



# week5

# 트레닝

2019122064 오다건

2021 / 02 / 12



순서

01 딥러닝이란?

02 인공신경망

03 딥러닝 용어  
&요소

04 RNN

05 LSTM

06 끝!

• • ○

”

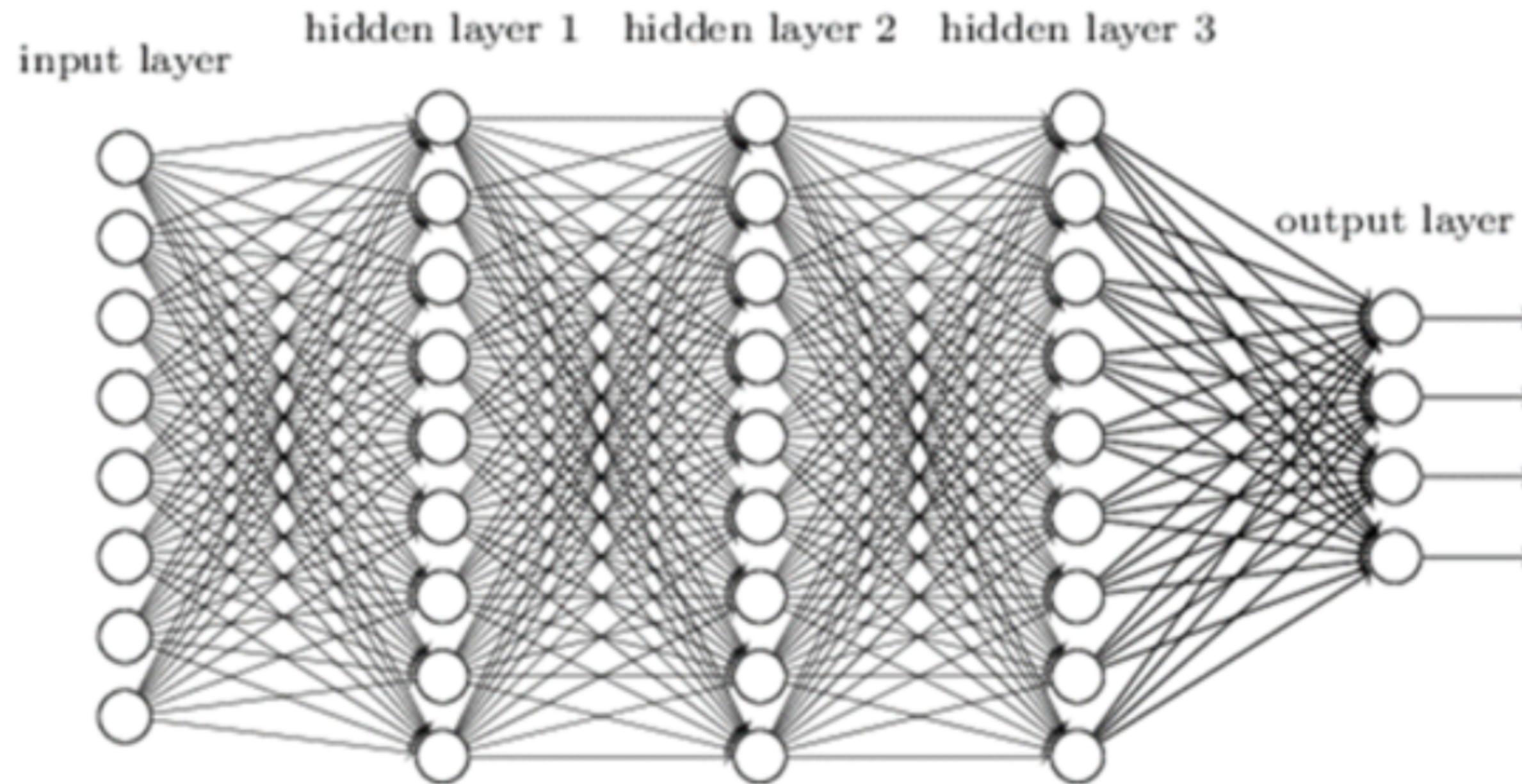
# 딥러닝이란?

딥러닝 개요/역사/ 머신러닝과의 차이

01

• • ○

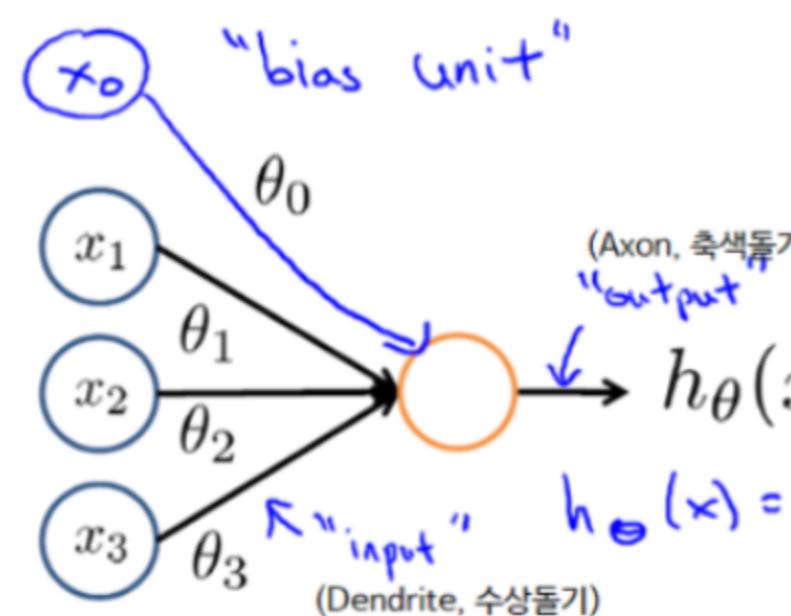
## Deep neural network



## 딥러닝:

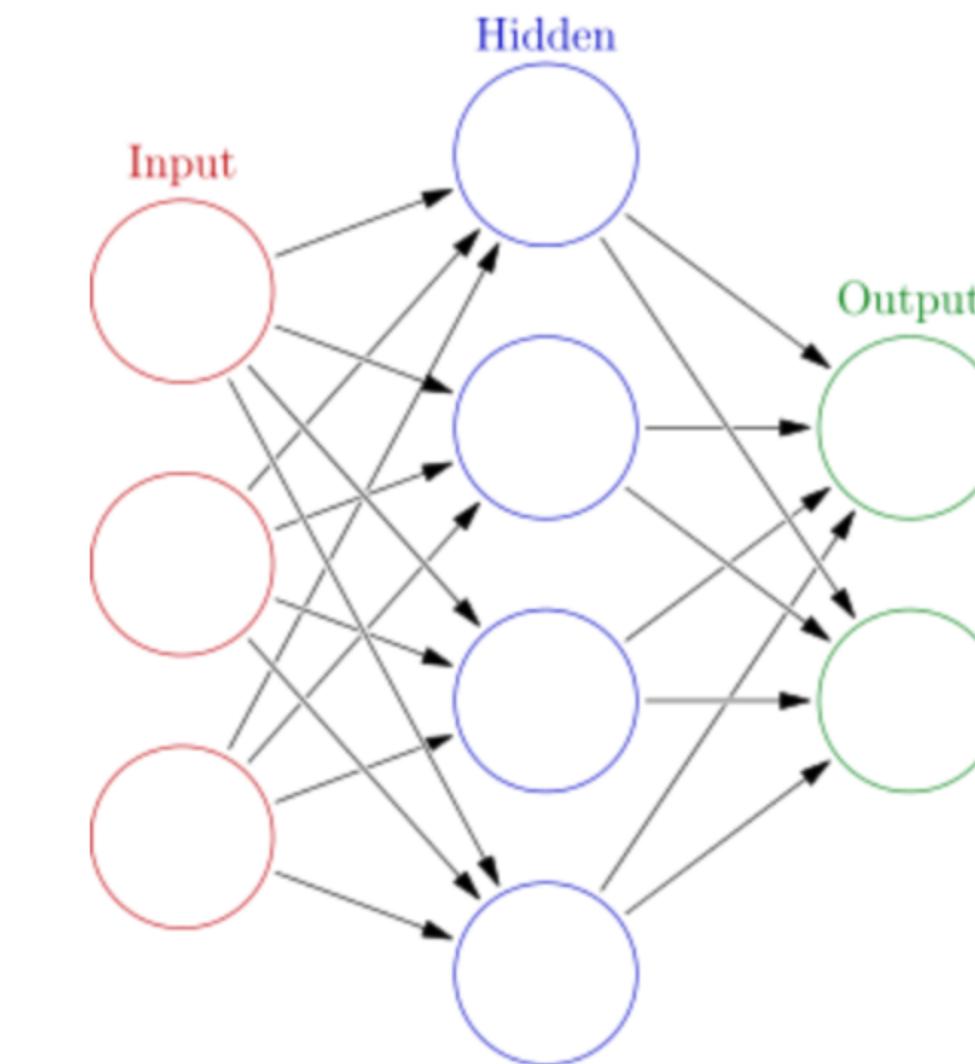
입력과 출력 사이에 있는 인공 뉴런들을 여러개 층층이 쌓고 연결한 인공신경망  
인간의 뇌 신경망을 모방.

### Neuron model:



$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}$$

$h_{\theta}(x) = \frac{1}{1 + e^{-\theta^T x}}$  (Activation Function, 활성화 함수)



인공지능 > 머신러닝 > 인공신경망 > 딥러닝

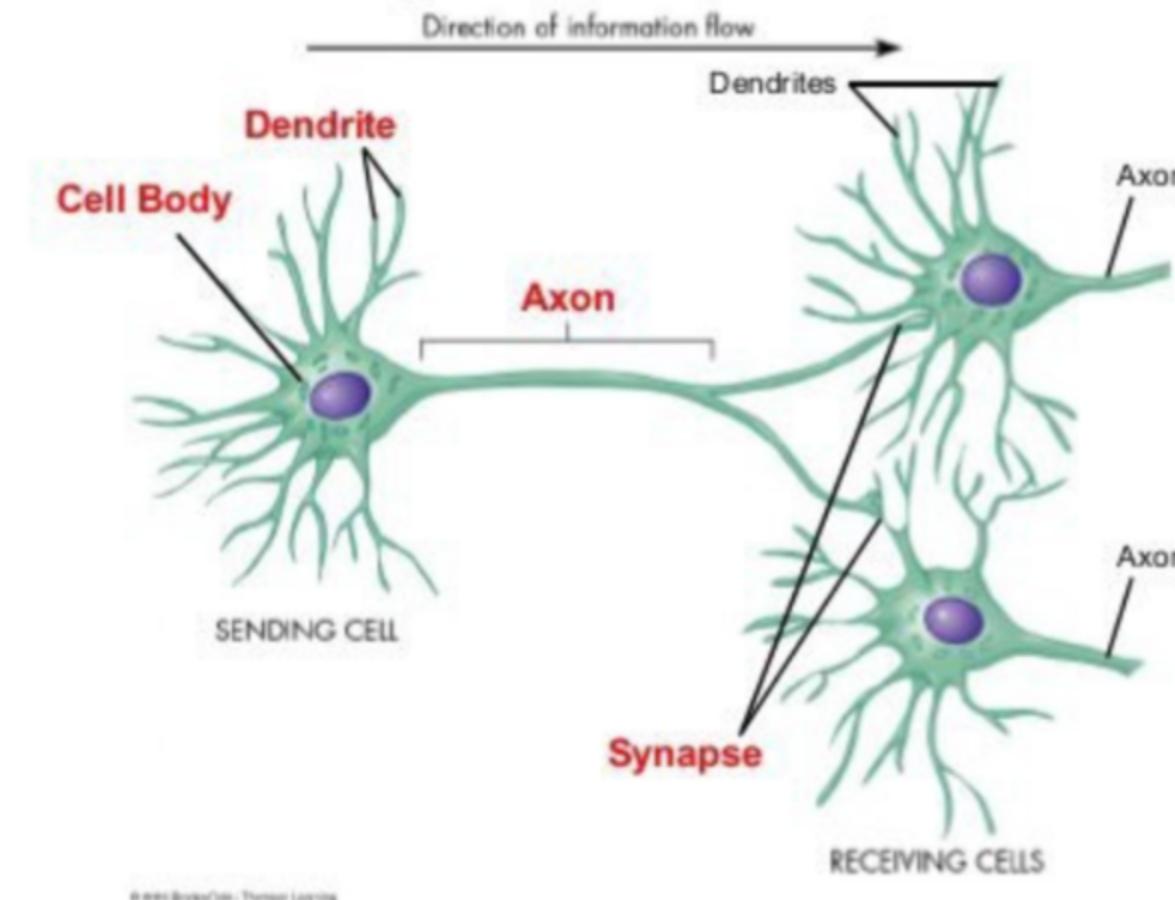


인공신경망

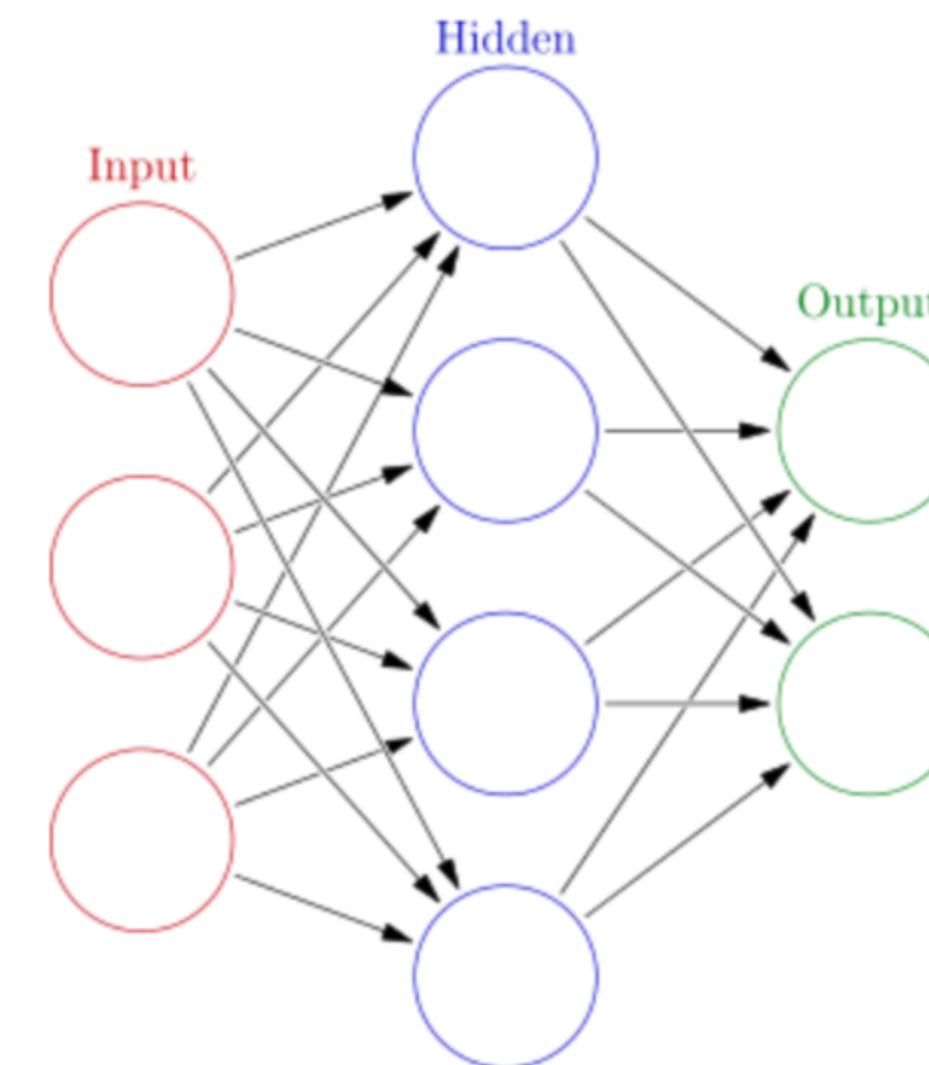
02

# 생물학적 인공신경망 & 딥러닝 인공신경망?

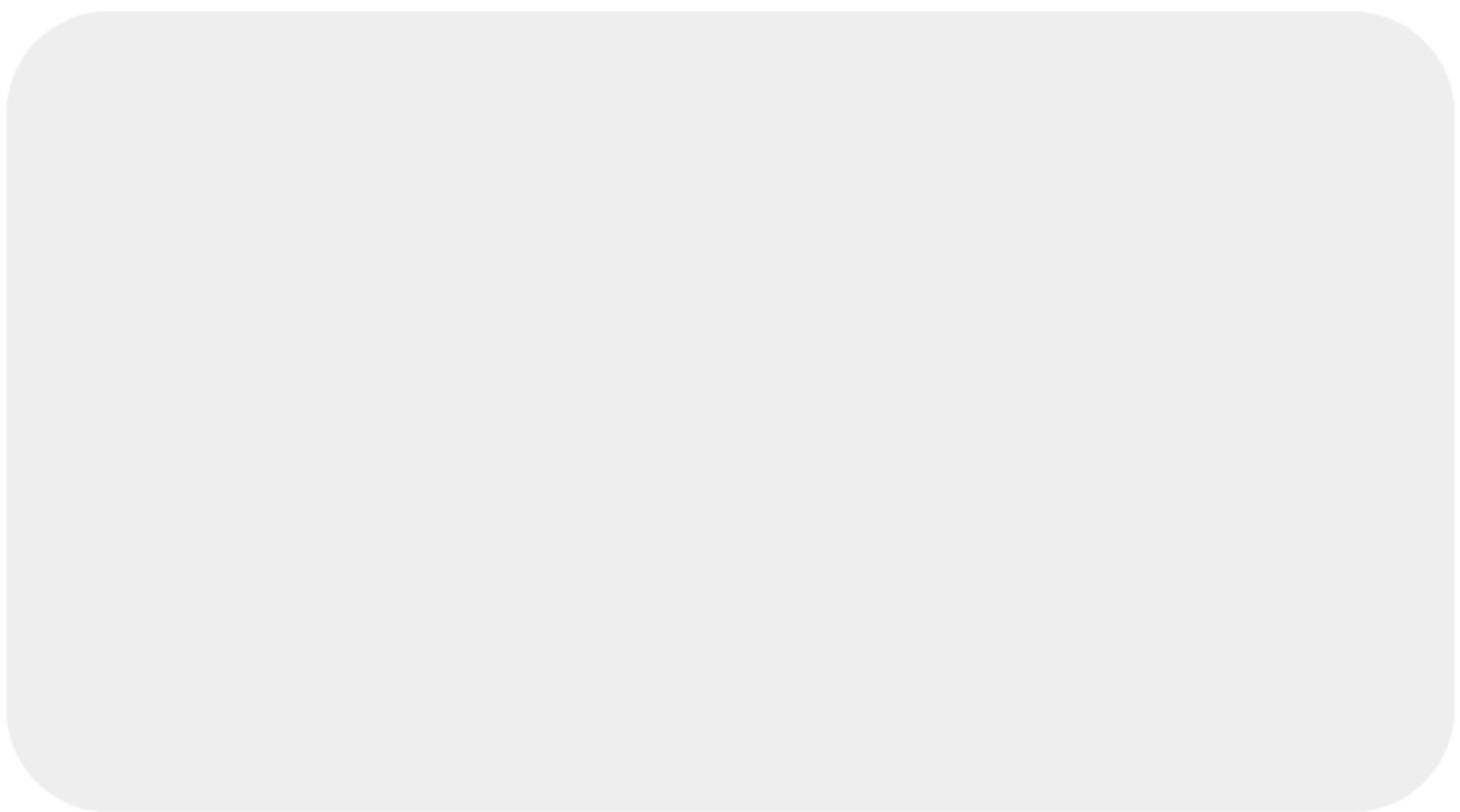
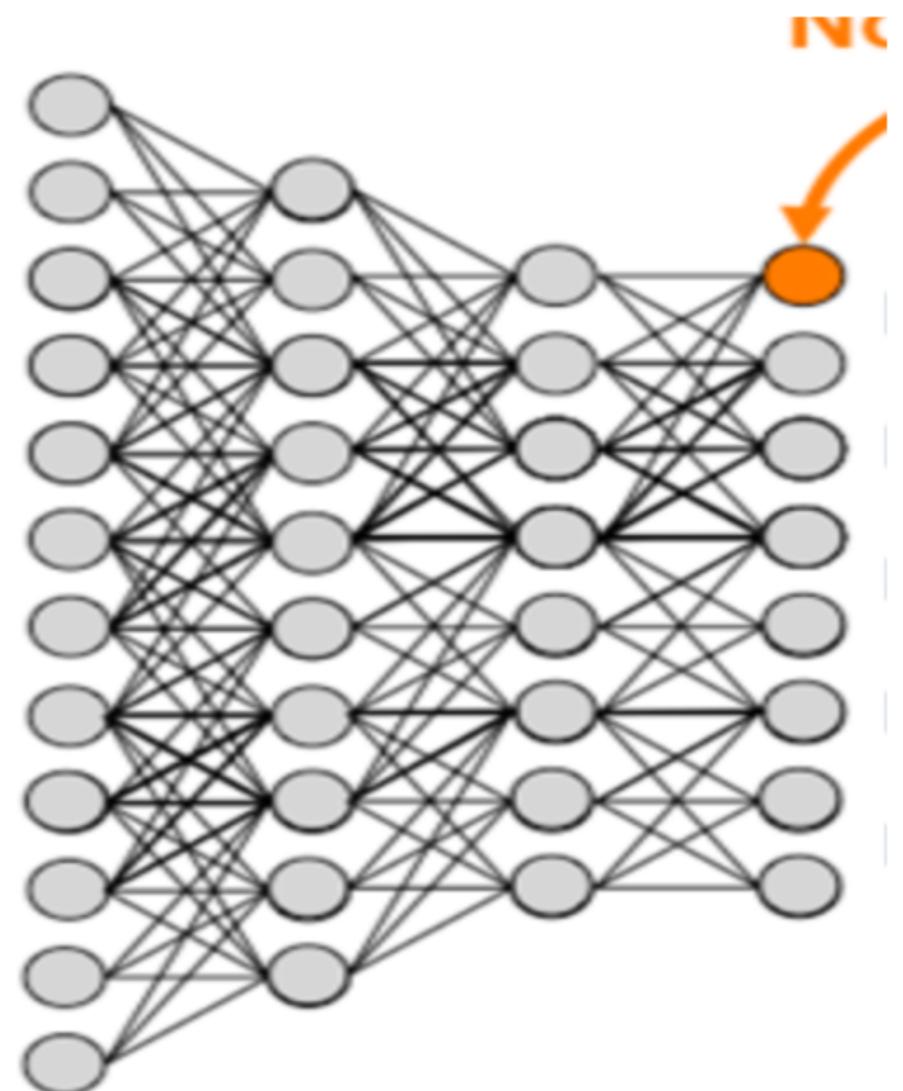
Neuron (Nerve cell) Anatomy



© 2009 Brooks/Cole - Thomson Learning



deep learning  
: 연속된 층으로 표현을 학습해서 deep

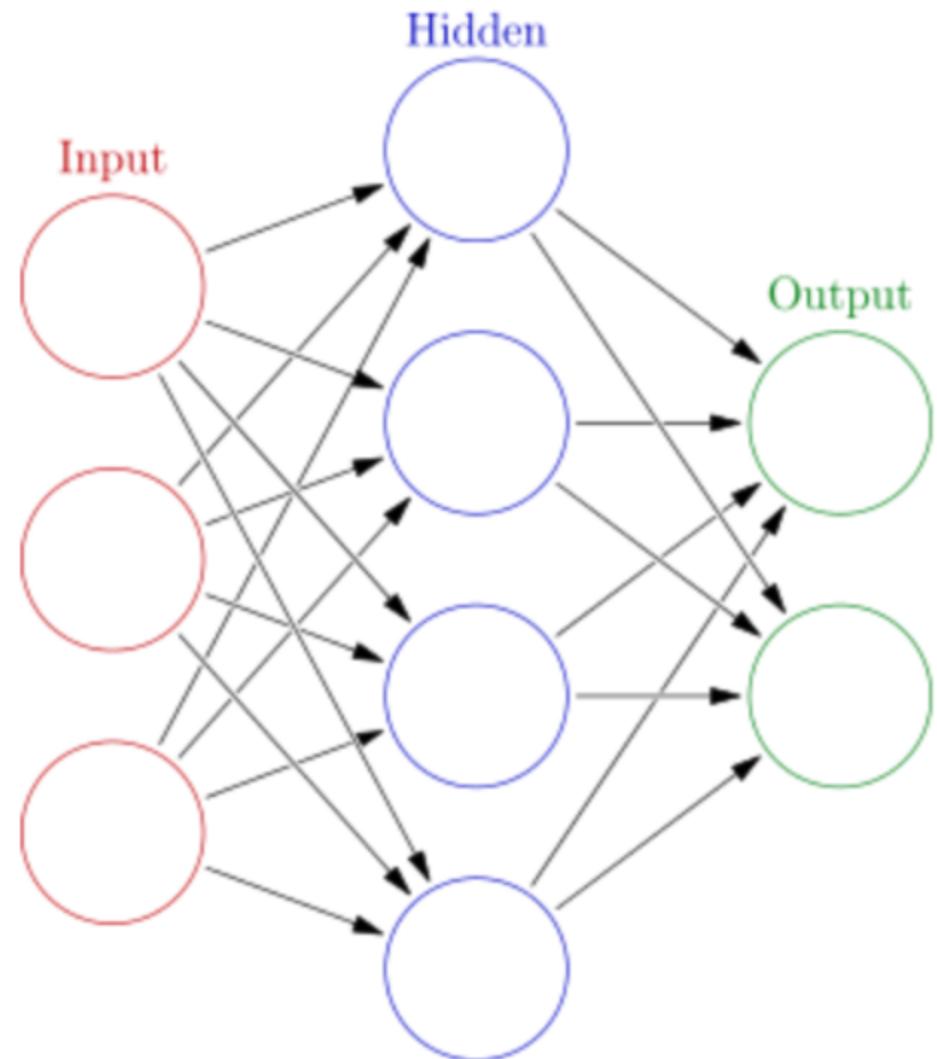




딥러닝 요소  
& 용어

03

## 딥러닝 레이어



**input layer :**

여러개의 Input nodes 존재  
독립변수의 값 입력받음

**hidden layer :**

종속변수를 예측하는데 필요한 정보 추출

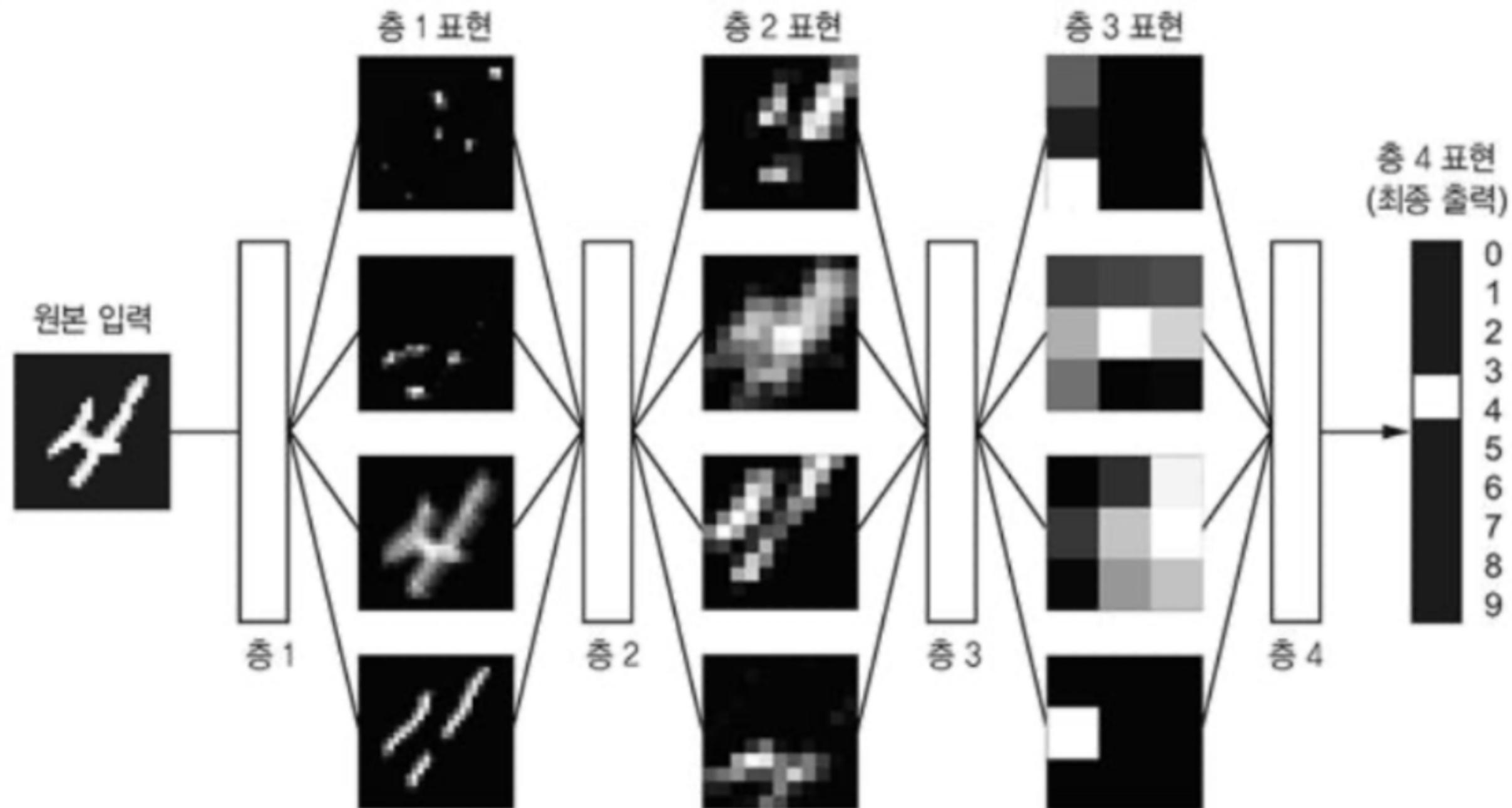
**output layer:**

예측치 출력

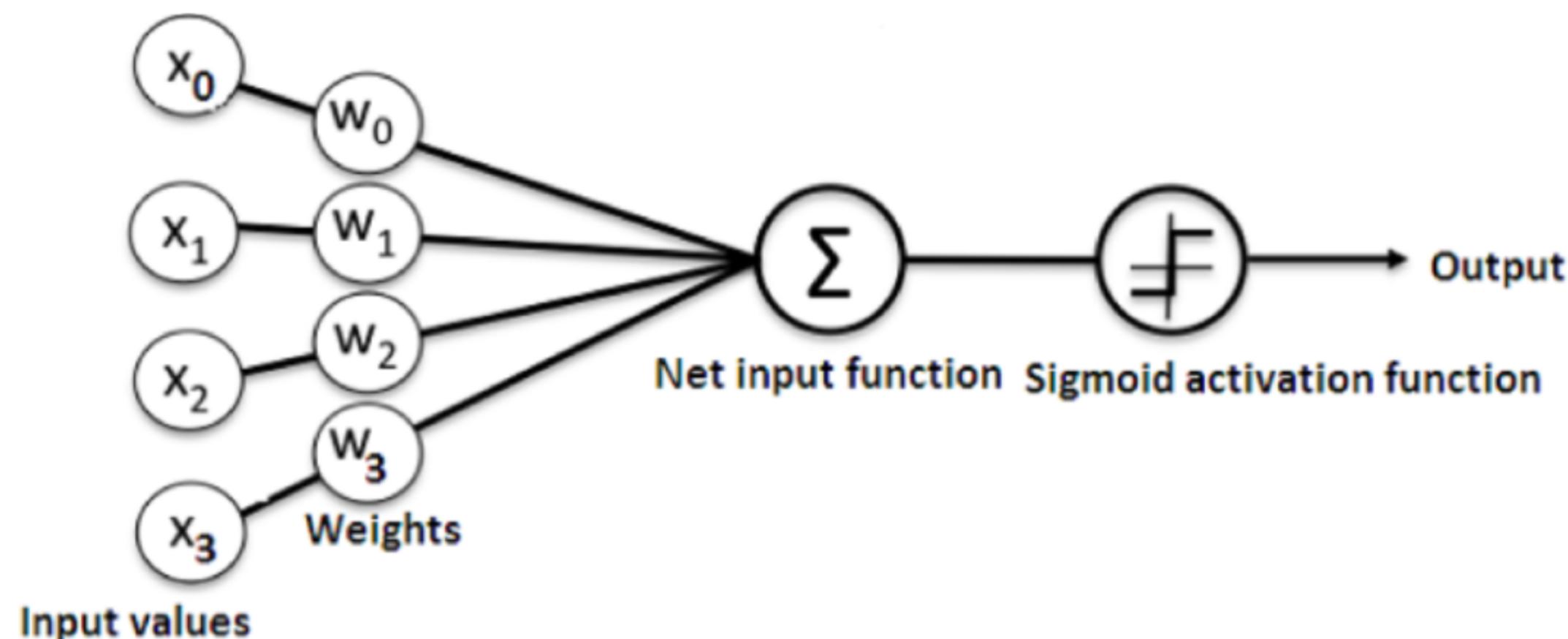
분류문제면 2개의 Output nodes,  
회귀문제이면 1개의 output node

● ● ○

## 딥러닝 학습 원리



## 딥러닝 학습 원리



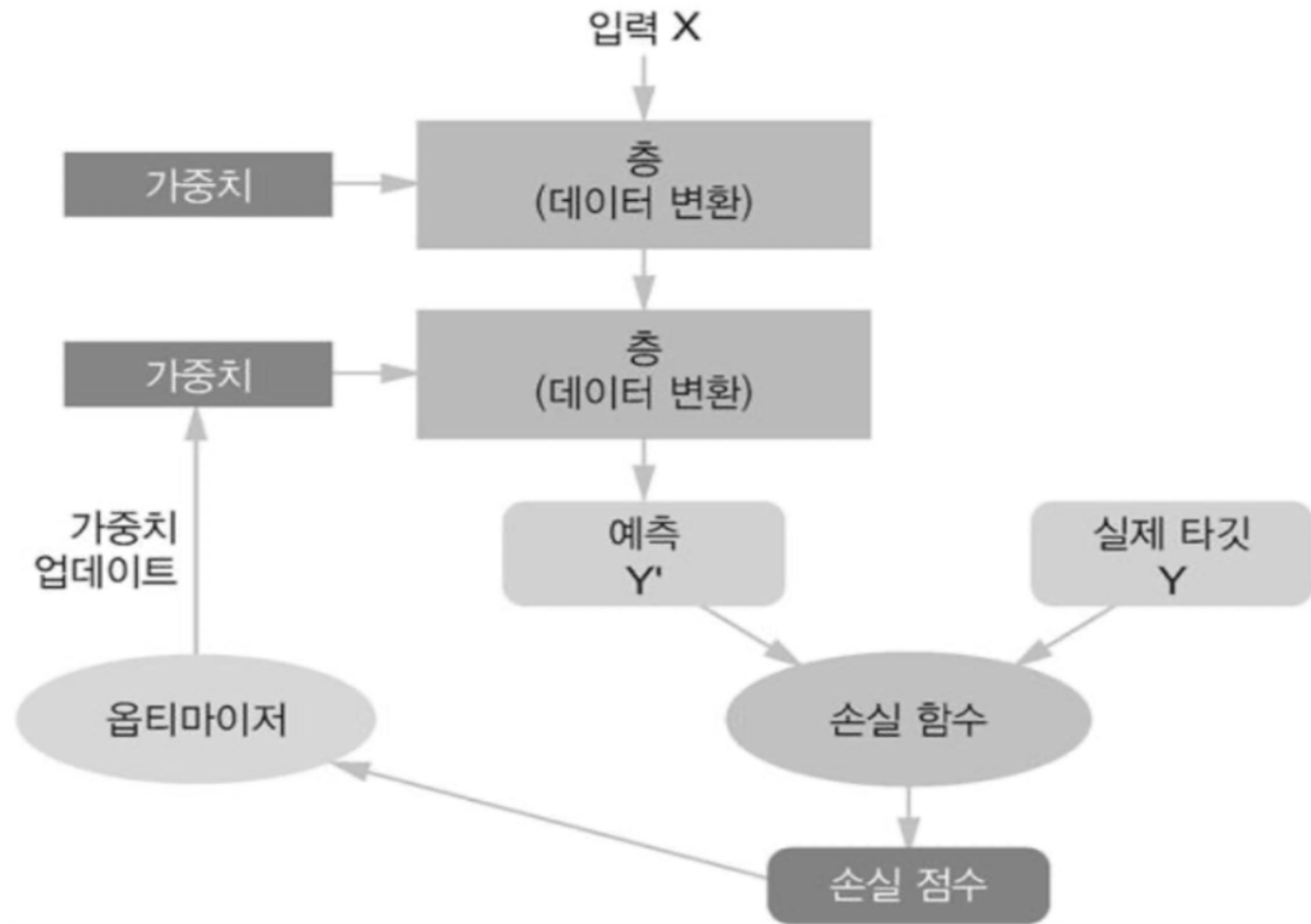
• • ○

## 딥러닝 학습 원리

- 1) 모형의 구조 결정
- 2) 관측치가 갖는 독립변수값을 가지고 예측치를 도출 (순전파)
- 3) 각 관측치가 갖는 실제 종속변수의 값과 모형을 통해서 나온 예측치와의 차이를 이용해 구한 비용함수 계산
- 4) 이 비용함수를 미분을 하여 기울기를 구하고
- 5) 비용함수의 값을 최소화하는 모형의 파라미터, 즉 가중치의 값을 계산하여 업데이트. (역전파 수행)

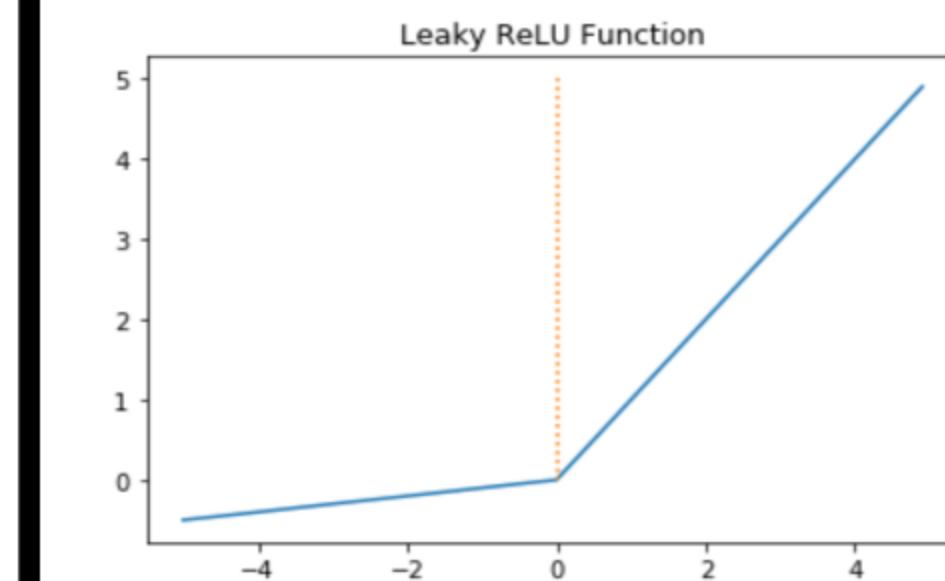
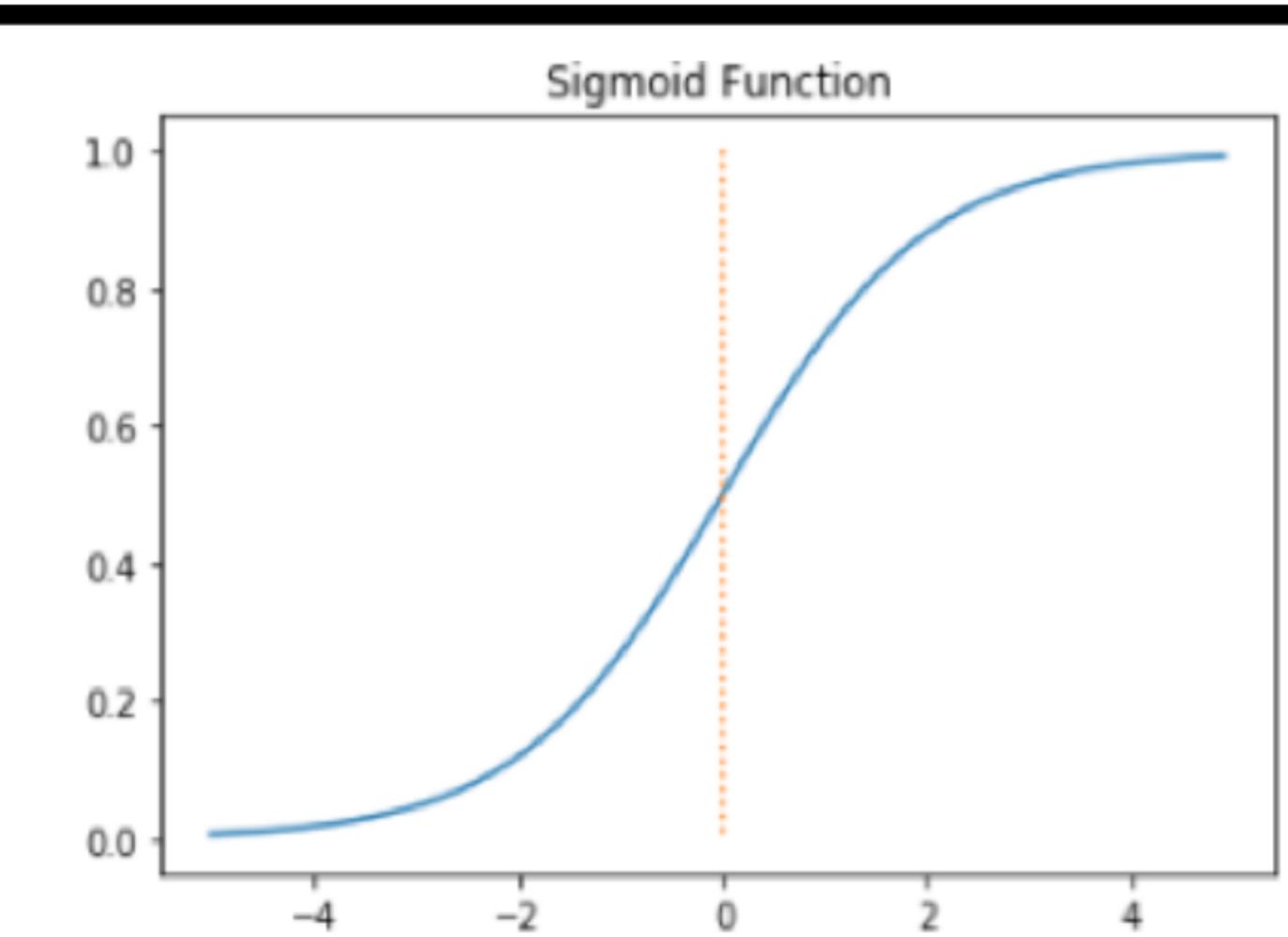
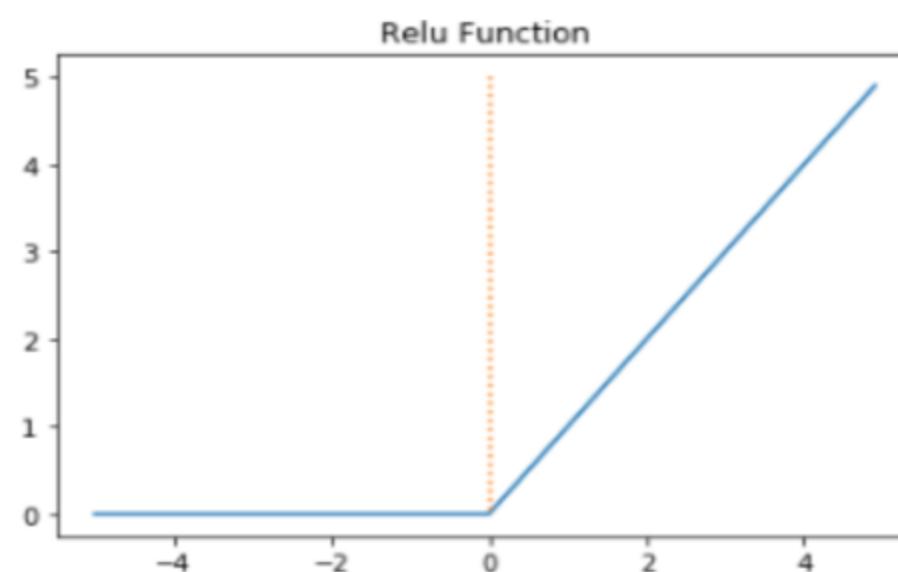
● ● ○

## 딥러닝 학습 원리



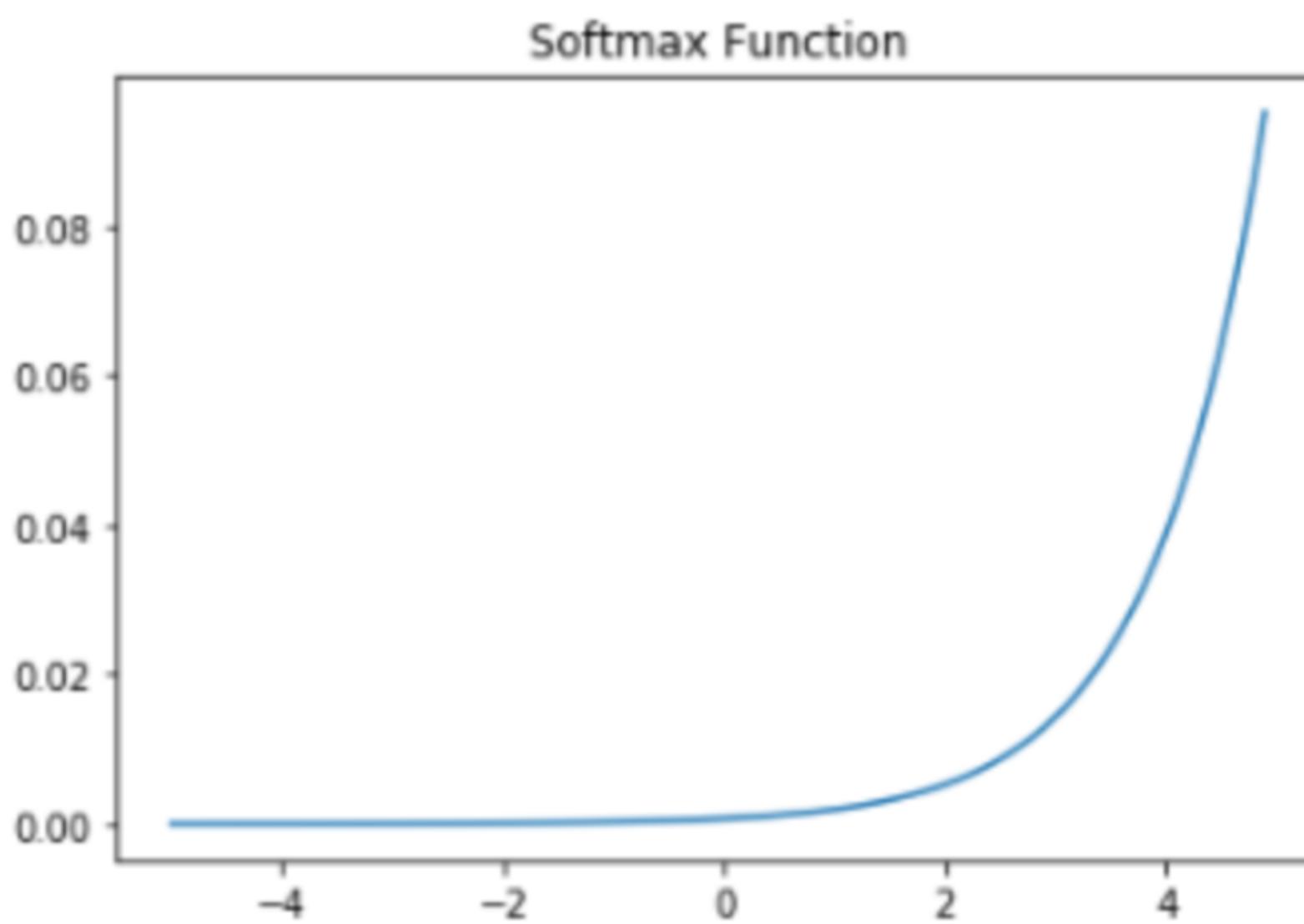
# vanishing gradient

역전파 과정에서 0에 가까운 아주 작은 기울기가 곱해지게 되면, 앞단에는 기울기가 잘 전달되지 않게 됨.  
그러면 sigmoid function을 대신하여 RELU, Leaky RELU, softmax function을 사용하고  
gradient explode가 일어난다면 gradient clipping, weight intialization, Batch normalization으로 대체



• • ○

# softmax function





## loss function

- 3) 각 관측치가 갖는 실제 종속변수의 값과 모형을 통해서 나온 예측치와의 차이를 이용해 구한 비용함수 계산
- 4) 이 비용함수를 미분을 하여 기울기를 구하고  
---- 의 손실함수!

회귀 문제: MSE  
분류 문제 : Cross Entropy

• • ○

# optimizer

손실 함수의 값을 줄여나가면서 학습하는 방법은 어떤 옵티마이저를 사용하느냐에 따라 달라짐.

01

BGD

배치 경사 하강법은 전체 데이터를 고려하고  
한 번의 에포크에  
모든 매개변수 업데이트를 단 한 번 수행

## Batch

가중치 등의 매개 변수의 값을  
조정하기 위해 사용하는 데이터의 양

02

SGD

확률적 경사 하강법은 매개변수 값을 조정 시  
전체 데이터가 아니라 랜덤으로 선택한  
하나의 데이터에 대해서만 계산하는 방법

03

Mini Batch

정해진 양에 대해서만 계산하여 매개 변수의  
값을 조정하는 경사 하강법을 미니 배치 경사  
하강법이라고 함. 가장 많이 사용



# Epoch, Batch Size, iteration

01

## Epoch

전체 데이터에 대해 순전파 + 역전파가 끝난 상태

02

## Batch Size

몇 개의 데이터 단위로 매개변수를 업데이트 하는지를 말함.

03

## Iteration

한 번의 에포크를 끝내기 위해서 필요한 배치의 수



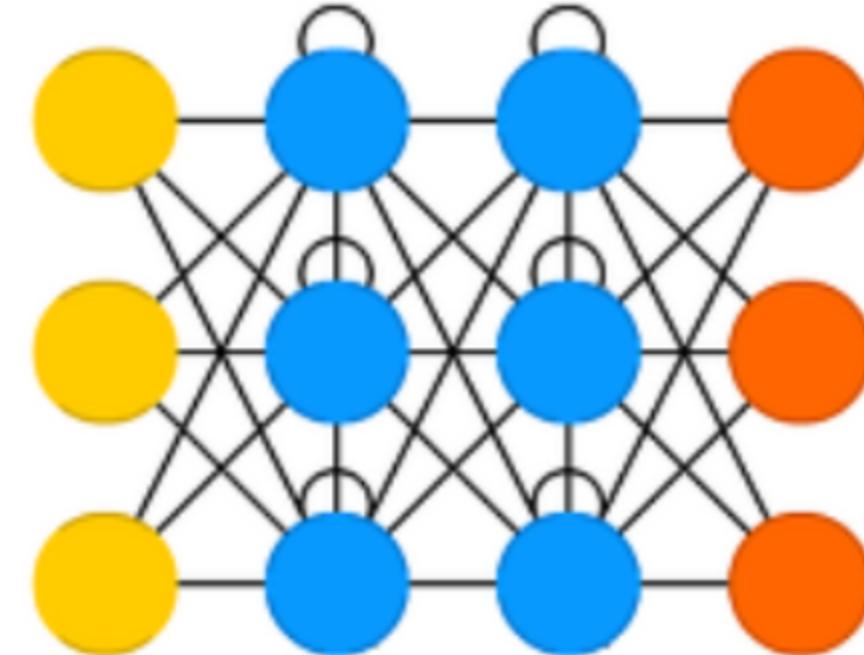
순환신경망

04

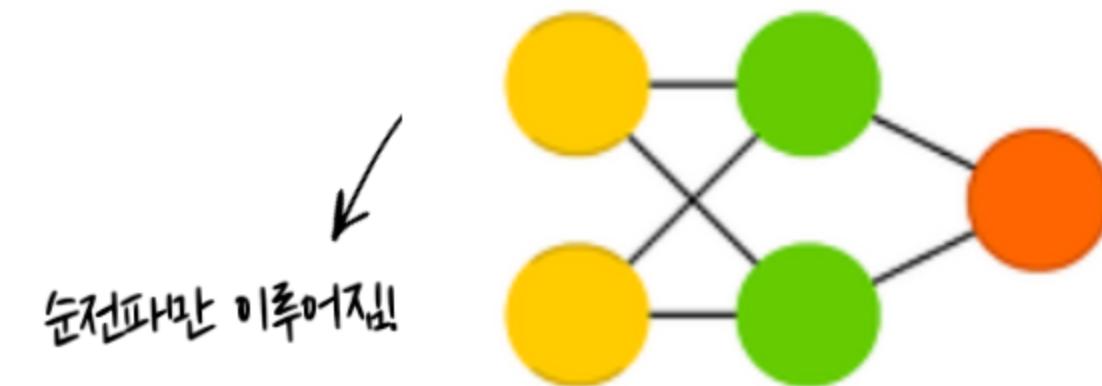
# RNN

입력과 출력을  
시퀀스 단위로 처리하는 모델

Recurrent Neural Network (RNN)

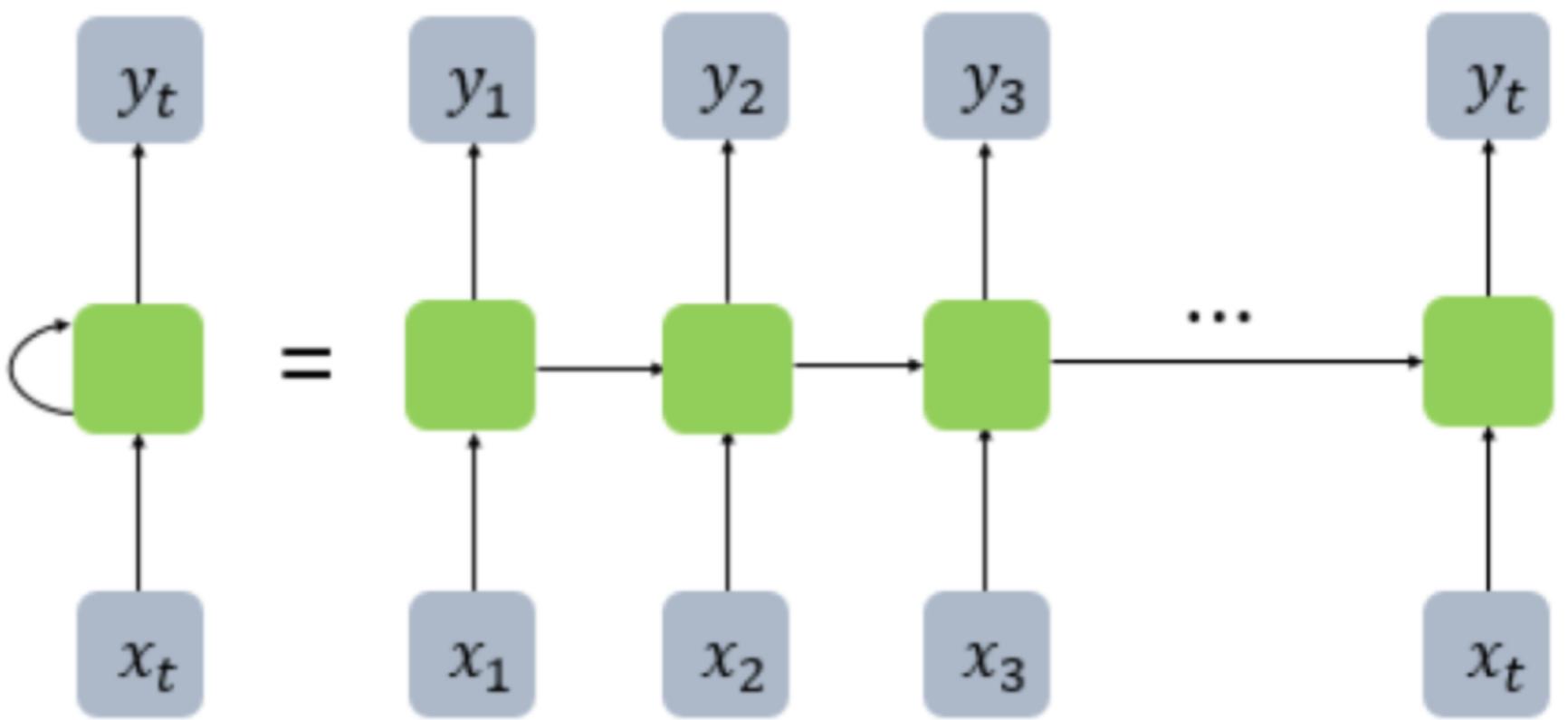
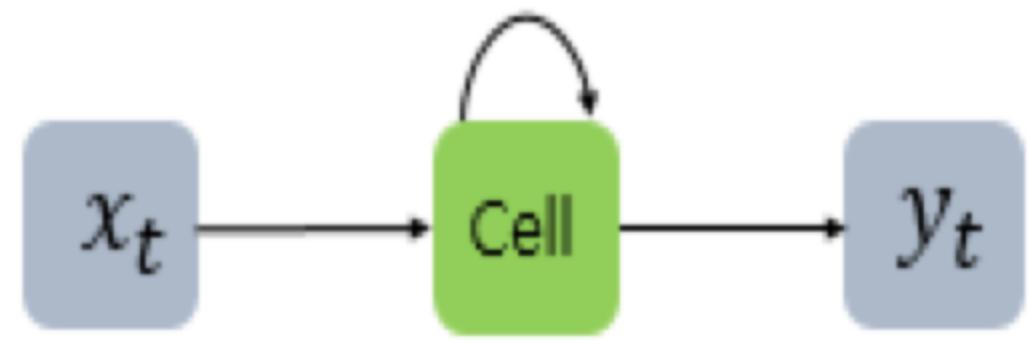


Feed Forward (FF)

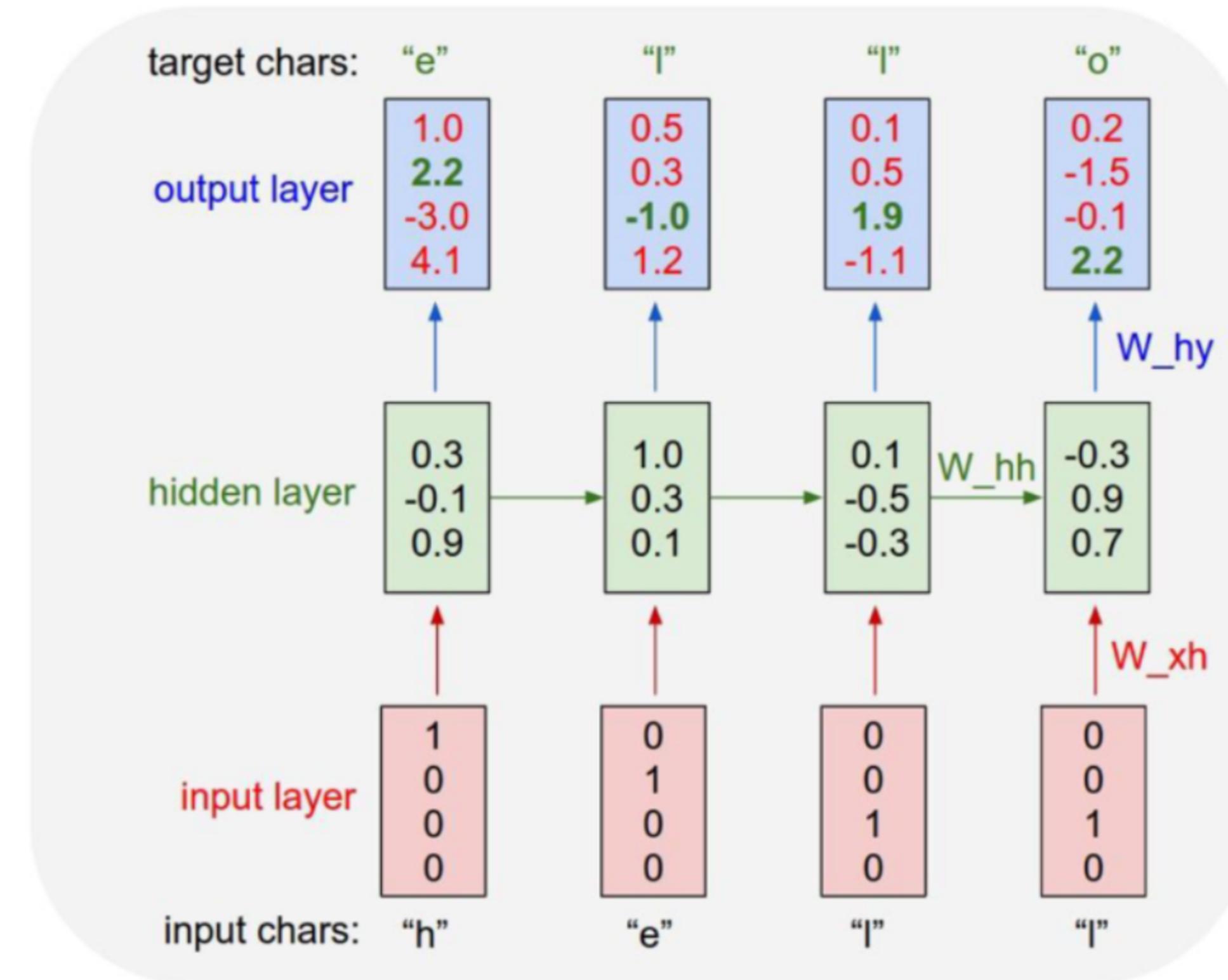


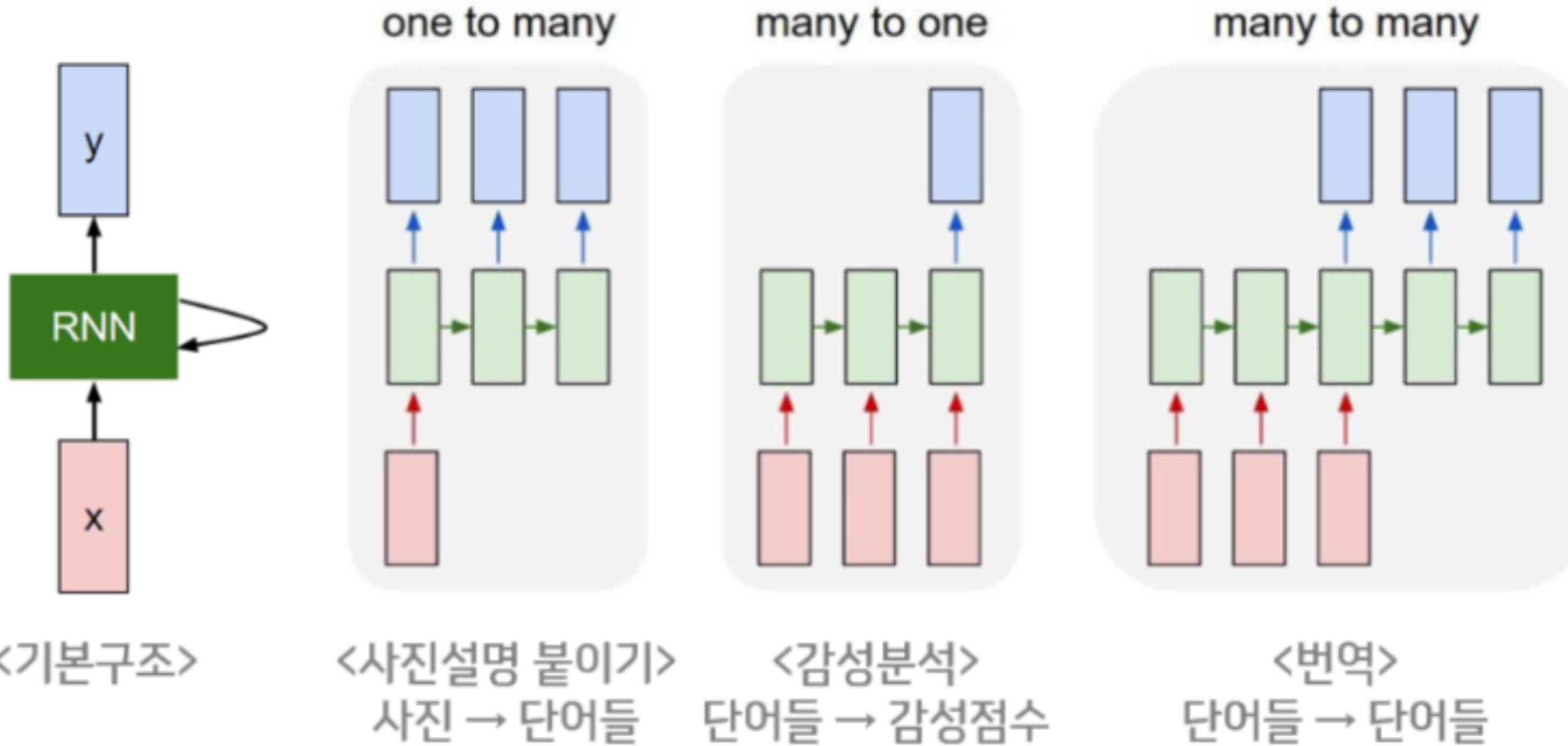
순전파만 이루어짐!

02



• • ○







LSTM

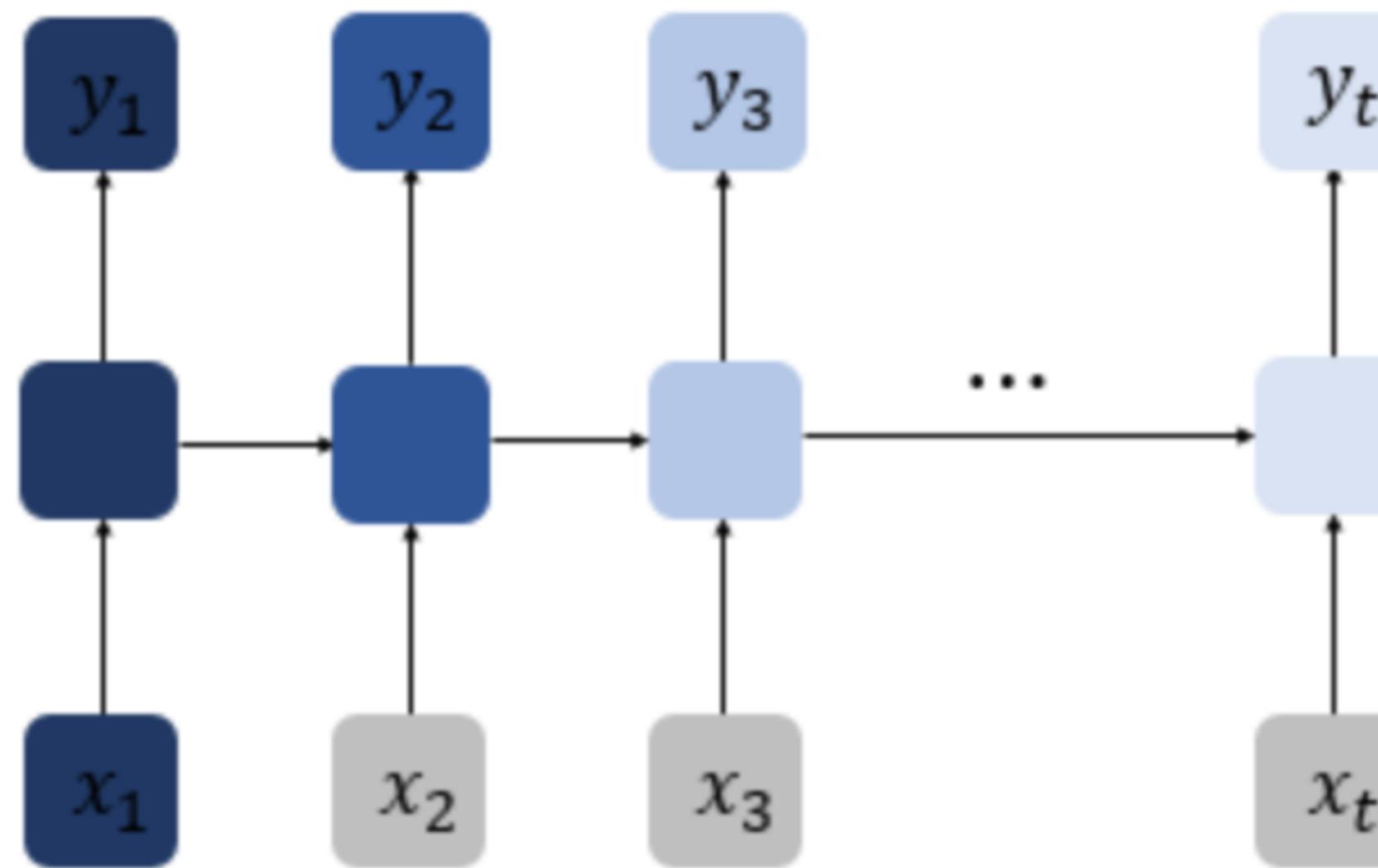
Long Term Short Memory

05

• • ○

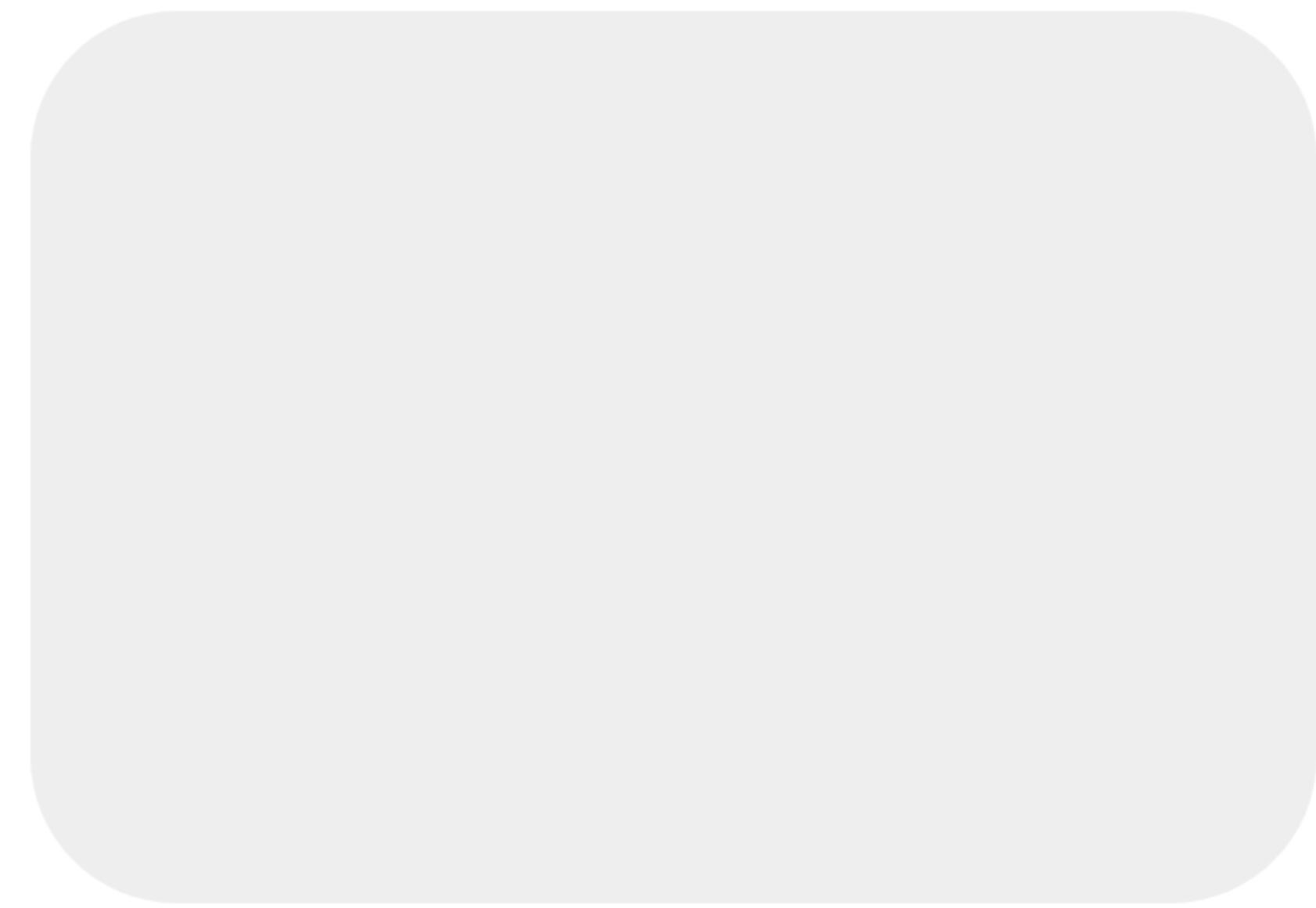
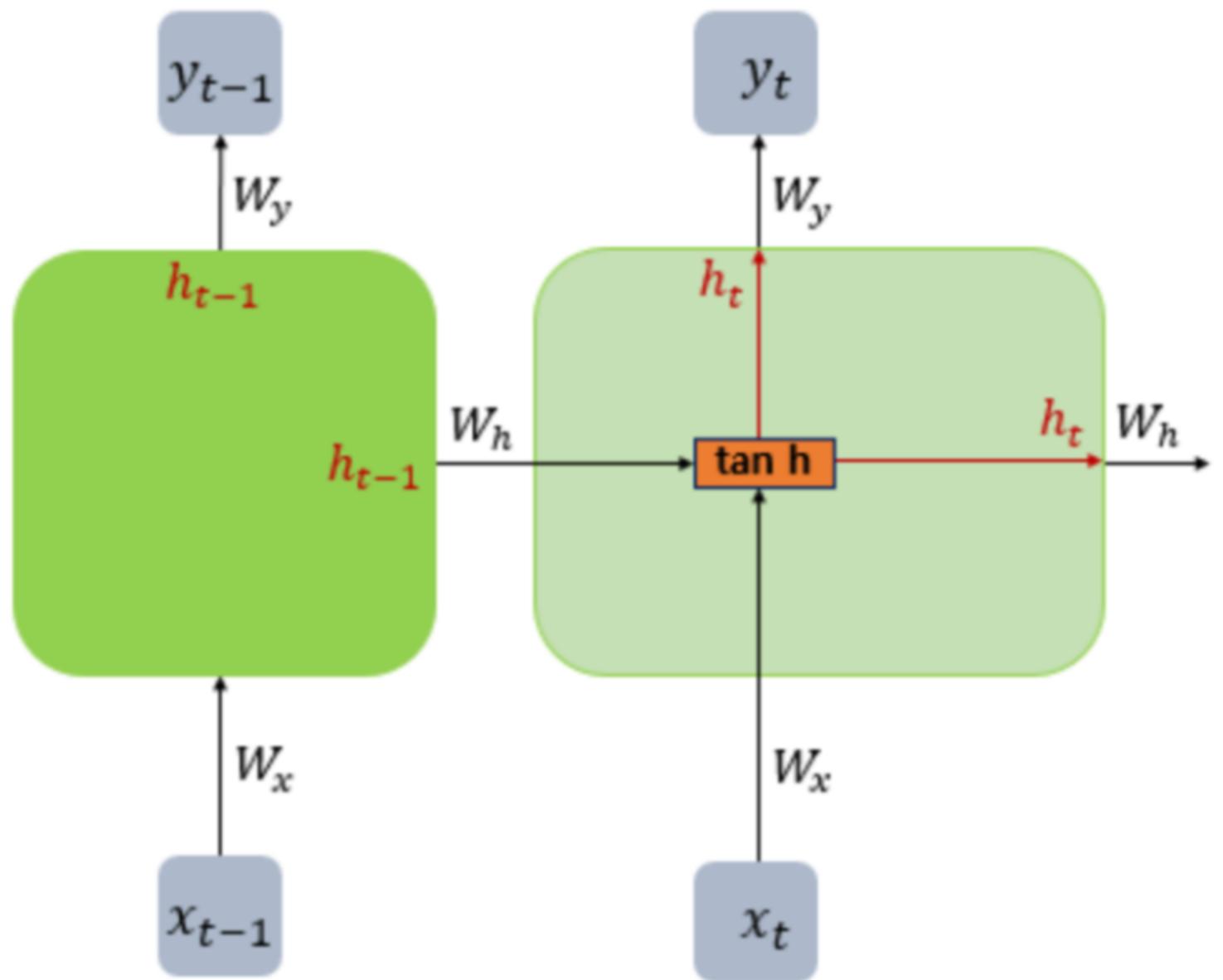
# LSTM의 필요성

RNN의 한계 : 시간이 흐르면 기억을 못한다!



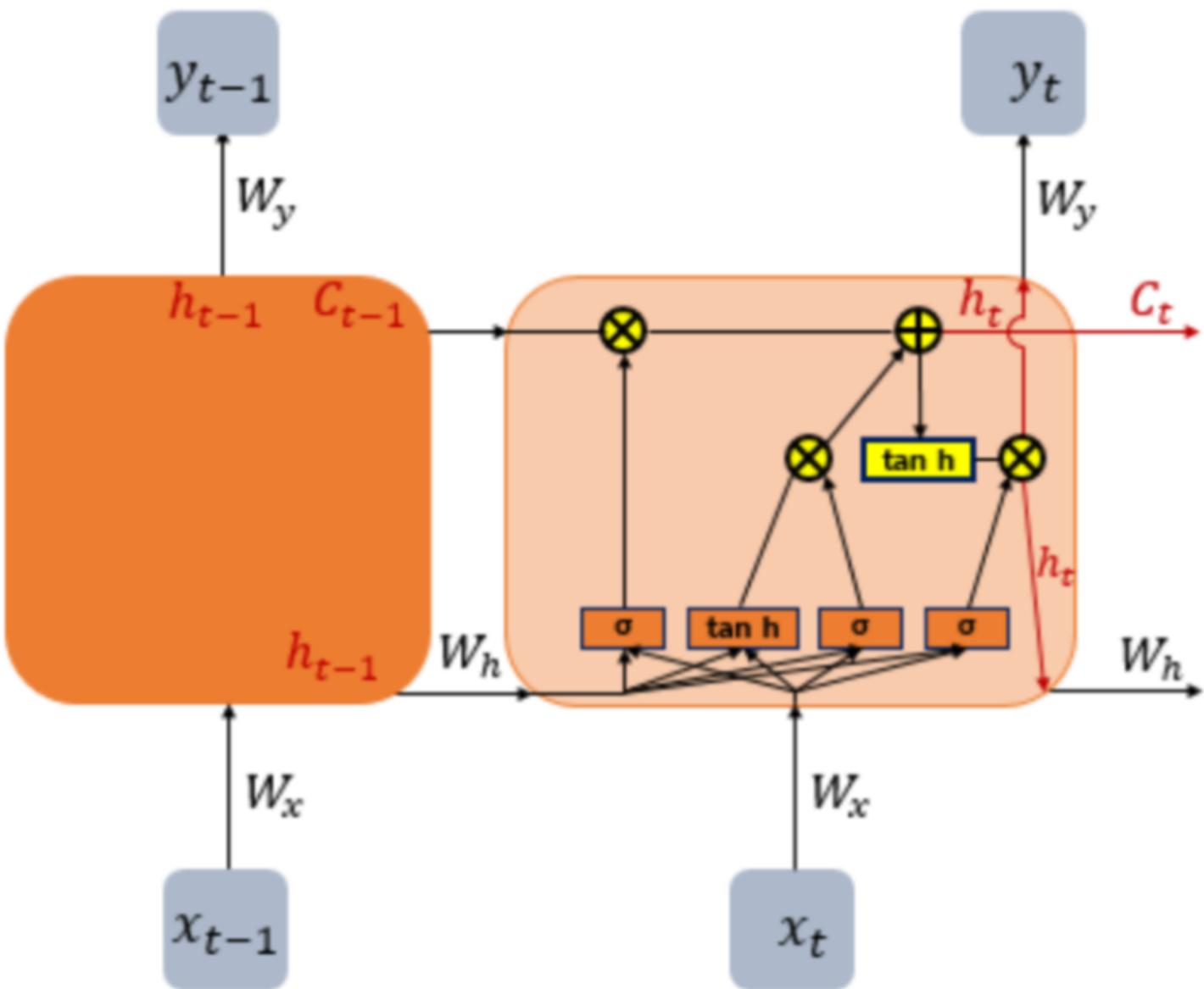
# LSTM의 필요성

RNN의 한계 : 시간이 흐르면 기억을 못한다!



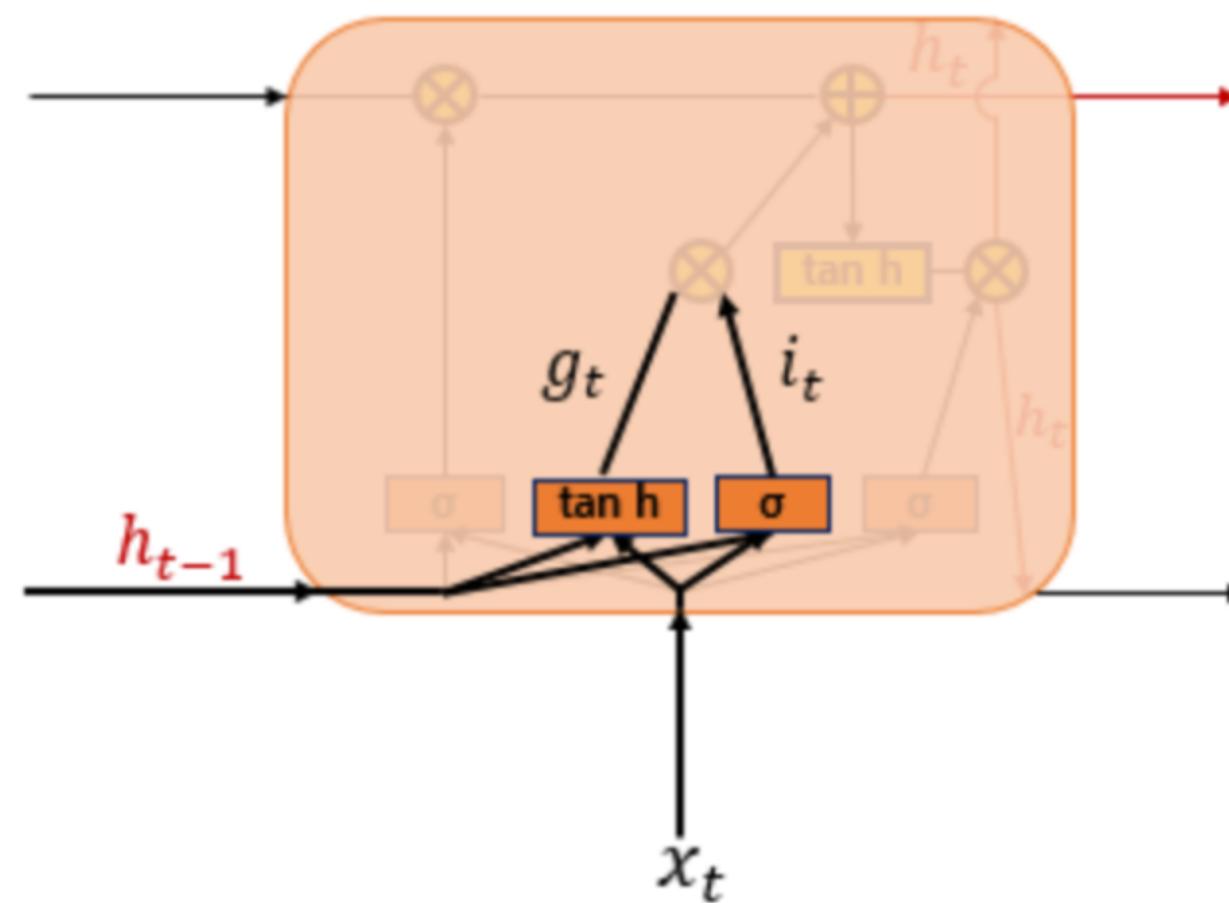
# LSTM

까먹지 않고 기억할 수 있게 하는 것.



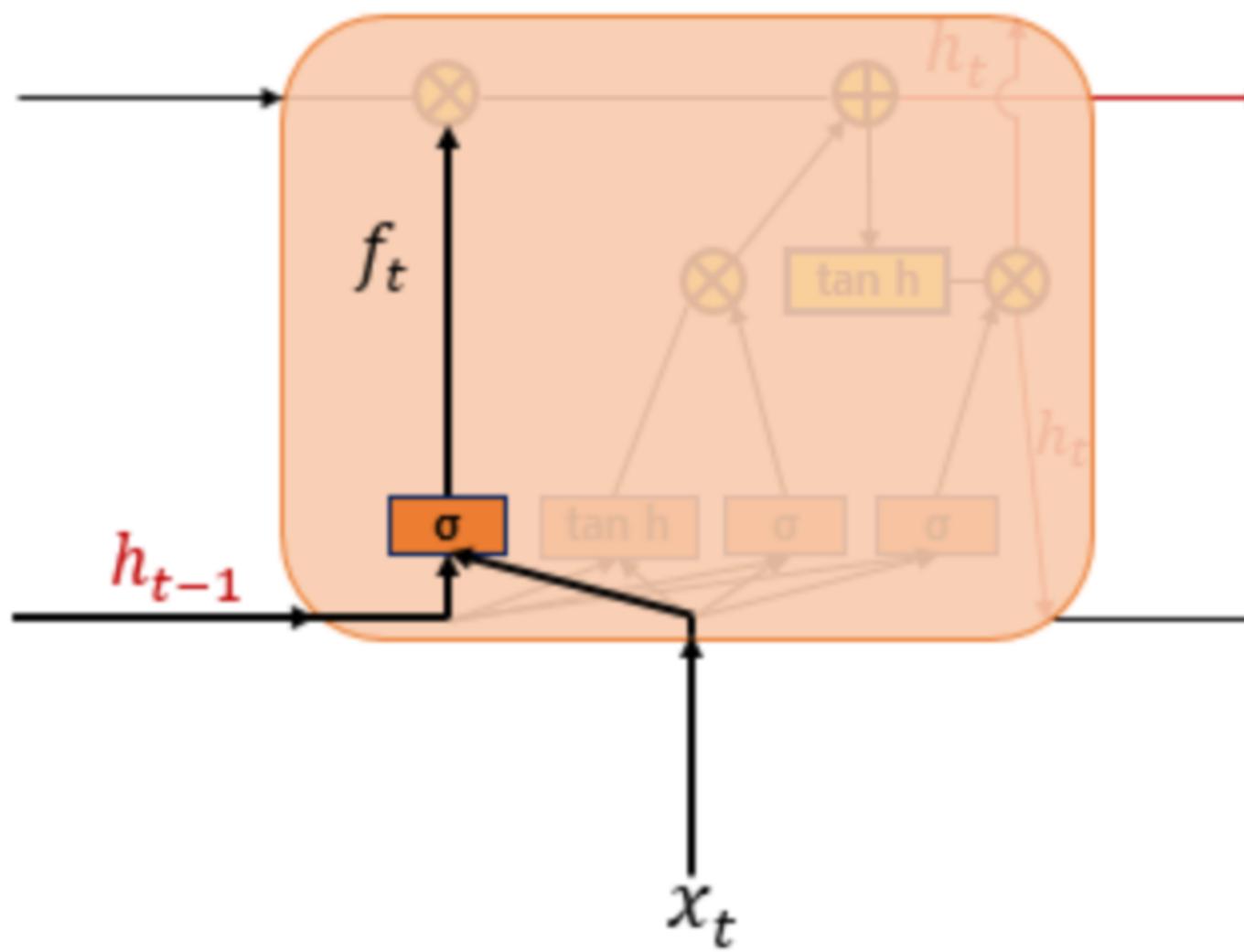
LSTM은 은닉층의 메모리 셀에  
입력 게이트, 망각 게이트, 출력 게이트를 추가하여  
불필요한 기억을 지우고, 기억해야 할 것들을 정한다.

# LSTM 입력게이트



현재 정보를 기억하기 위한 게이트

# LSTM 삭제 게이트



삭제 게이트는 기억을 삭제하기 위한 게이트 시그모이드 함수를 지나면 0과 1 사이의 값이 나오게 되는데, 이 값이 곧 삭제 과정을 거친 정보의 양

0에 가까우면 삭제가 많이 되었다는 뜻

• • ○

# LSTM 장기 상태/ 단기 상태

