# BOAZ

*Bigdata is Our A to Z*

## 21-2 학기 스터디. 파이토치 마스터

4장 인공신경망

16기 김영은

# CONTENTS

01 신경망이란

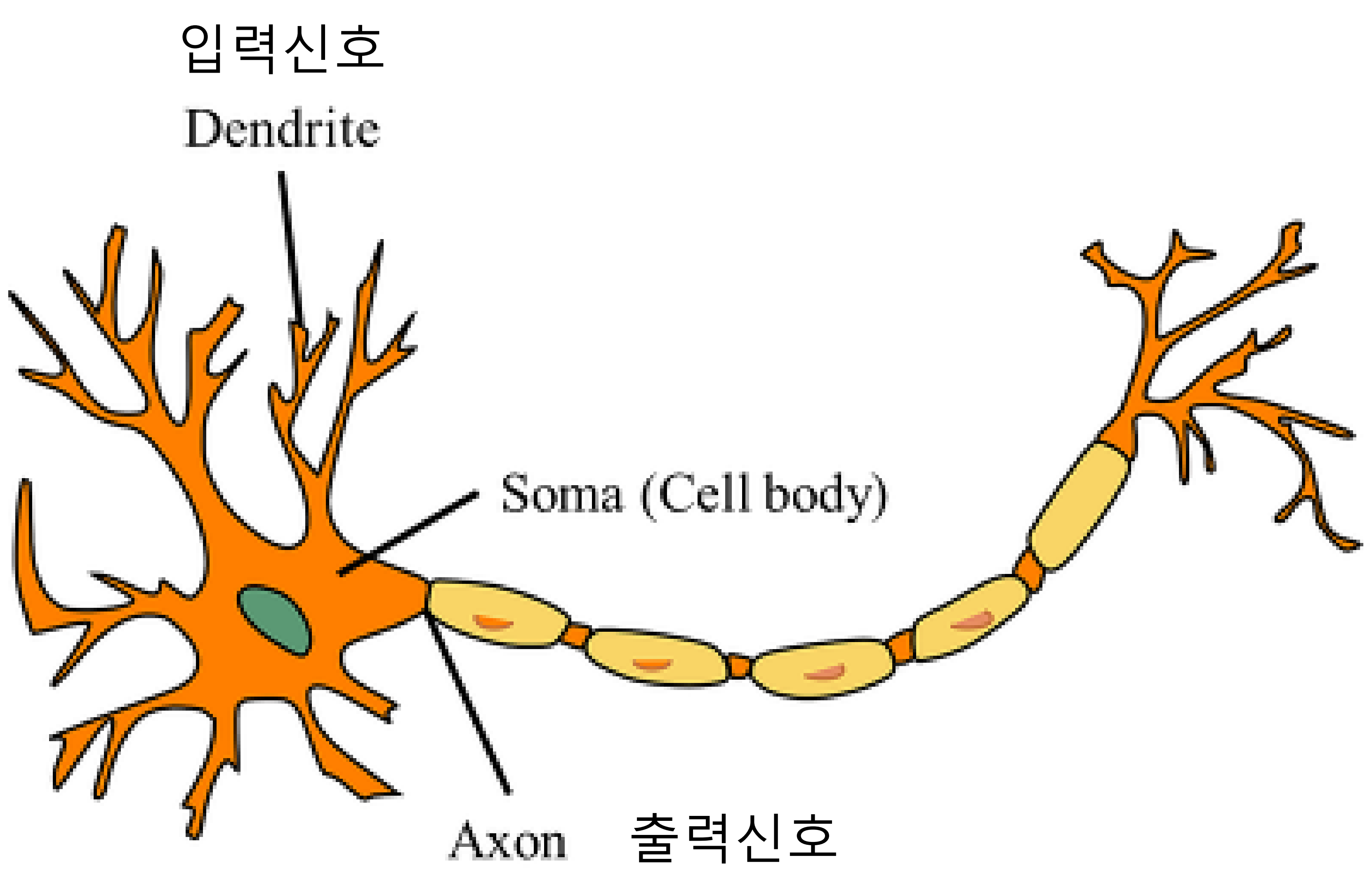
02 인공신경망 요소

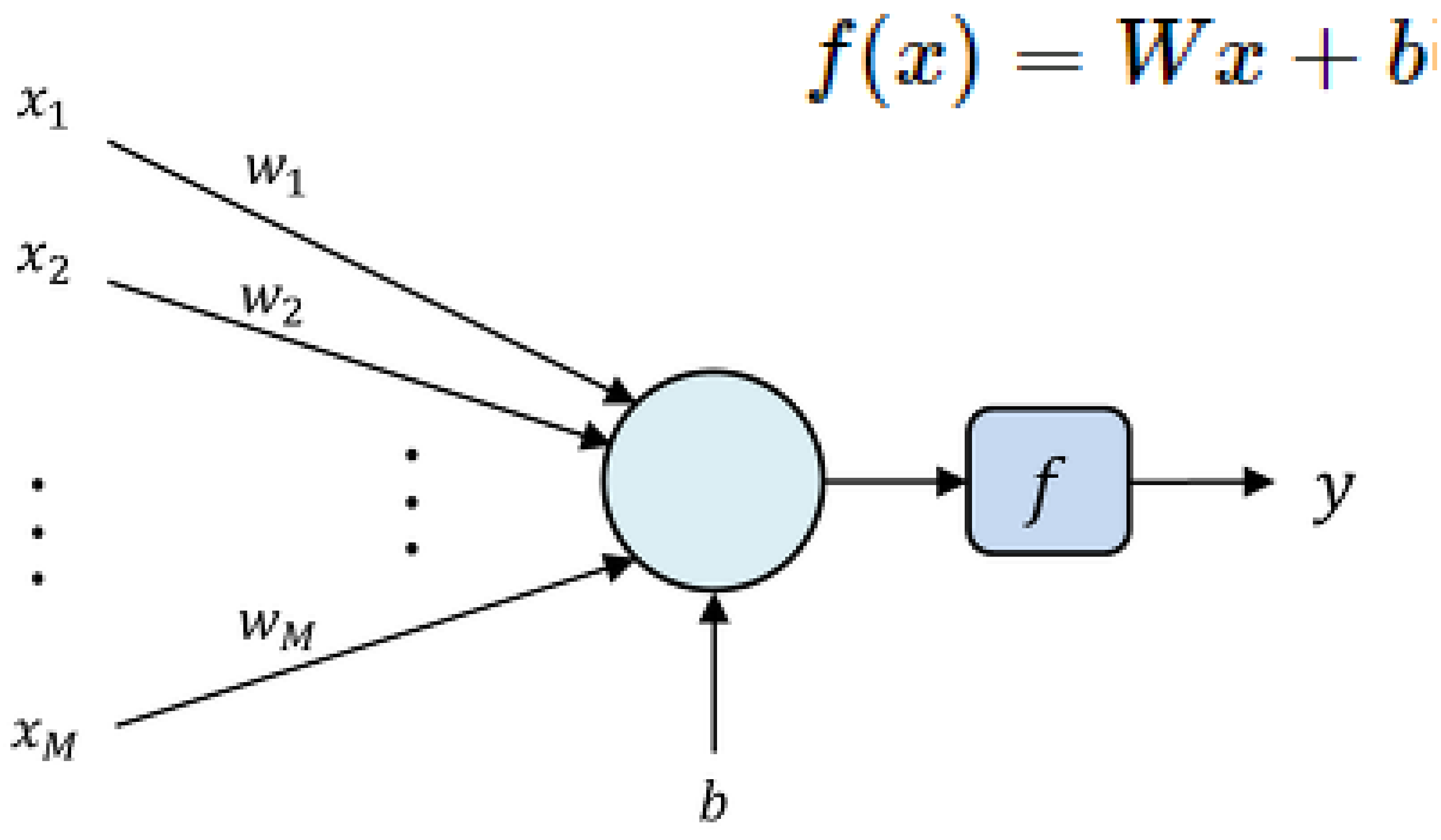
03 전파와 역전파

04 모델 구현, 학습

05 발표 및 과제 QnA

# 인공 신경망 Artificial Neural Network, ANN




**생물학적 신경망**

여러 가지돌기dendrite들을 통해 신경세포로 들어오고 어느 정도 이상의 자극이 들어오면 이를 축삭axon을 통해 다른 세포로 전달하는 구조

**인공 신경망**

여러 자극 혹은 입력x이 들어오면 각각 가중치w를 곱해 더해주고 추가적으로 편차b도 더해준다

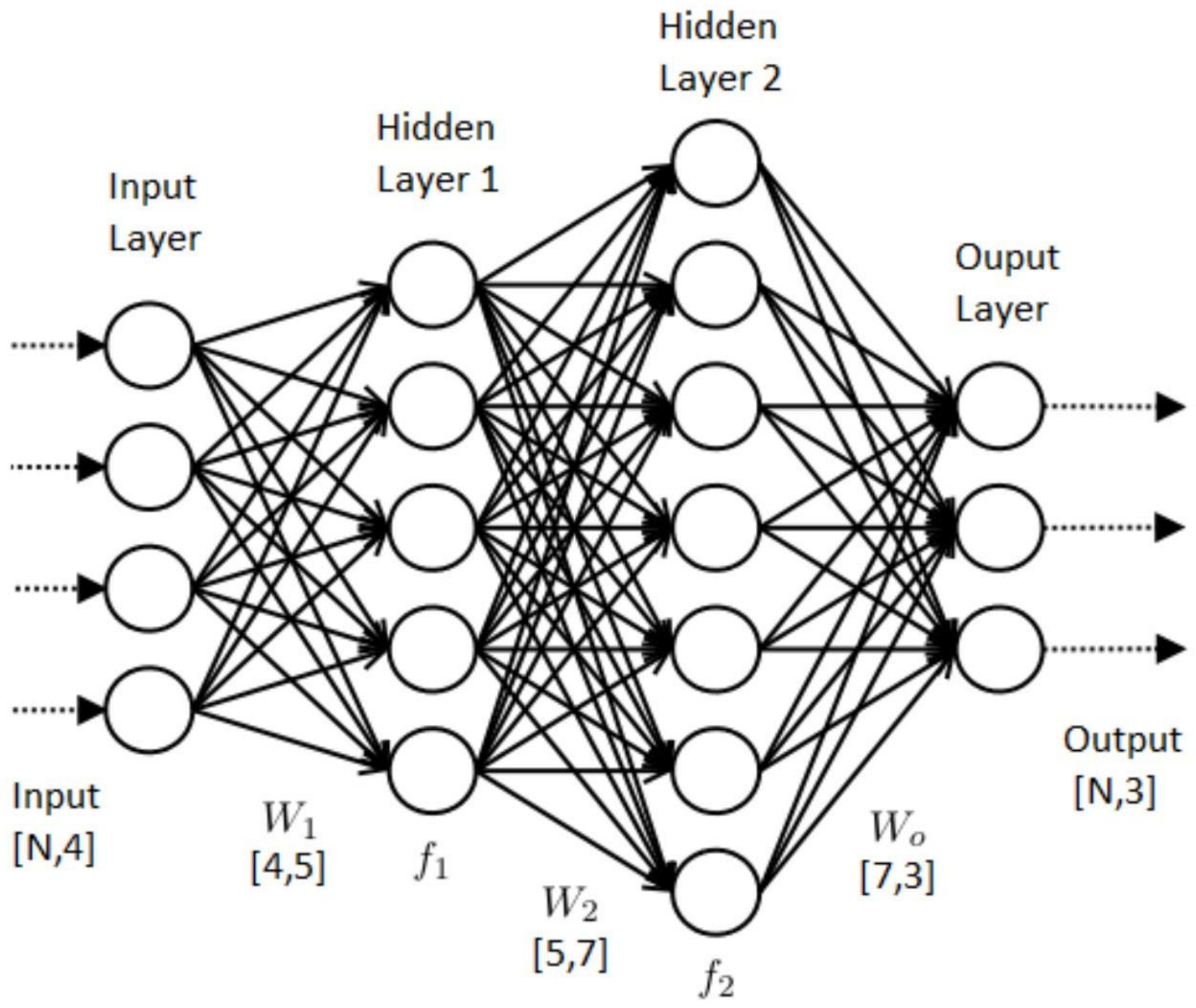
이렇게 다 더한 값을 활성화 함수f를 통해 변형하여 전달하는 단위를 인공 뉴런y이라고 하고 이러한 뉴런들이 모인 네트워크

# 인공 신경망 종류

인공신경망(ANN)은 인간의 뇌를 묘사한 기계학습 예측 모델

최근 딥러닝의 도약으로 그 성능이 재조명되고 활발하게 사용되는 중

- 퍼셉트론 Perceptron : 각 노드의 가중치 합이 0보다 크면 활성화되는 기본적인 인공신경망 알고리즘
- CNN : 시신경 구조를 모방한 인공신경망 알고리즘
- RNN : 뉴런 출력이 다시 입력으로 재귀하는 연결 구조의 인공신경망 알고리즘




https://bioinformaticsandme.tistory.com/231

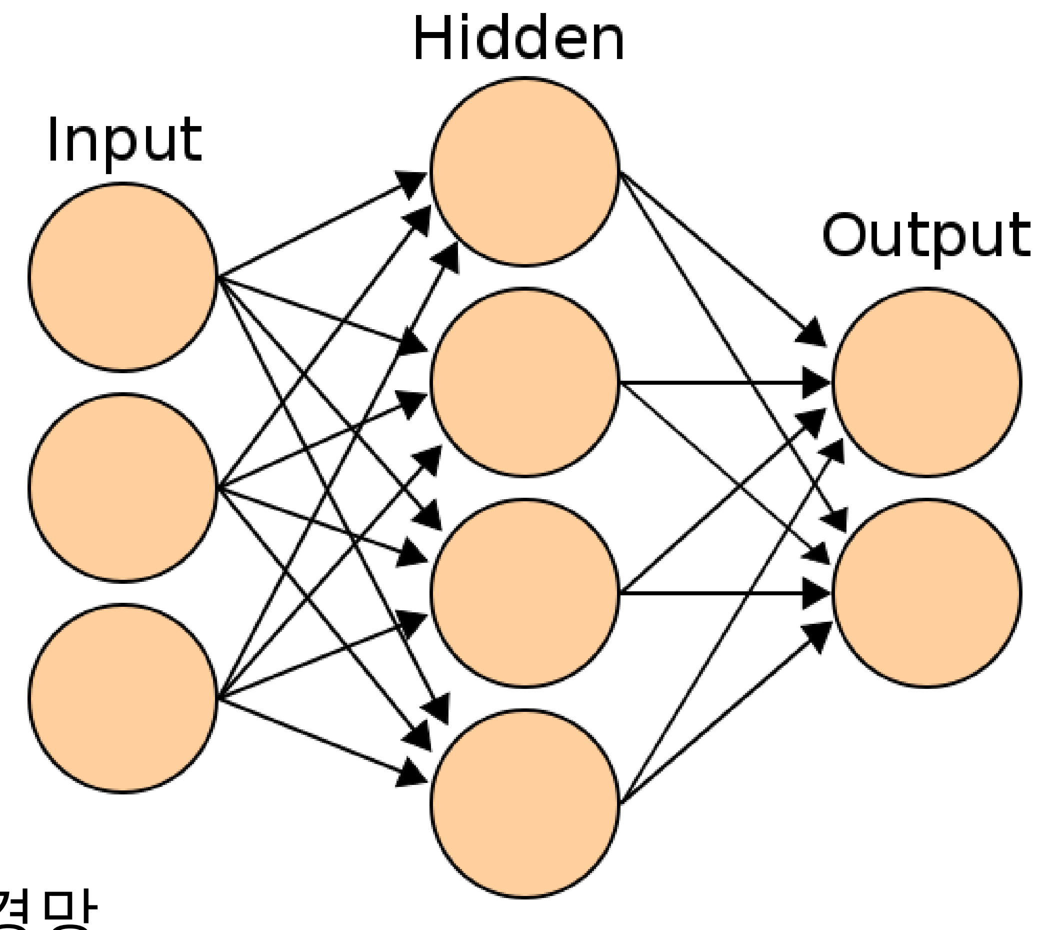
## 입력층, 은닉층, 출력층

- 입력층 : 들어온 신호를 그대로 다음 노드에 전달하는 창구 역할 ( 가중합, 활성함수 계산 X ). 시스템 외부로부터 입력자료를 받아들여 시스템으로 이들을 전송

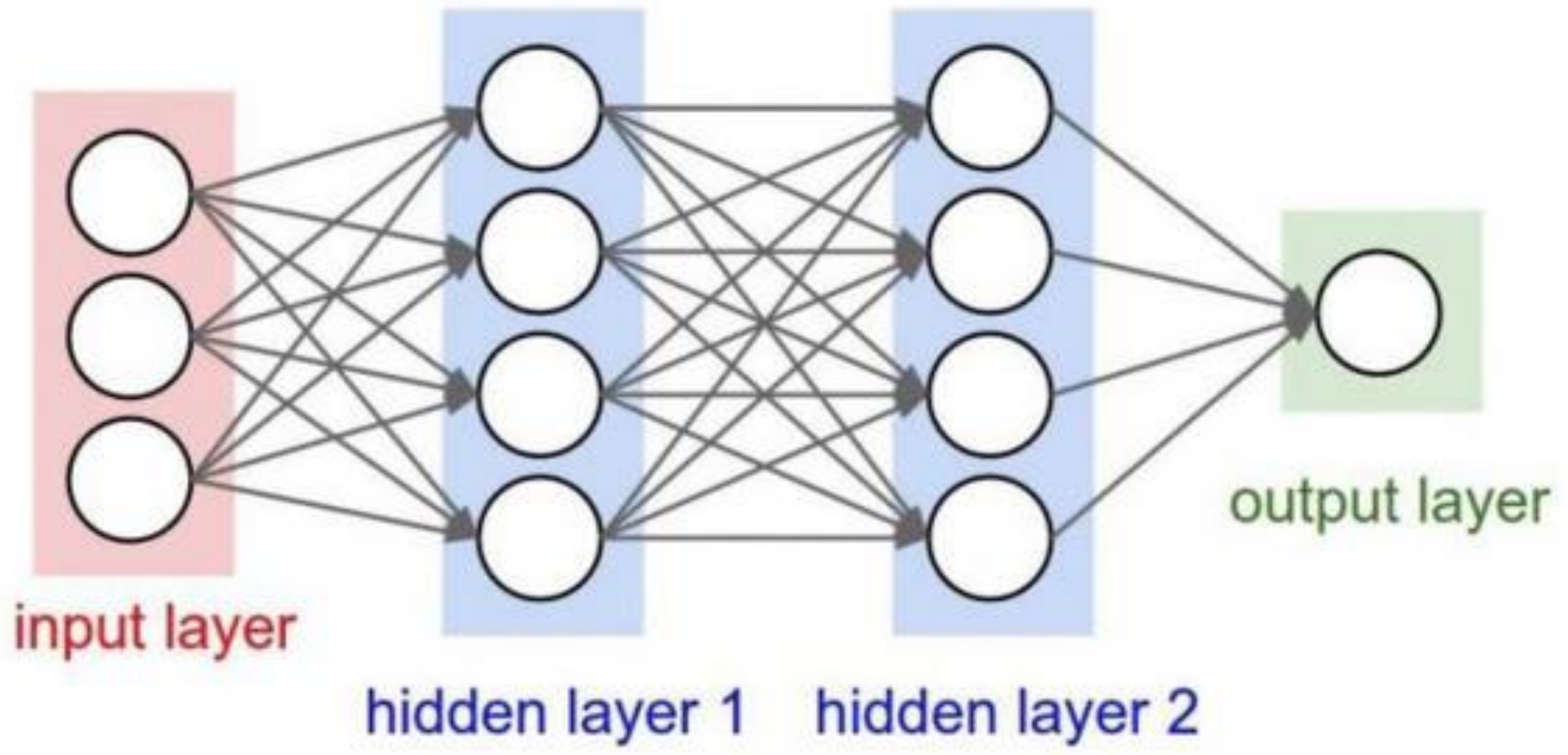
- 출력층 : 노드의 출력이 신경망의 최종 결과값이 된다. 입력 값과 현재 시스템 상태에 기준하여 시스템 출력 값을 산출

- 은닉층 : 입력층과 출력층 사이에 있는 계층들. 신경망의 외부에서는 이 계층의 노드들에 직접 접근할 수 없다. 시스템 안쪽에 자리잡고 있으며 입력 값을 넘겨받아 그것들을 처리한 뒤 결과 산출

Hidden layer의 개수에 따라
1개 ) 신경망




2개 이상) 심층신경망



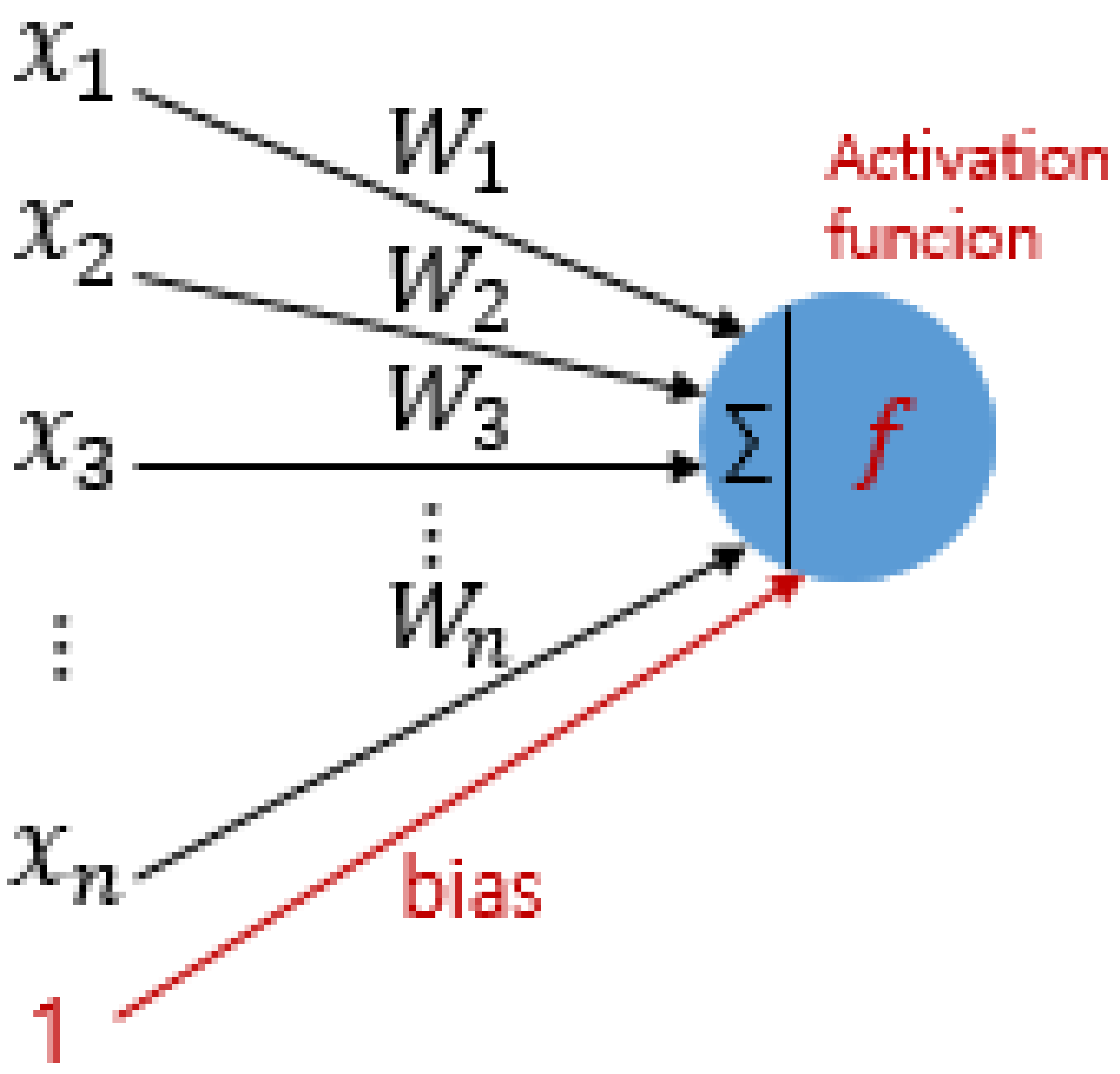

https://blog.naver.com/PostView.nhn?blogId=darda751&logNo=222278145663

## 활성화 함수 activation function

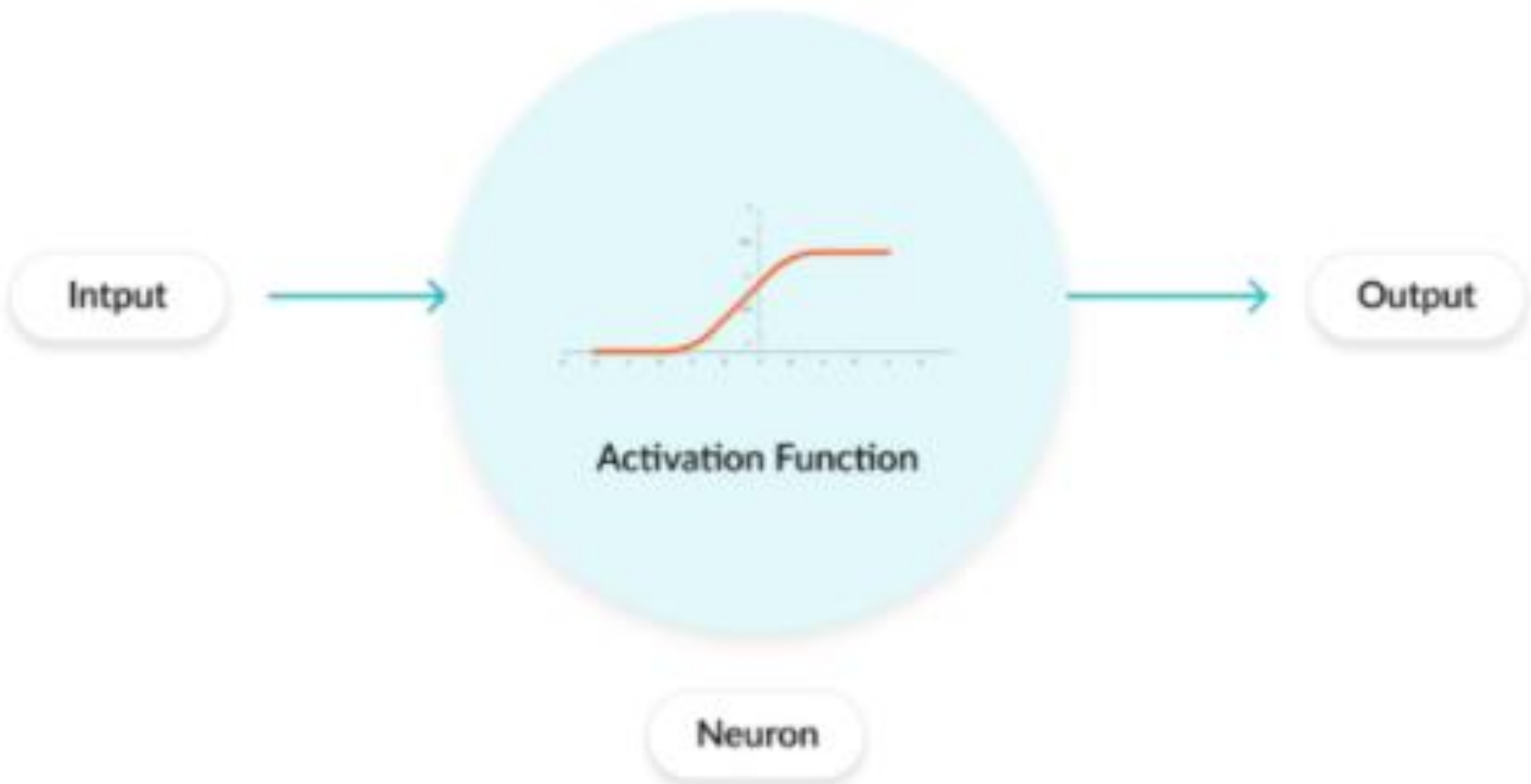
신경망의 output을 결정하는 식

현재 뉴런의 input을 feeding하여 생성된 output이 다음 레이어로 전해지는 과정 중 역할을 수행하는 수학적인 게이트

인공 신경망에서 활성화 함수는 반드시 비선형 함수여야 한다.




$$f(x) = Wx + b$$

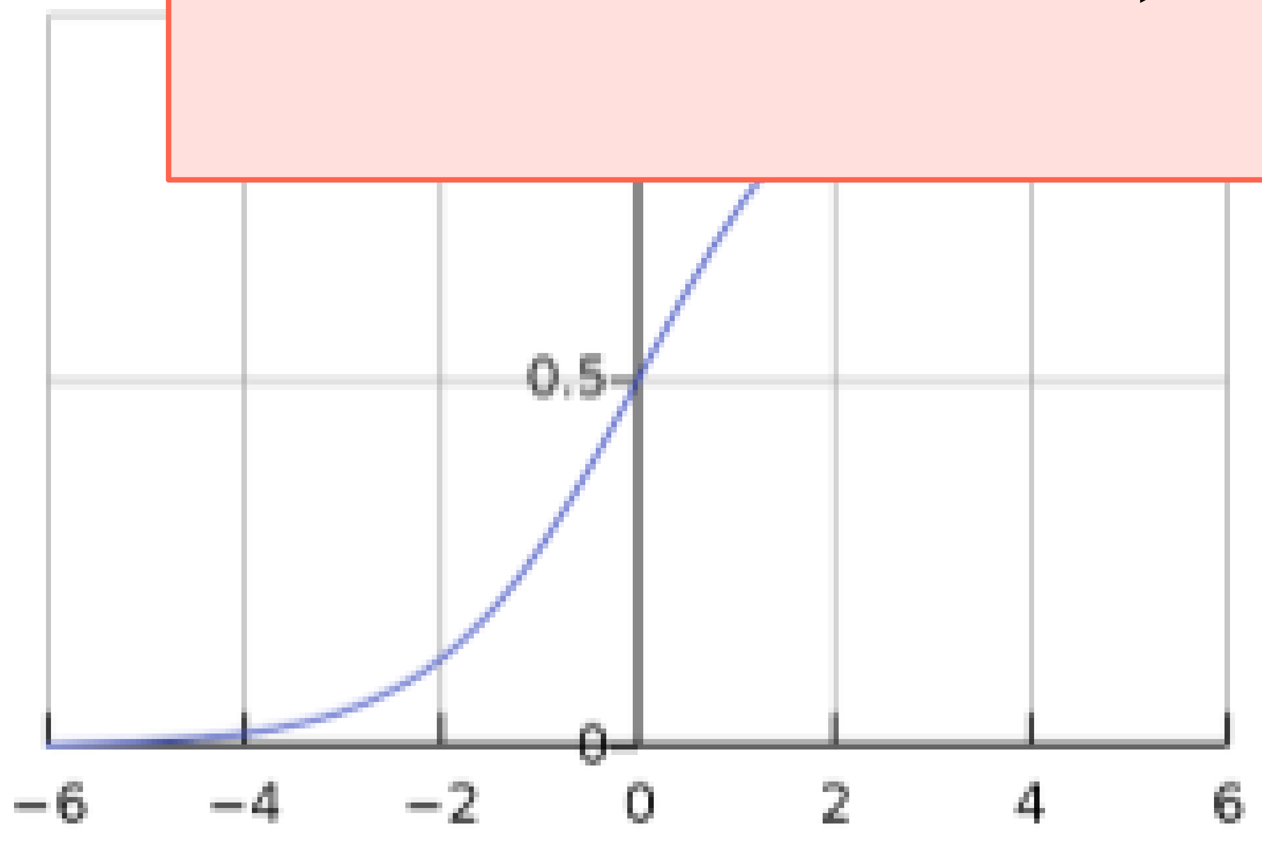


# 활성화 함수 종류

## 1. 시그모이드 sigmoid 함수

결과값 0 ~ 1 완만한 곡선 형태.

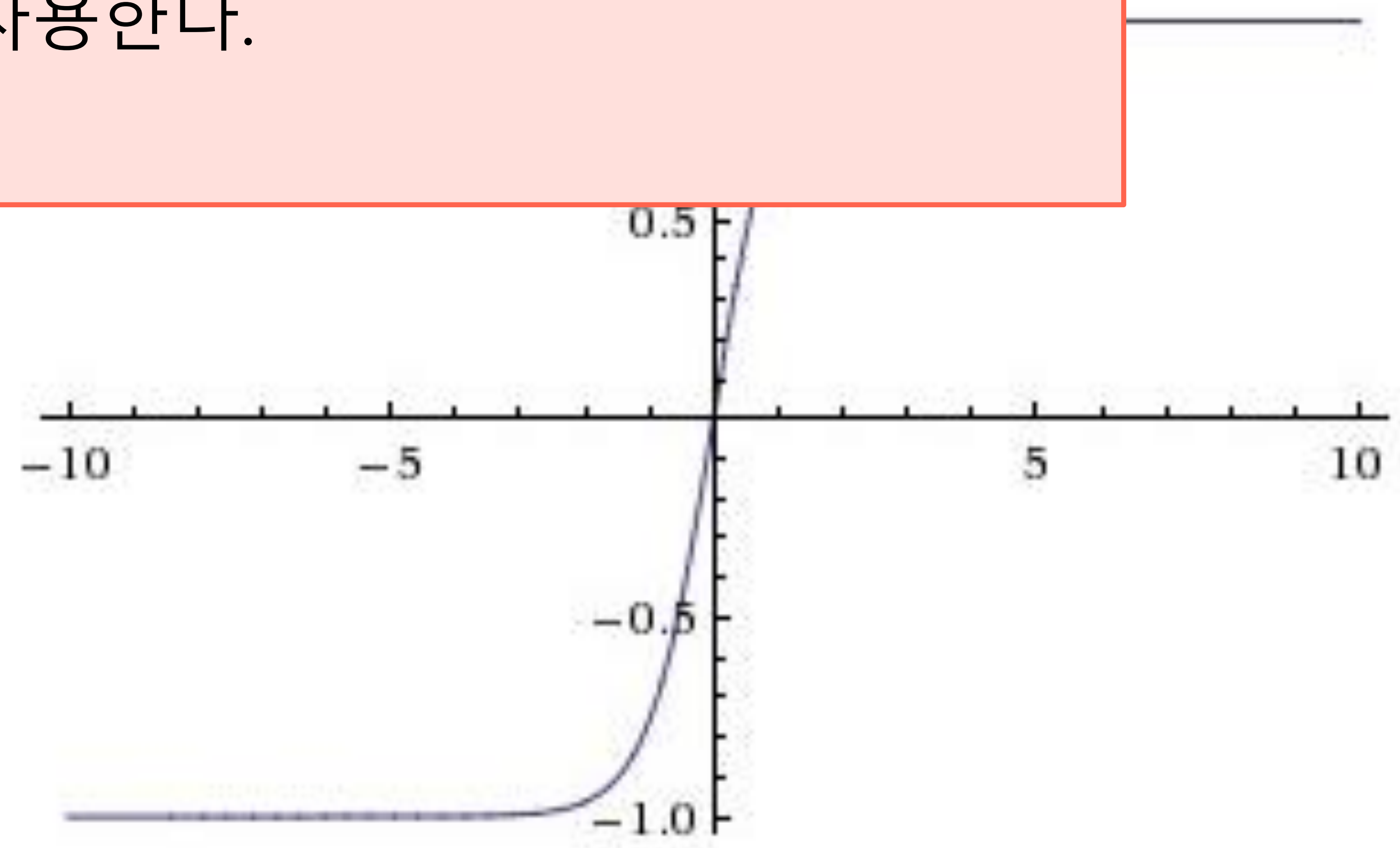
보통 이진분류(




## 2. 하이퍼볼릭 탄젠트 Tanh 함수

결과값 -1 ~ 1. 수학적으로는 시그모이드를 평행이동한 것과 같다.

때문에 다음 층의




X가 매우 크거나 매우 작을 때,

함수의 Gradient(미분값, 기울기)가 거의 0이 된다는 단점이 존재한다.

따라서 Gradient Vanishing 문제가 발생한다.

-> Backpropagation(역전파)로 인해, 활성화 함수를 미분하는데, Sigmoid와 Tanh 함수는 미분시

Gradient값이 급격히 작아지기 때문에, 여러 번 미분시 Gradient가 사라지는 문제가 발생한다.

-> 이를 해결하기 위해 RELU 함수를 사용한다.

자세한 내용 : https://wooono.tistory.com/209

## 활성화 함수 종류

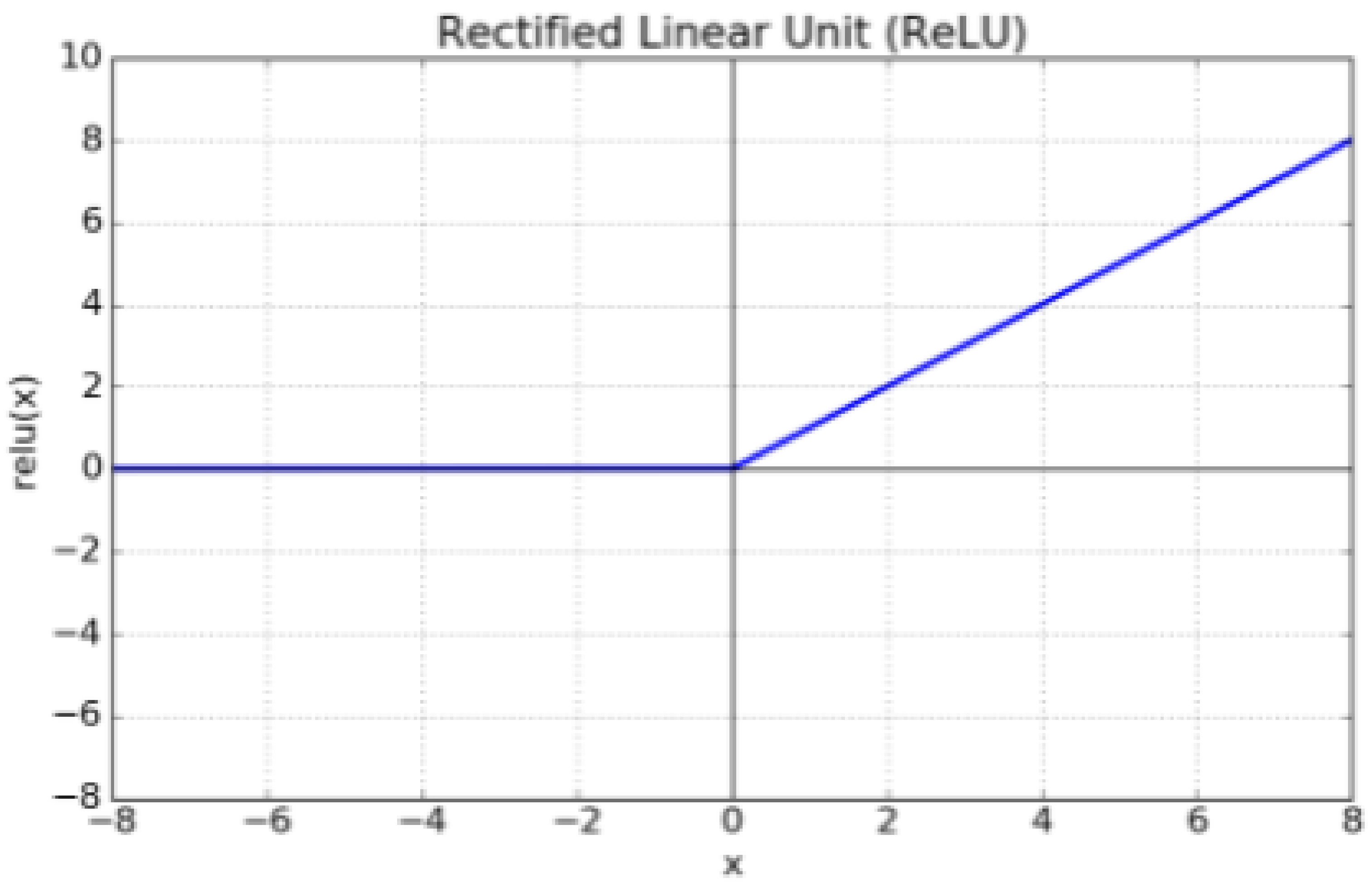
### 3. 렐루 ReLU 함수

x가 양수면 미분값이 1이고, x가 음수면 미분값이 0인 함수.

x가 음수일때 미분값이 0이라는 것이 한 가지 단점

$$R(z) = \begin{Bmatrix} z & z > 0 \\ \alpha z & z <= 0 \end{Bmatrix}$$

$$R'(z) = \begin{Bmatrix} 1 & z > 0 \\ \alpha & z < 0 \end{Bmatrix}$$

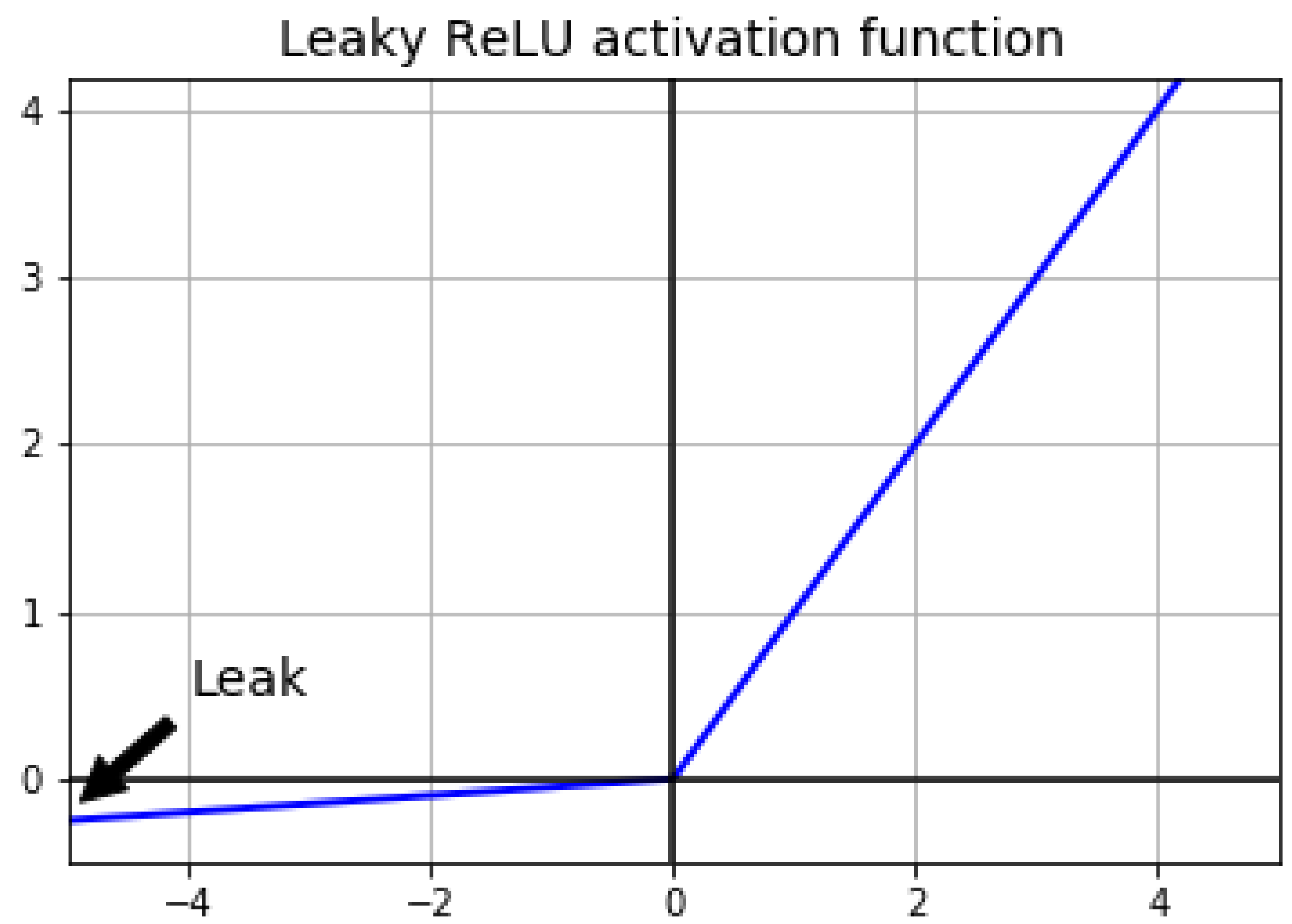



### 4. 리키 렐루 Leaky RELU 함수

RELU 단점 극복. x가 음수일 때 약간의 기울기를 갖는 함수.

잘 작동하지만 많이는 사용 X

$$f(x) = max(0.01x, x)$$

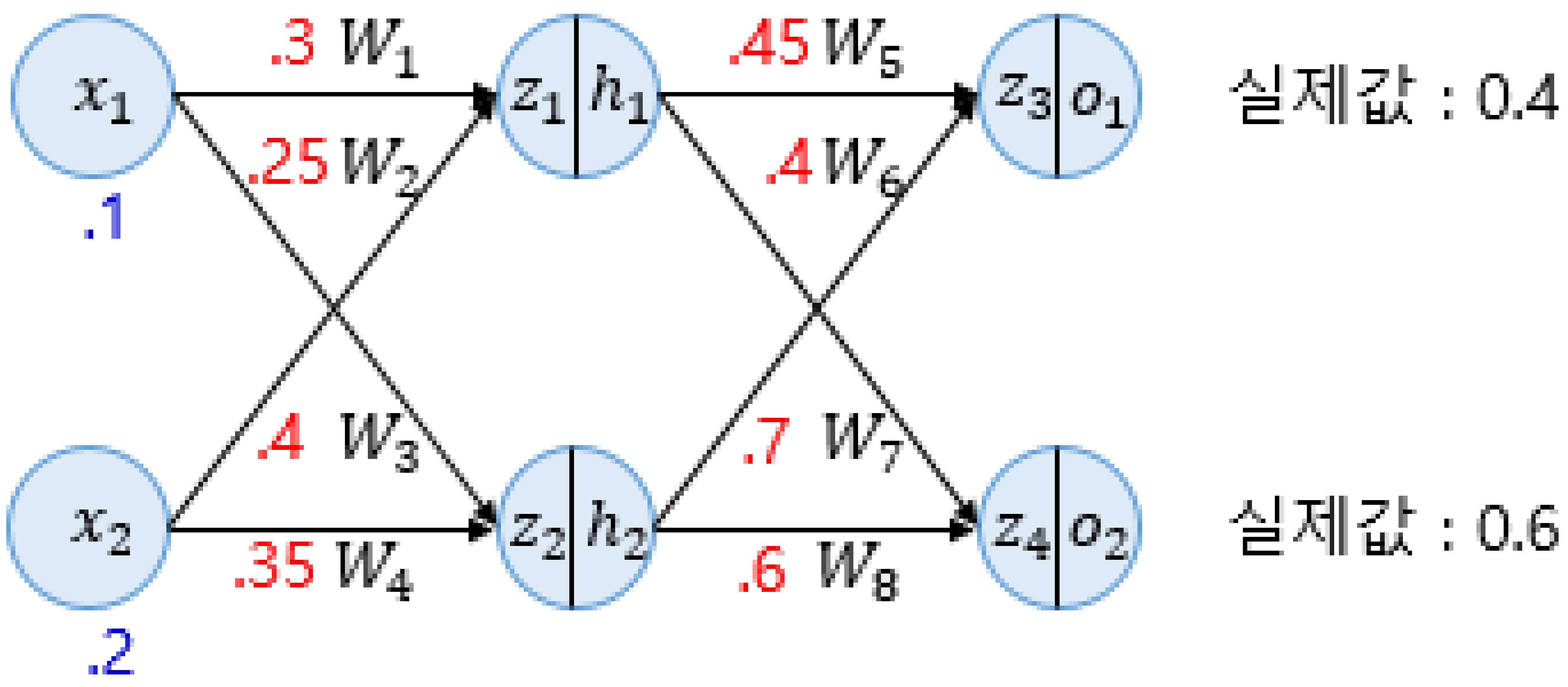


## 전파, 순전파 Froward Propagation

입력값이 들어오면 여러 개의 은닉층을 순서대로 거쳐 결과값을 내는 과정

파랑 입력값 / 빨강 가중치값 / 활성화함수는 시그모이드 함수

순전파 출력값과 오차 직접 계산해보기




$$z_1 = W_1x_1 + W_2x_2 = 0.3\times0.1 + 0.25\times0.2 = 0.08$$

$$z_2 = W_3x_1 + W_4x_2 = 0.4\times0.1 + 0.35\times0.2 = 0.11$$

$$h_1 = sigmoid(z_1) = 0.51998934$$

$$h_2 = sigmoid(z_2) = 0.52747230$$

$$z_3 = W_5h_1 + W_6h_2 = 0.45\times h_1 + 0.4\times h_2 = 0.44498412$$

$$z_4 = W_7h_1 + W_8h_2 = 0.7\times h_1 + 0.6\times h_2 = 0.68047592$$

$$o_1 = sigmoid(z_3) = 0.60944600$$

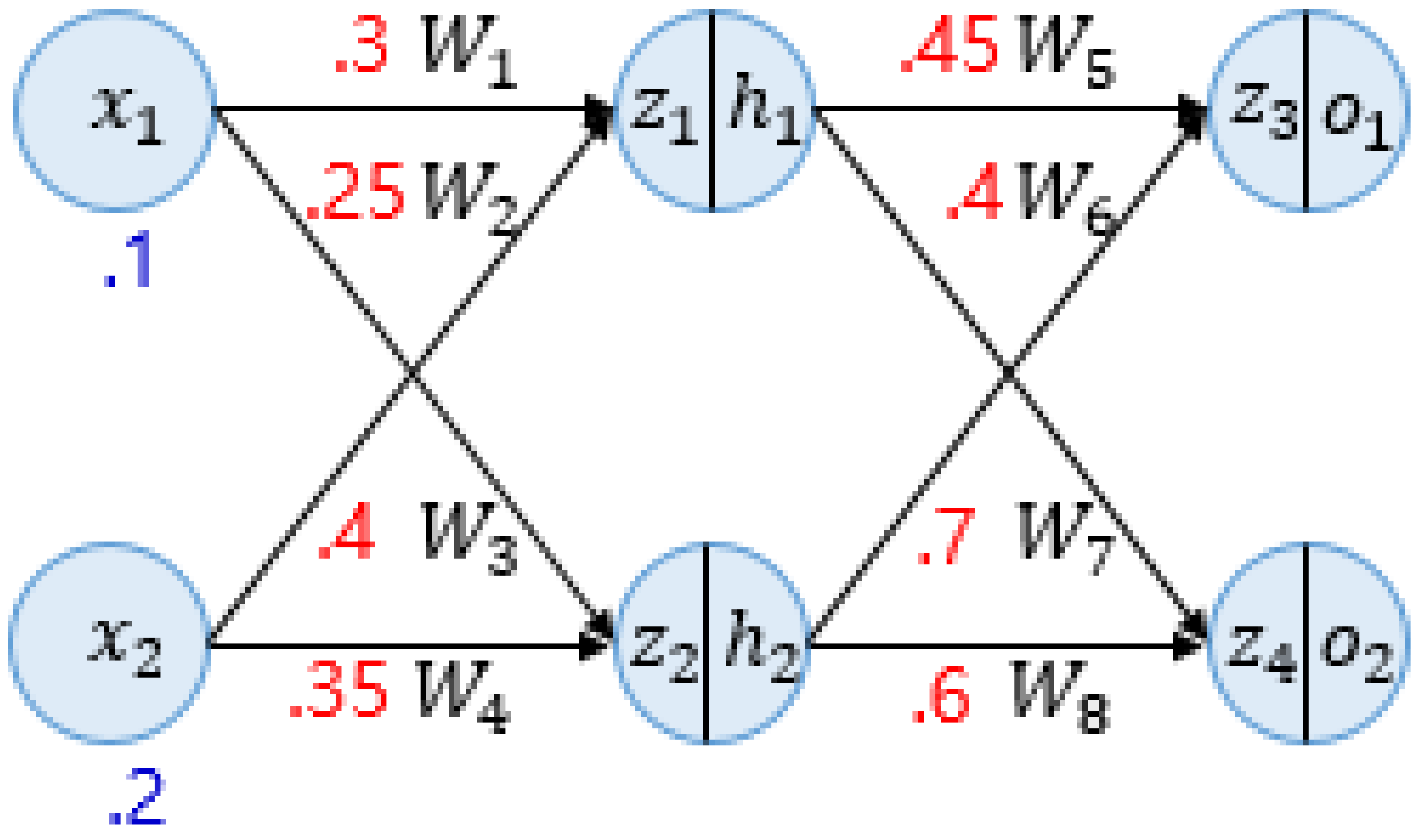
$$o_2 = sigmoid(z_4) = 0.66384491$$

## 전파, 순전파 Froward Propagation

전파. 순전파 Forward Propagation : 입력값이 들어오면 여러 개의 은닉층을 순서대로 거쳐 결과값을 내는 과정

파랑 입력값 / 빨강 가중치값 / 활성화함수는 시그모이드 함수

순전파 출력값과 오차 직접 계산해보기




실제값 : 0.4

실제값 : 0.6

$$E_{o1} = \frac{1}{2}(target_{o1} - output_{o1})^2 = 0.02193381$$

$$E_{o2} = \frac{1}{2}(target_{o2} - output_{o2})^2 = 0.00203809$$

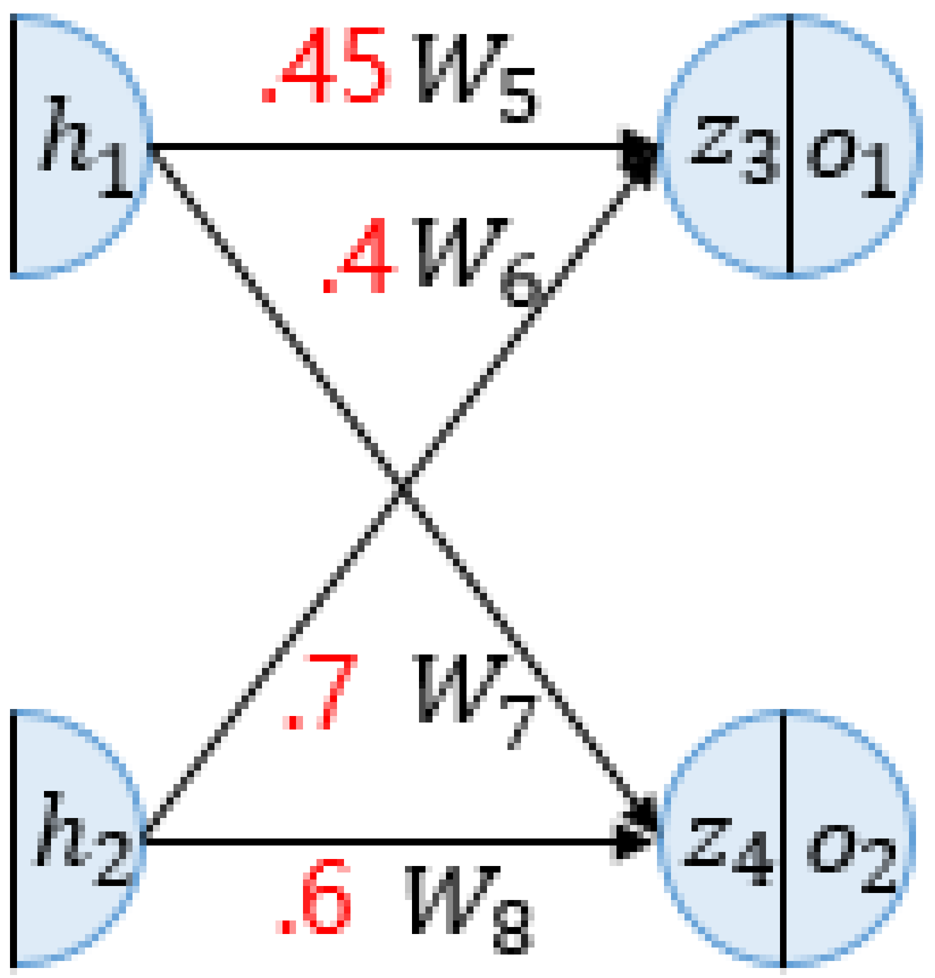
$$E_{total} = E_{o1} + E_{o2} = 0.02397190$$

https://wikidocs.net/37406

## 역전파 BackPropagation

결과와 정답의 차이로 계산된 손실을 연쇄법칙을 이용하여 입력 단까지 다시 전달하는 과정

역전파 직접 해보기 1단계. 출력층에서 바로 이전의 은닉층의 가중치를 연쇄법칙을 이용하여 업데이트하는 단계




실제값 : 0.4

실제값 : 0.6

업데이트 해야 할 가중치 W5,6,7,8 4개

$$\frac{\partial E_{total}}{\partial W_5} = \frac{\partial E_{total}}{\partial o_1} \times \frac{\partial o_1}{\partial z_3} \times \frac{\partial z_3}{\partial W_5}$$

$$E_{total} = \frac{1}{2}(target_{o1} - output_{o1})^2 + \frac{1}{2}(target_{o2} - output_{o2})^2$$

$$\frac{\partial E_{total}}{\partial o_1} = 2 \times \frac{1}{2}(target_{o1} - output_{o1})^{2-1} \times (-1) + 0$$

$$\frac{\partial E_{total}}{\partial o_1} = -(target_{o1} - output_{o1}) = -(0.4 - 0.60944600) = 0.20944600$$

## 역전파 BackPropagation

$$\frac{\partial E_{total}}{\partial W_5} = \frac{\partial E_{total}}{\partial o_1} \times \frac{\partial o_1}{\partial z_3} \times \frac{\partial z_3}{\partial W_5} \qquad E_{total} = \frac{1}{2}(target_{o1} - output_{o1})^2 + \frac{1}{2}(target_{o2} - output_{o2})^2$$

시그모이드 함수의 미분 f(x) * (1-f(x)) → $\frac{\partial o_1}{\partial z_3} = o_1 \times (1 - o_1) = 0.60944600(1 - 0.60944600) = 0.23802157$

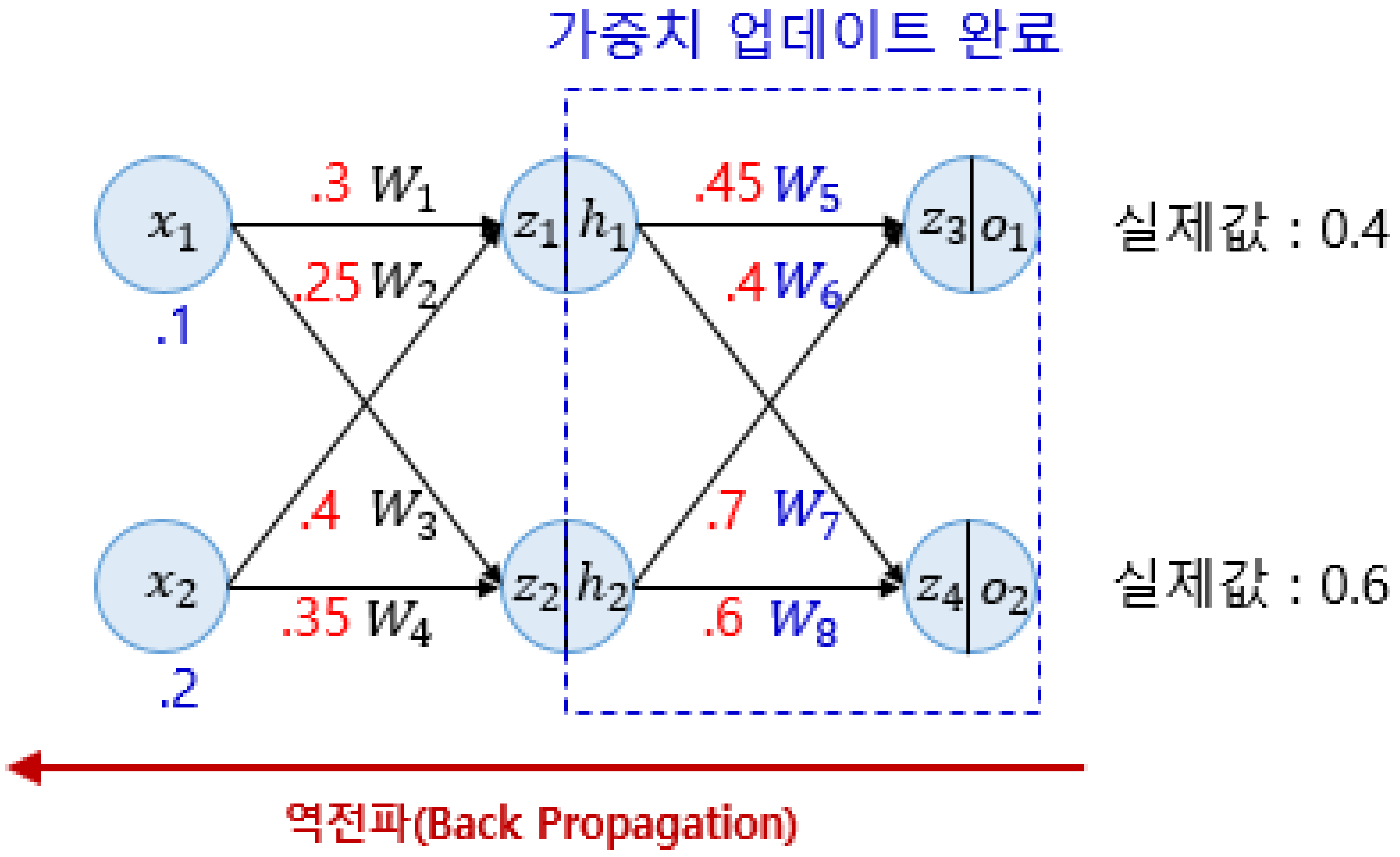
마지막 항 미분 → $z_3 = W_5 h_1 + W_6 h_2$ $\frac{\partial z_3}{\partial W_5} = h_1 = 0.51998934$

$$\frac{\partial E_{total}}{\partial W_5} = 0.20944600 \times 0.23802157 \times 0.51998934 = 0.02592286$$

학습률 0.5로 가정. 경사하강법을 통해 가중치 업데이트→ $W_5^{+} = W_5 - \alpha \frac{\partial E_{total}}{\partial W_5} = 0.45 - 0.5 \times 0.02592286 = 0.43703857$

# 역전파 BackPropagation

업데이트 해야 할 가중치 W1,2,3,4 4개




$$\frac{\partial E_{total}}{\partial W_1} = \frac{\partial E_{total}}{\partial h_1} \times \frac{\partial h_1}{\partial z_1} \times \frac{\partial z_1}{\partial W_1}$$

$$\frac{\partial E_{total}}{\partial h_1} = \frac{\partial E_{o1}}{\partial h_1} + \frac{\partial E_{o2}}{\partial h_1}$$

# 역전파 BackPropagation

Backpropagation: a simple example

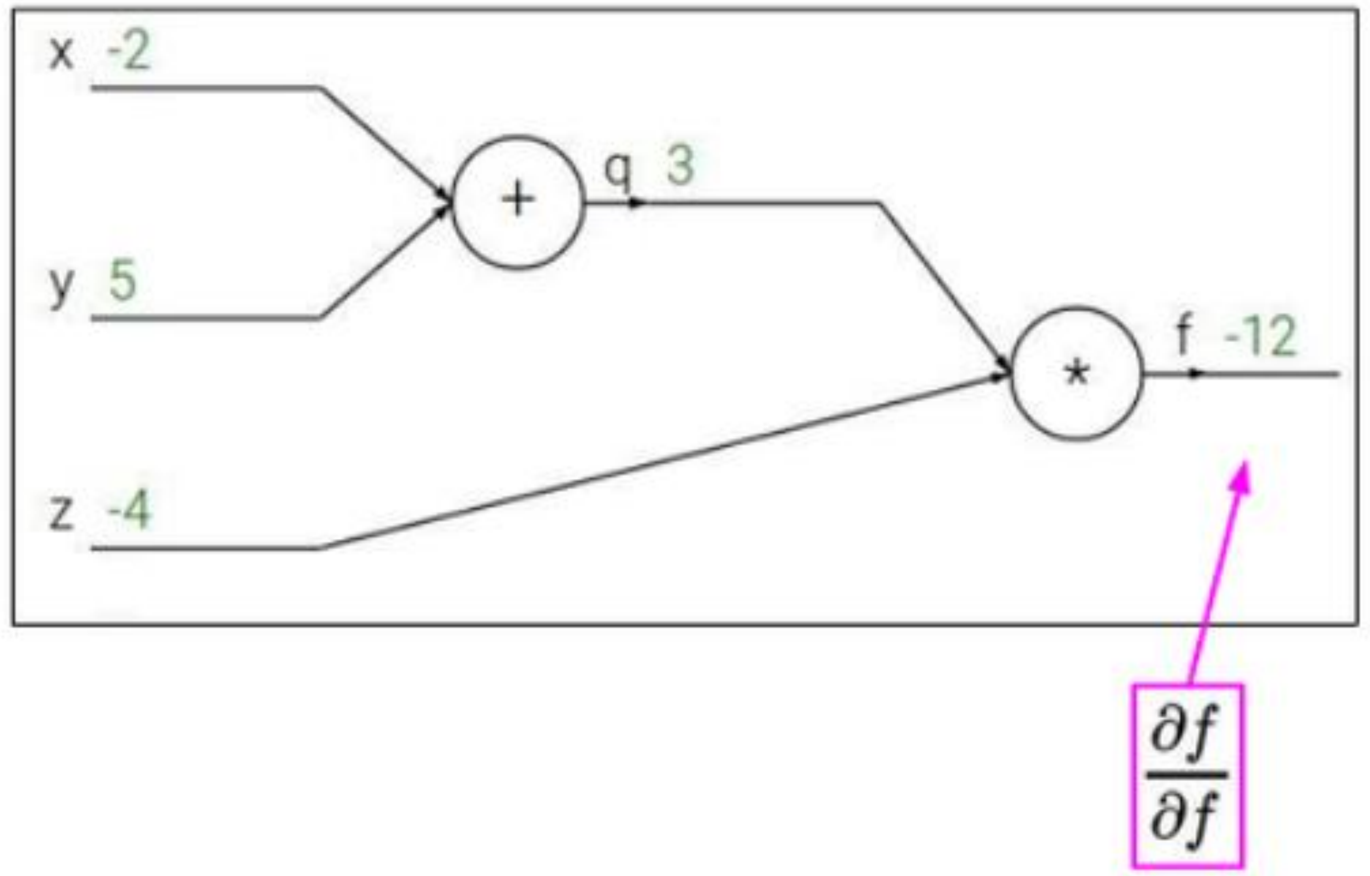
$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

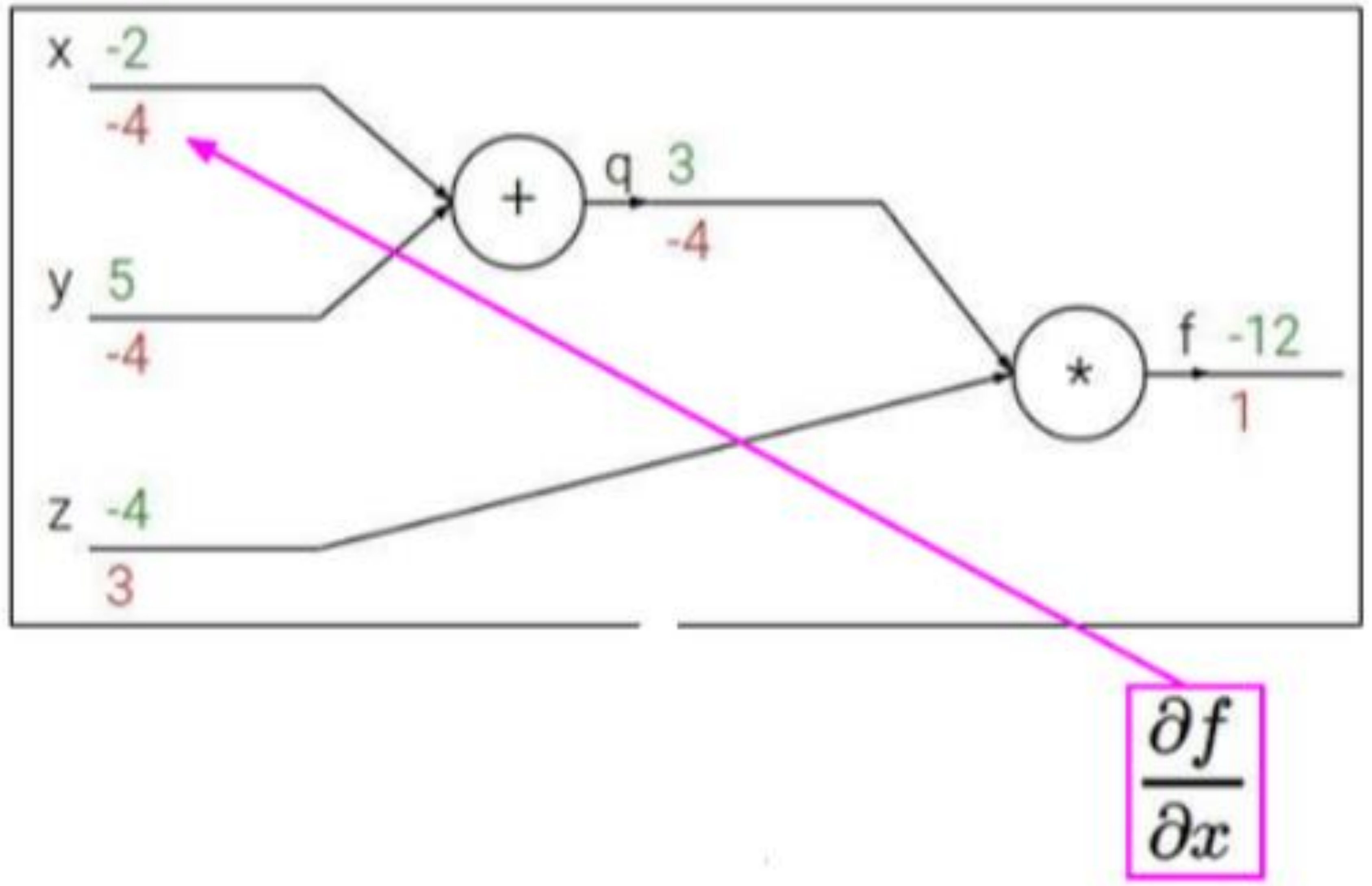
$$q = x + y \qquad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$




Fei-Fei Li & Justin Johnson & Serena Yeung Lecture 4 - 14 April 13, 2017




https://hwi-doc.tistory.com/entry/4-%EC%98%A4%EC%B0%A8%EC%97%AD%EC%A0%84%ED%8C%8Cpropagation%EB%A5%BC-%EC%9D%B4%EC%9A%A9%ED%95%9C-gradient-%EA%B3%84%EC%82%B0

## 모델 구현, 학습

자세한 주석 달아서 깃허브 실습코드 파일에 업로드 해두었습니다 !

Loss.backward() : 각 변수별 기울기 모두 계산

nn.Sequential 클래스 : 다양한 모듈을 담을 수 있는 일종의 리스트.

nn.Linear, nn.RELU 같은 모듈들을 인수로 받아서 순서대로 정렬해놓고 입력값이 들어오면 이 순서대로 모듈을

실행하여 결괏값을 리턴. 많은 연산을 묶어서 한 번에 관리할 수 있어서 편리.

# 발표 및 과제 QnA

질문 있으신 분 ~

없으면 바로 과제 풀이 !

# BOAZ

*Thank You*