^{(2)})$$



Unit 02 | Backpropagation

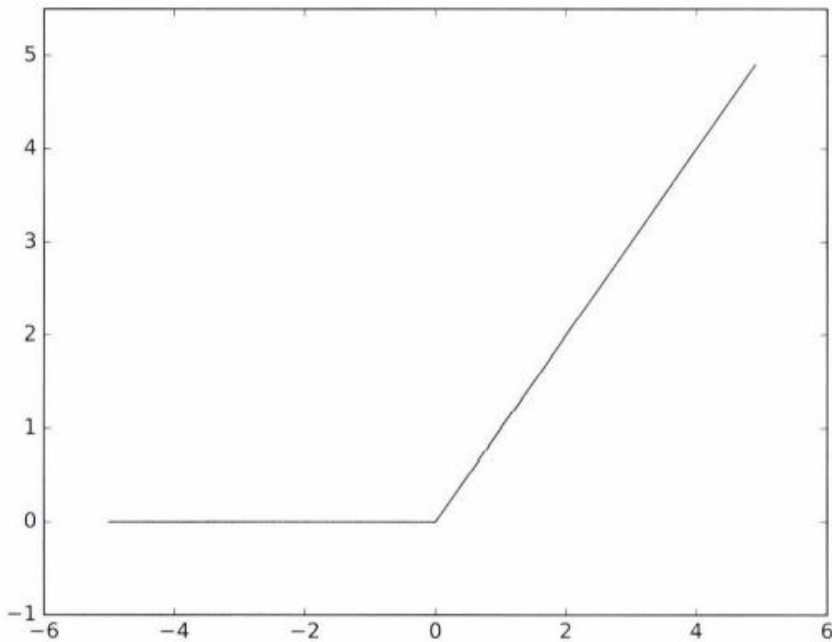
Vanishing Gradient Problem

깊이가 깊은 심층신경망에서는 역전파 알고리즘이 입력층으로 전달됨에 따라
그래디언트가 점점 작아져 결국 가중치 매개변수가 업데이트 되지 않는 경우가 발생



Unit 02 | Backpropagation

활성화함수 - ReLU 함수



$$h(x) = \begin{cases} x & (x > 0) \\ 0 & (x \leq 0) \end{cases}$$

- Vanishing gradient problem 해결!
- 입력이 0 넘으면 값이 클수록 많이 활성화
- 입력이 0보다 작으면 무조건 비활성화

Unit | 과제

“week3_NeuralNetworkBasic_assignment.pdf” 파일의
문제들을 상세한 풀이과정과 함께 풀어주세요.

Reference

참고자료

- 투빅스 15기 조준혁님 강의자료
- 밑바닥부터 시작하는 딥러닝
- 건국대학교 김학수 교수님 '기계학습' 강의
- <https://heeya-stupidbutstudying.tistory.com/38>
- <https://code4human.tistory.com/44>
- 역전파: <https://m.blog.naver.com/samsjang/221033626685>
- <https://22-22.tistory.com/35>
- <https://dailyheumsi.tistory.com/57>
- <https://m.blog.naver.com/laonple/220527647084>
- <https://m.blog.naver.com/laonple/220527647084>

Q & A

들어주셔서 감사합니다.