



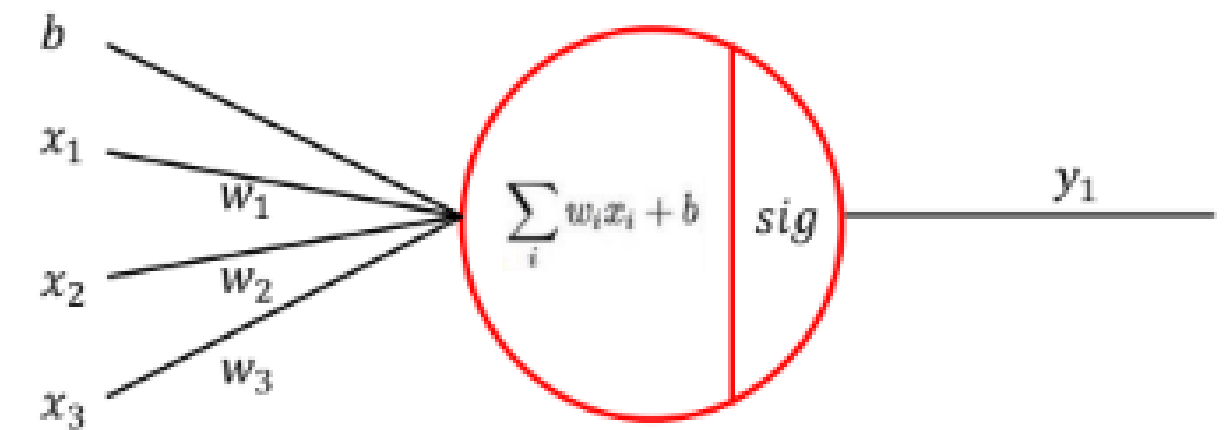
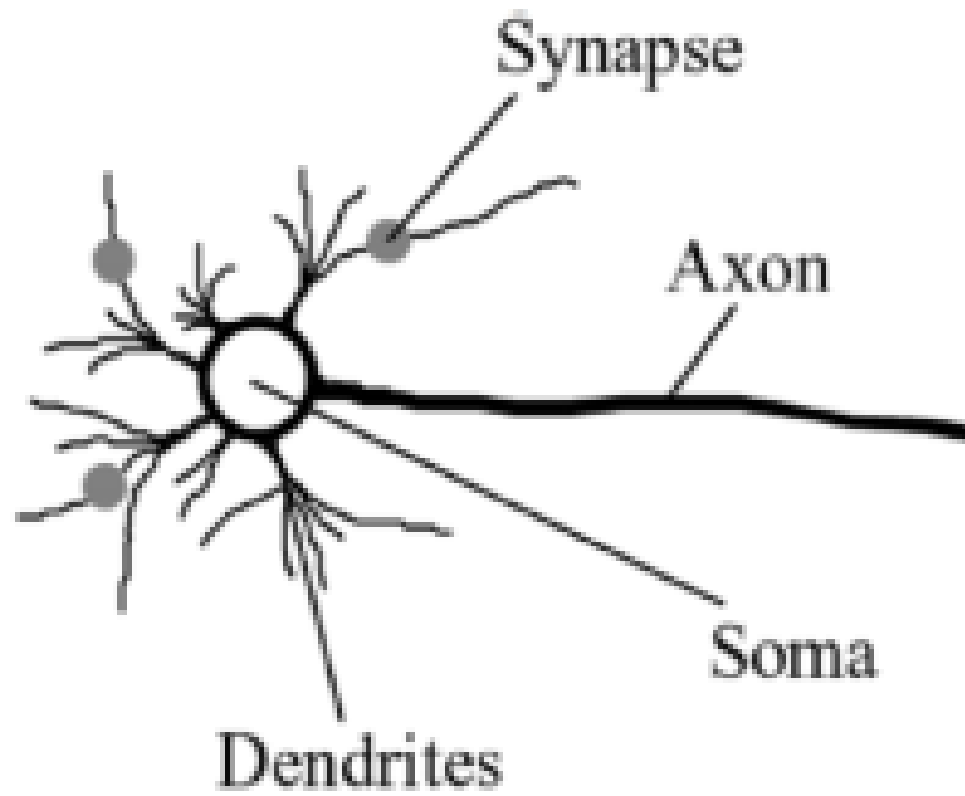
ML/DS 지식 따라가기

Perceptron, MLP, Loss Function



PERCEPTRON

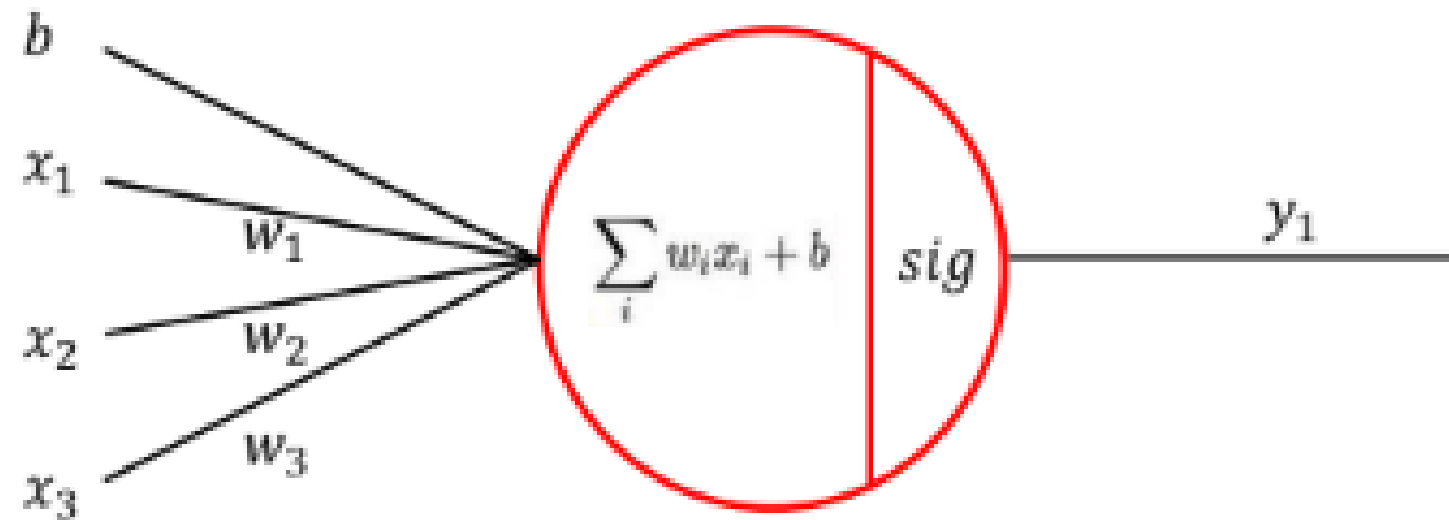
Neuron을 모방하여 만들어진 인공신경망



PERCEPTRON

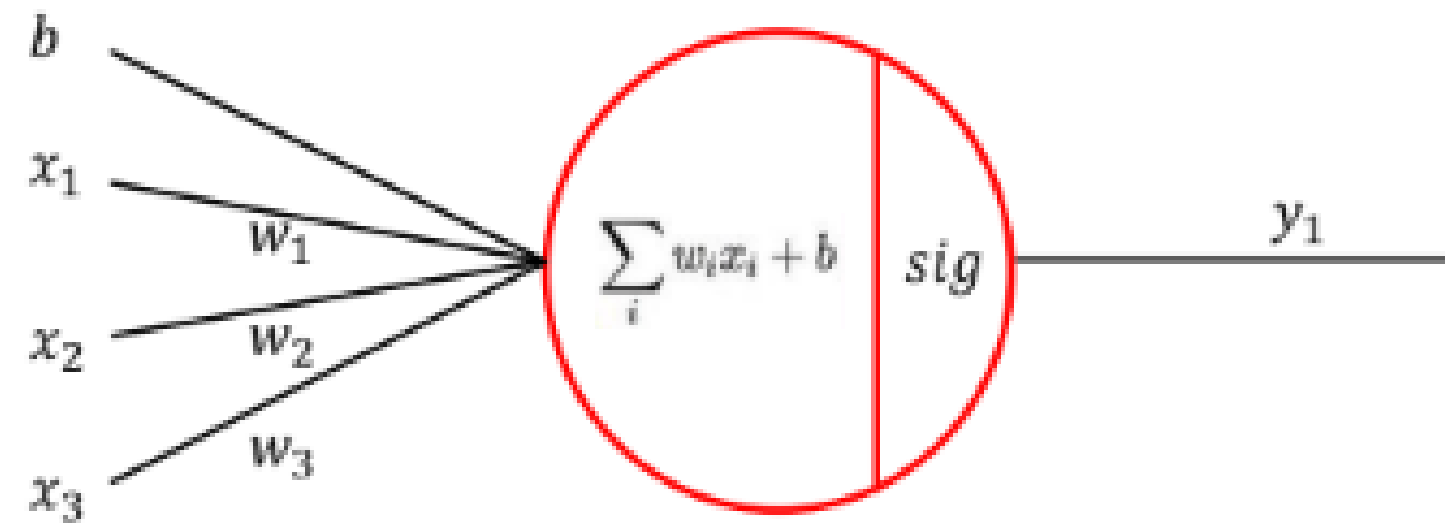
이항 분류를 위한 알고리즘

- 뉴런의 돌기들은 각각 input과 output이 됨
- 뉴런에서 각 synapse의 강도는 가중치가 됨



PERCEPTRON

Perceptron에서의 2가지 요소



(1) $Y=WX+B$

(2) Activation function

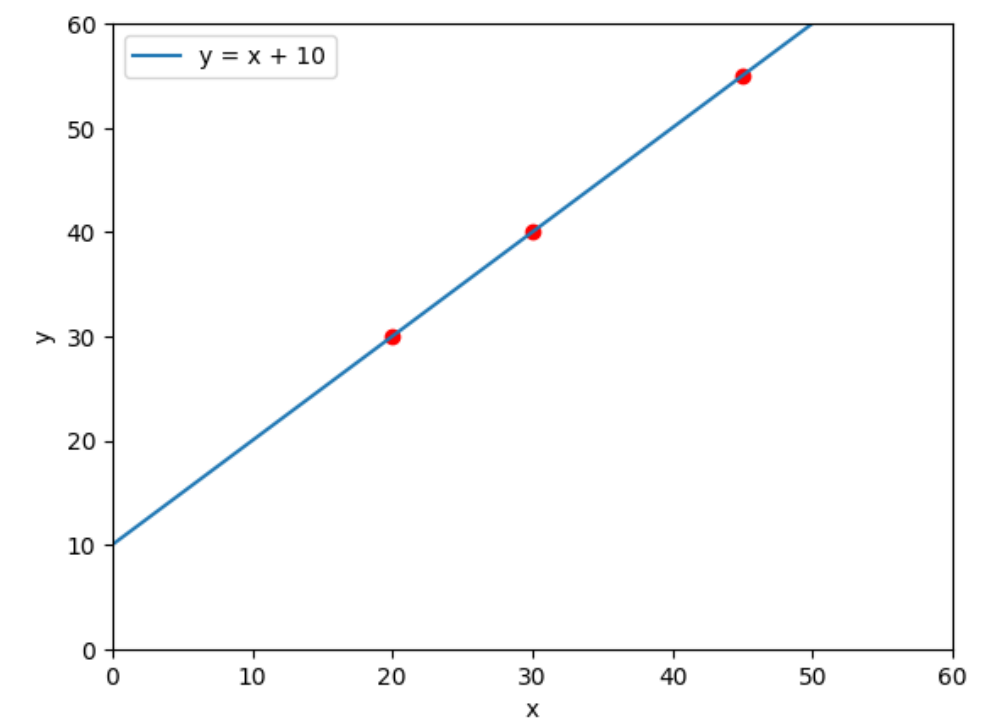
PERCEPTRON

$$(1) Y=WX+B$$

(1) Linear Regression

$$Y=WX+B$$

X	Y
20	30
30	40
45	???



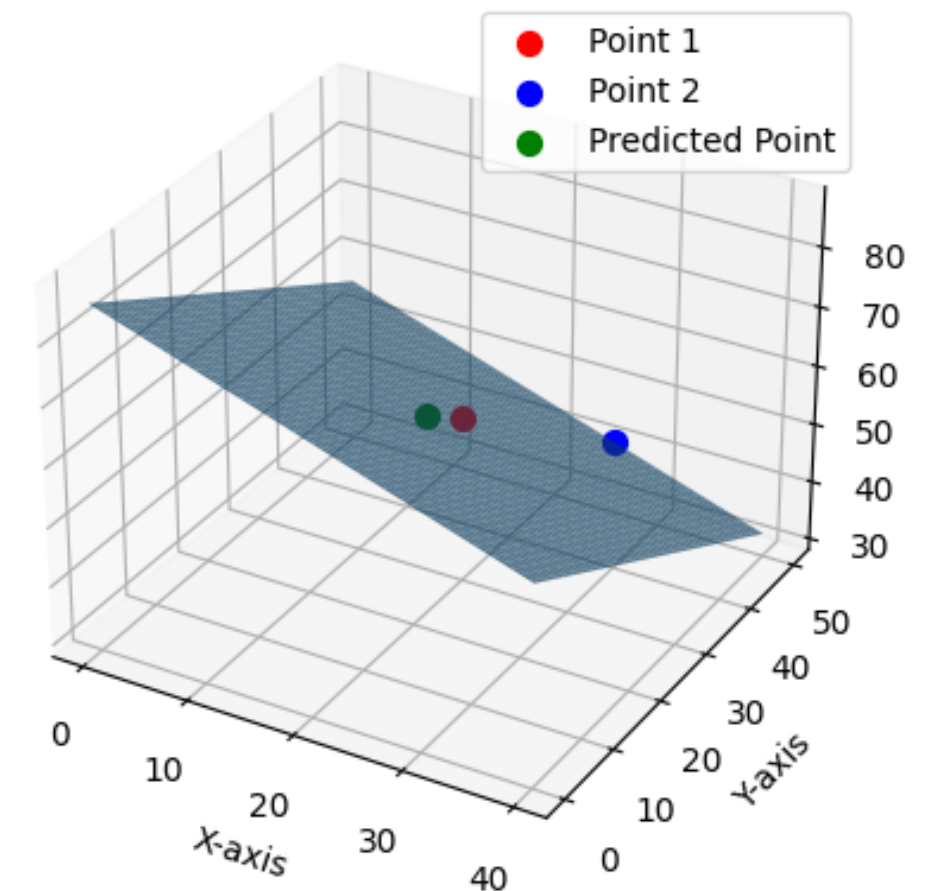
PERCEPTRON

(1) $Y = WX + B$

(2) Multivariate Regression

$$Y = W_1X_1 + W_2X_2 + \dots + B$$

X_1	X_2	Y
20	30	56
30	40	50
19	25	??



PERCEPTRON

(1) $Y=WX+B$

$$\left\{ \begin{array}{ccc} & X & \\ 20 & 30 & 23 \\ 30 & 40 & 47 \end{array} \right\} \quad \left\{ \begin{array}{c} Y \\ 56 \\ 50 \end{array} \right\}$$

(3) Regression in ML

$$Y=WX+B$$

$$W = \begin{bmatrix} 0.94849648 \\ 2.54056302 \\ -1.70377479 \end{bmatrix}$$

$$B = \begin{bmatrix} -4.97379915e-14 \\ -2.13162821e-14 \end{bmatrix}$$

PERCEPTRON

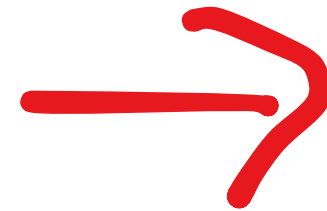
(2) Activation Function

$$Y=WX+B$$

(선형 회귀 모델)



Activation
Function



$$Y = \text{sigmoid}(WX+B)$$

PERCEPTRON

(2) Activation Function

- 앞선 입력들을 합쳐서 출력할 수 있도록 변환함

Sigmoid의 특성

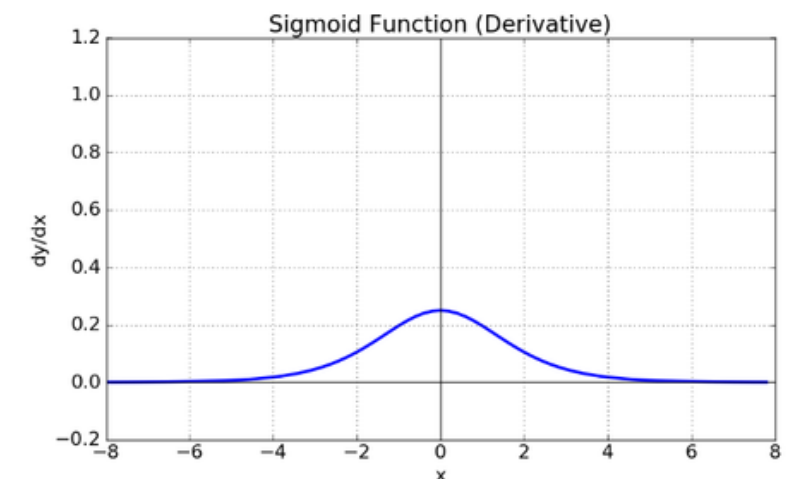
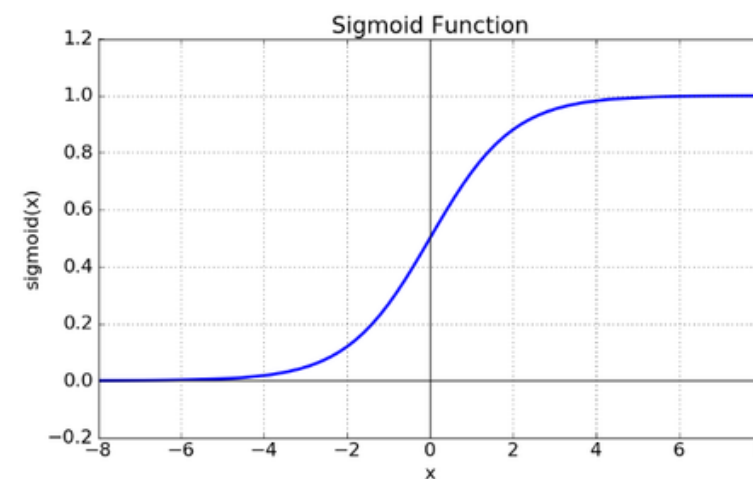
- Logistic이라고도 함
- S자형 구조임
- x값에 따라 0~1의 값을 가짐

1. Sigmoid

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

Sigmoid의 한계점

- Vanishing Gradient 문제



PERCEPTRON

2. Tanh

(2) Activation Function

- 앞선 입력들을 합쳐서 출력할 수 있도록 변환함

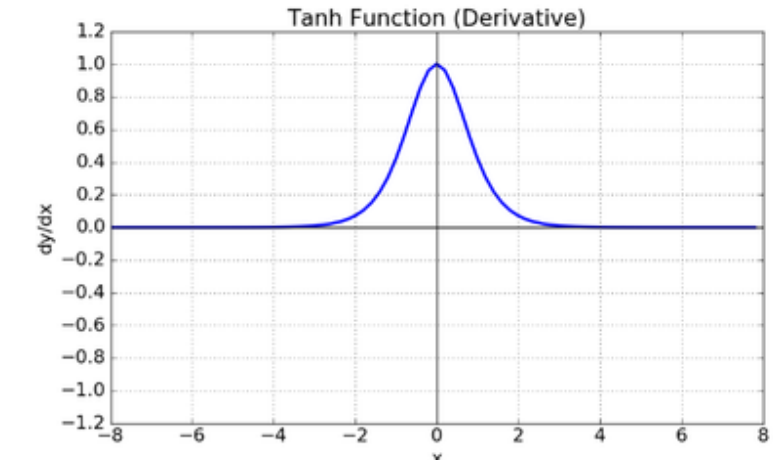
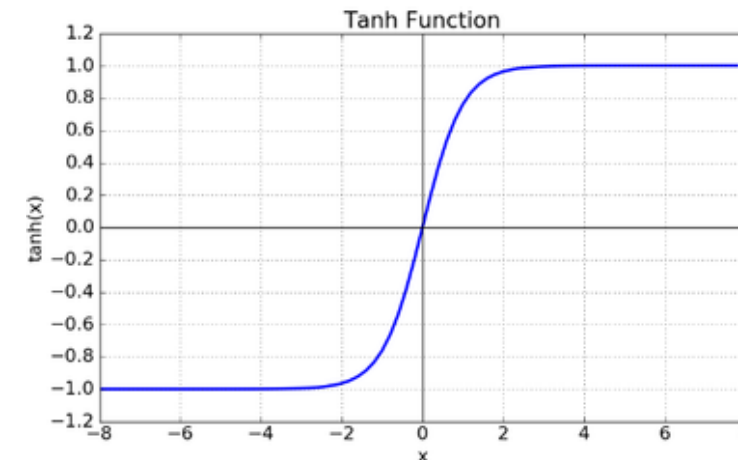
$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Tanh의 특성

- '쌍곡선'과 관련 있는 개념임.
- S자형 구조임
- 함수의 중심점이 원점임

Tanh의 한계점

- Vanishing Gradient 문제



PERCEPTRON

(2) Activation Function

- 앞선 입력들을 합쳐서 출력할 수 있도록 변환함

ReLU의 특성

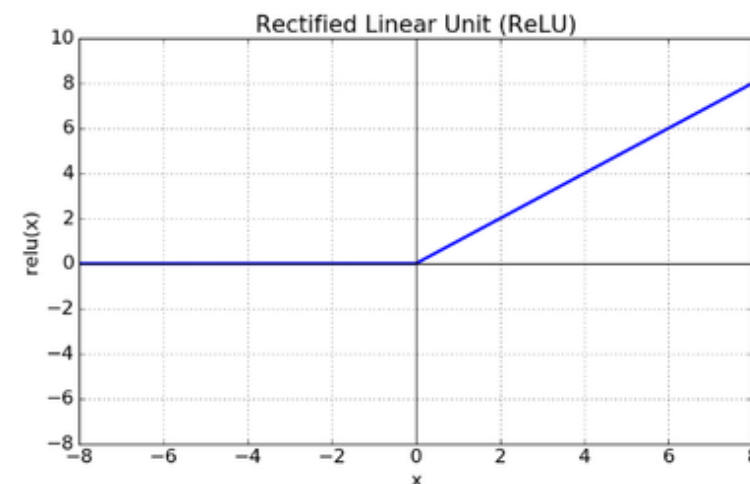
- Vanishing Gradient를 해결함
- 0보다 작으면 값이 0임
- 학습이 비교적 빠르고, 연산 비용이 적음

3. ReLU (Rectified Linear Unit)

$$f(x) = \max(0, x)$$

ReLU의 한계점

- Dying ReLU 문제



PERCEPTRON

(2) Activation Function

- 앞선 입력들을 합쳐서 출력할 수 있도록 변환함

Leaky ReLU의 특성

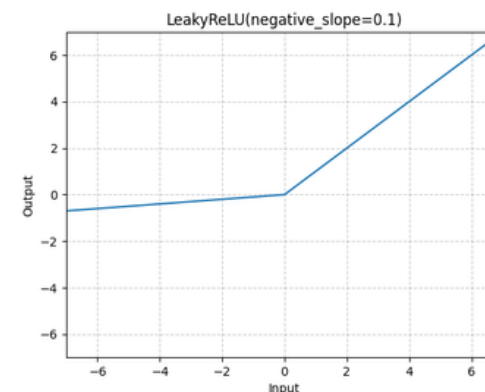
- Dying ReLU 문제를 해결함
- 0보다 작아도 기울기가 존재함
- 학습이 비교적 빠르고 연산 비용이 적음

4. Leaky ReLU

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

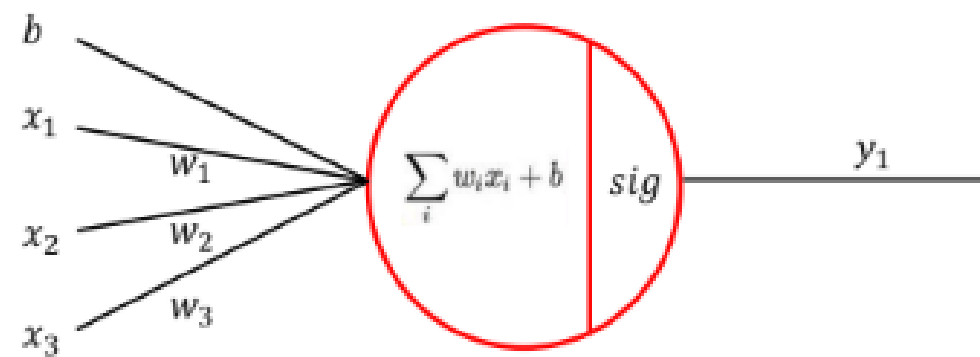
Leaky ReLU의 한계점

- 음수 파트가 중요할 때 쓰기 어려움

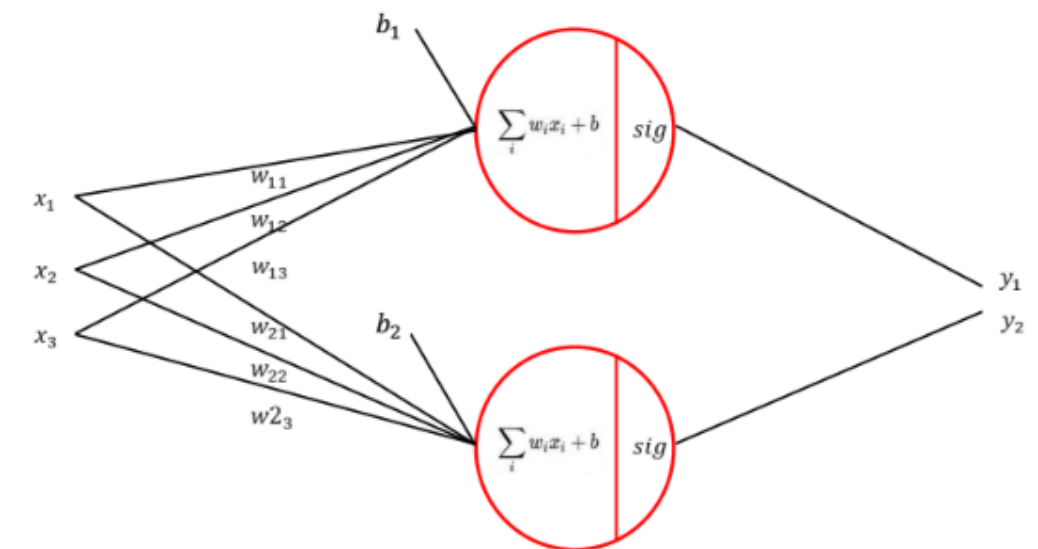


MLP

(1) MLP란? Multi-Layer Perceptron



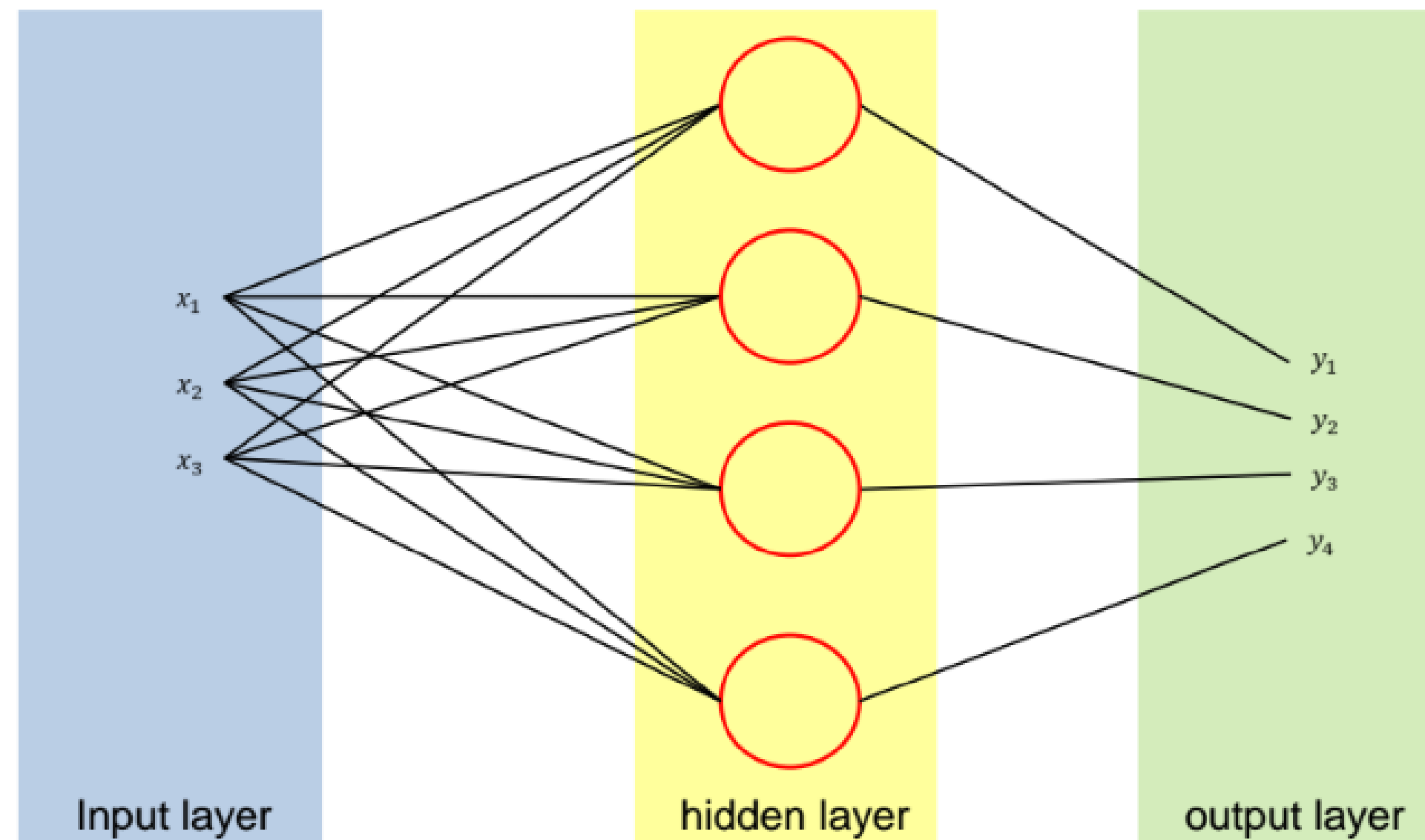
단층



다층

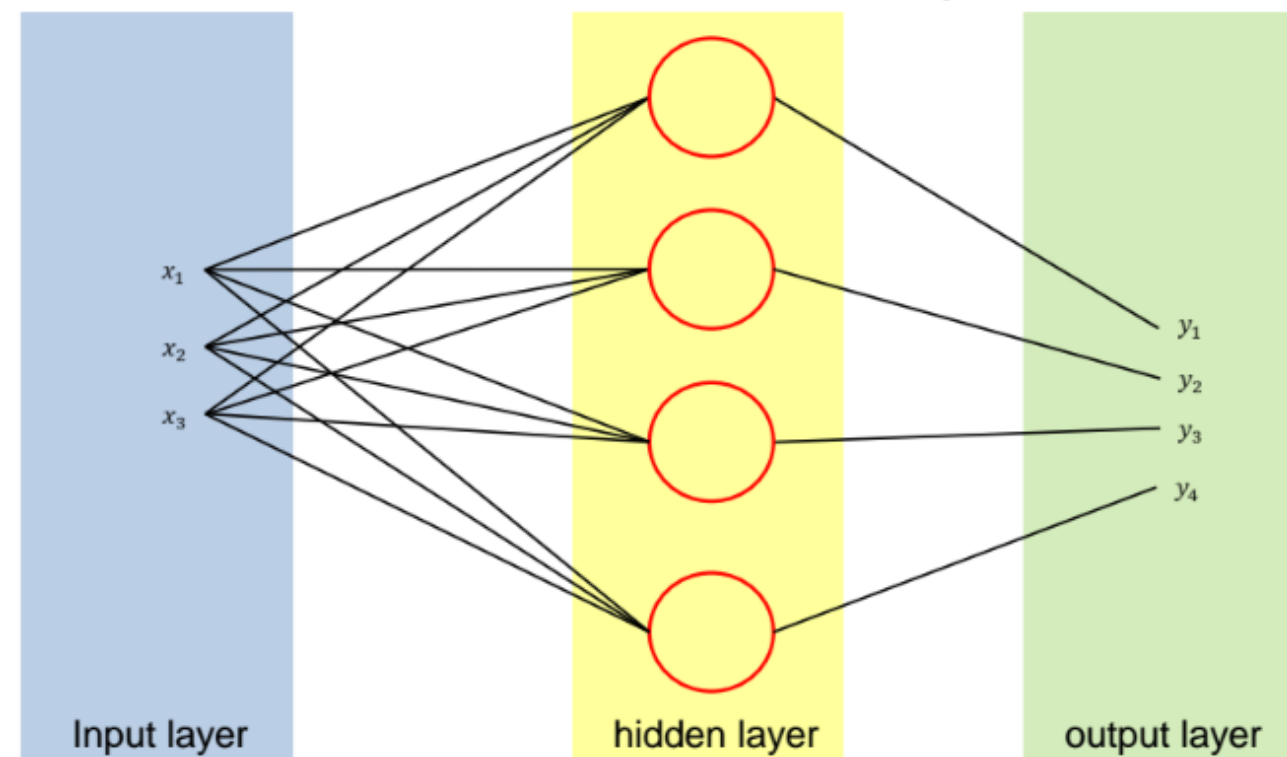
MLP

(1) MLP란?

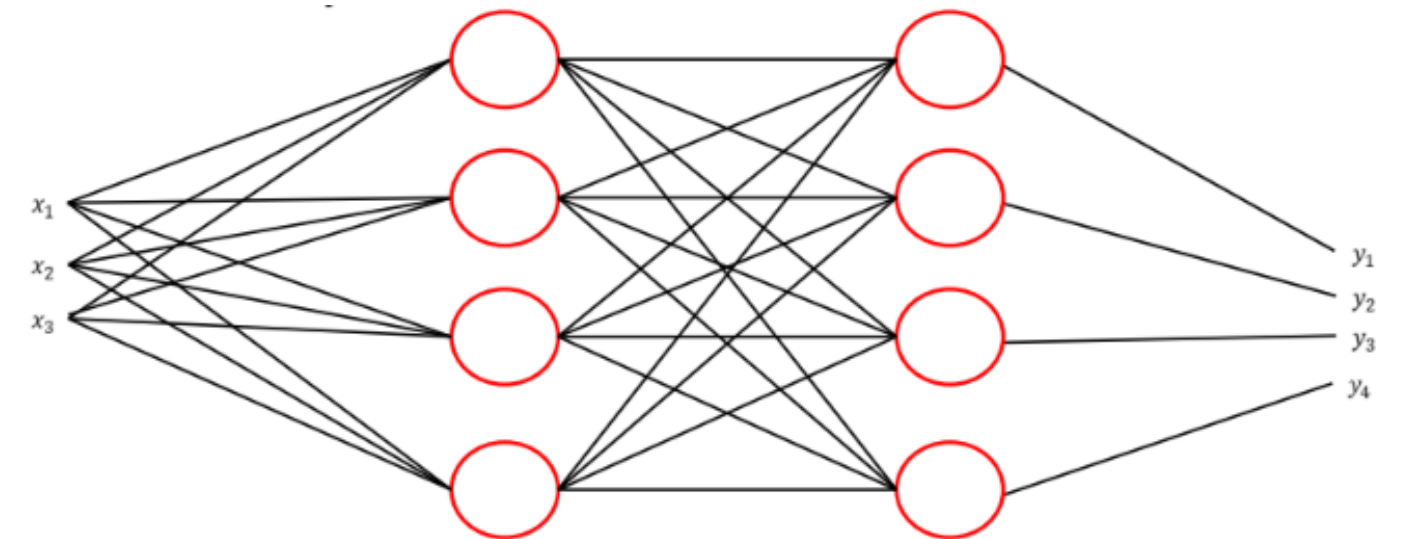


MLP

(1) MLP란?



ANN



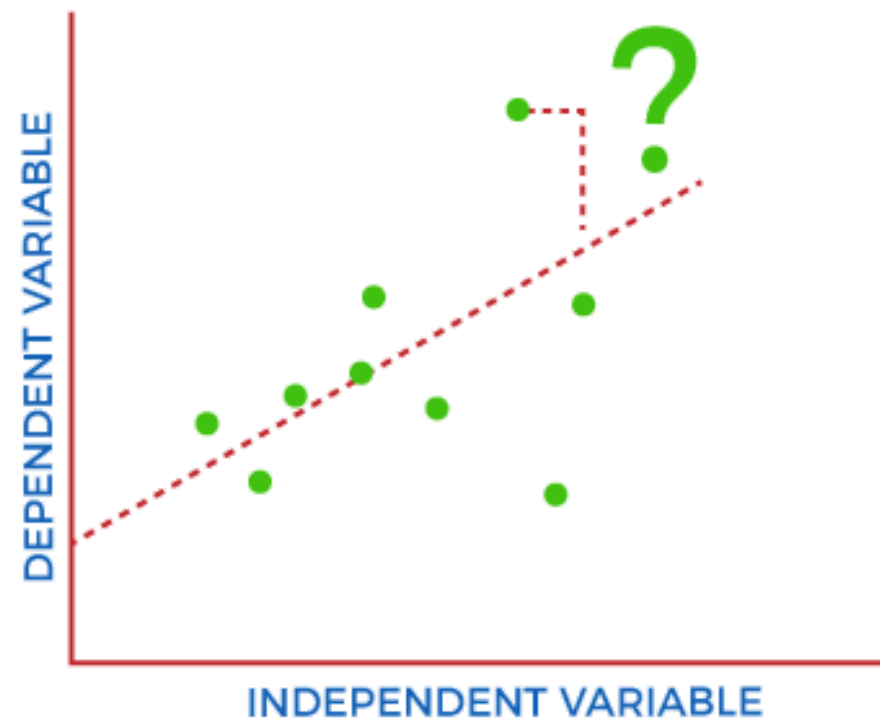
DNN

COST FUNCTION

(1) Cost Function 이란?

'실제 답'과 '학습의 결과'의 차이를 수치화한 것

COST FUNCTION IN MACHINE LEARNING



COST FUNCTION

(1) Cost Function의 종류

3주차의 Evaluation Technique을 참고

EVALUATION

Thank you