

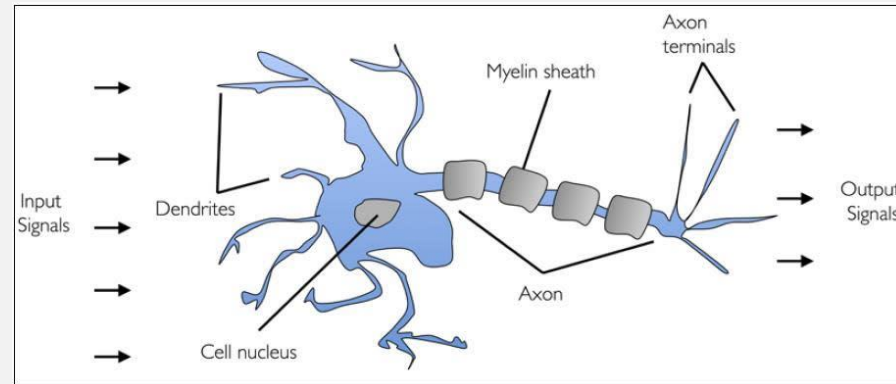
누구나 할 수 있는!

# 인터페이스 AI 오픈 스터디

4주차. 퍼셉트론과 심층 신경망 (DNNs)

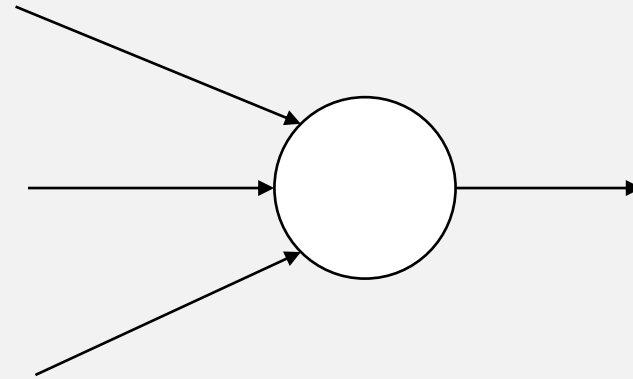
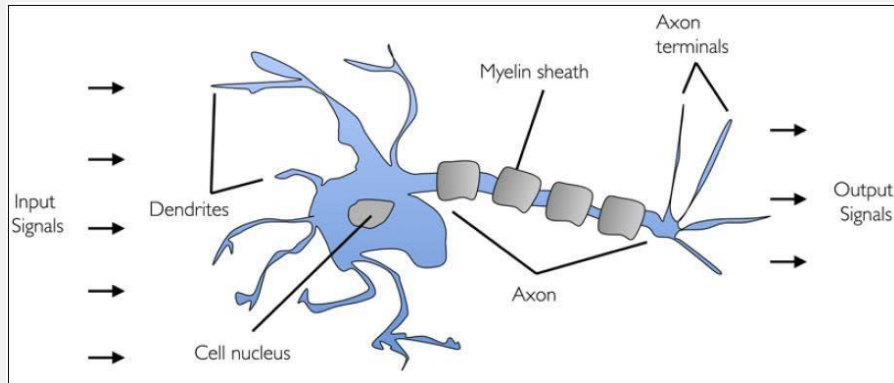
# How Our Brain Works

- 전기/화학적 자극을 전달받아, 이 자극의 합이 특정 임계점을 돌파하면 뒤쪽 뉴런을 향해 활성화 신호를 보냄. (Fire of Neuron 이라고 함)
- 이런 뉴런들이 여러 계층으로 모여 구성된 것이 바로 뇌.



# 퍼셉트론 (Perceptron)

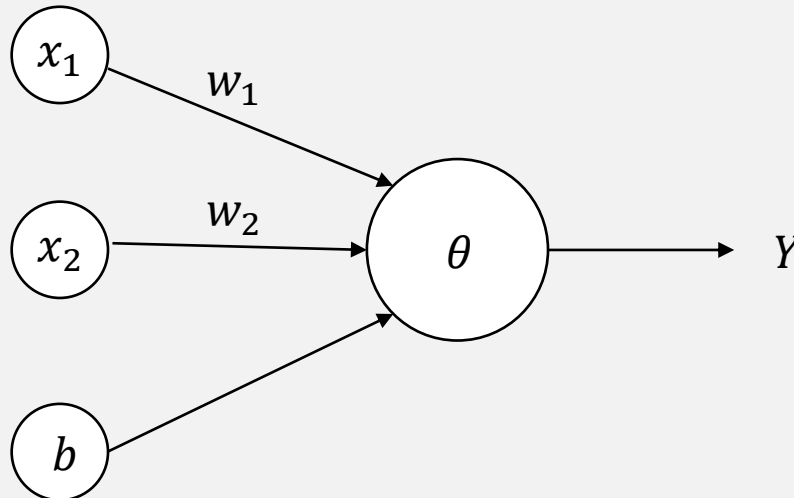
- 뇌를 컴퓨터에 구현하기 위해, 뇌의 부품이라고도 할 수 있는 뉴런을 구현한 퍼셉트론 모델 등장 (1958년)



# 퍼셉트론 (Perceptron)

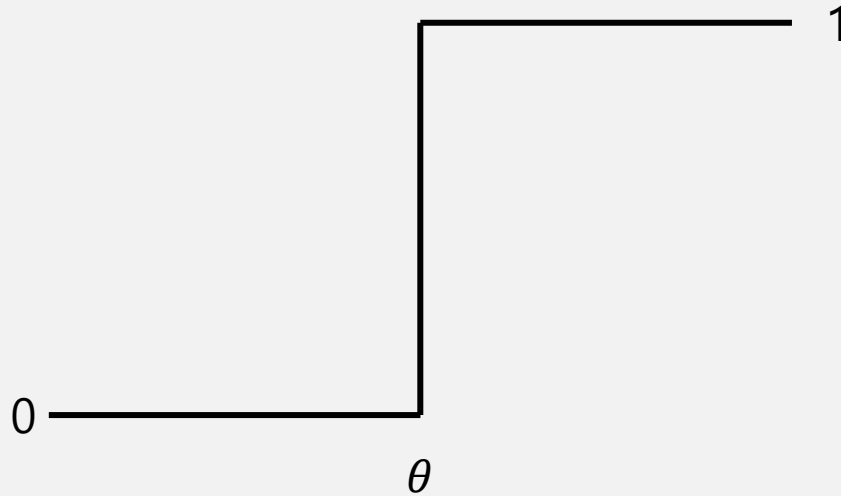
- 입력에 일정한 가중치( $W$ )를 곱하고 더하여, 이게 임계점 ( $\theta$ ) 을 넘으면 fire. (return 1)

$$Y(x) = \begin{cases} 1 & (wx + b > \theta) \\ 0 & (wx + b \leq \theta) \end{cases}$$



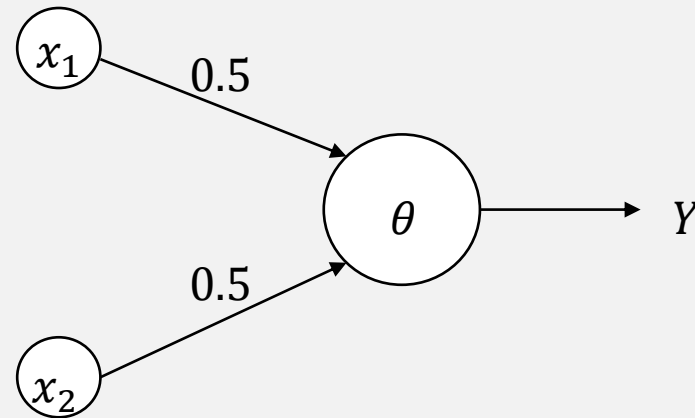
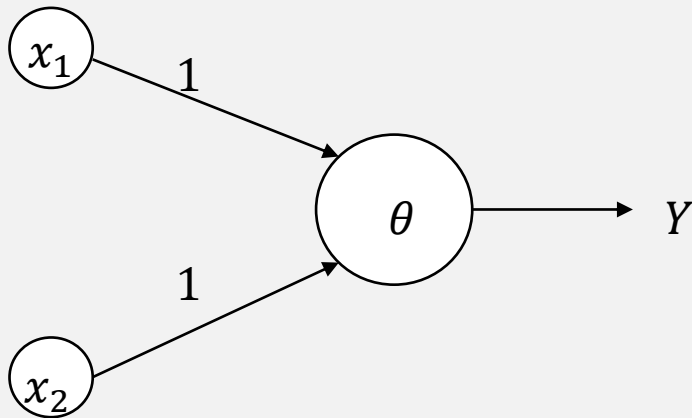
# 퍼셉트론 (Perceptron)

- 임계점을 넘으면 1, 아니면 0인 함수를 계단 함수 (Step Function)이라 한다.



# 퍼셉트론으로 다양한 문제 해결하기

- 아래는 각각 AND, OR문제 등을 해결하는 퍼셉트론  
(임계값이 모두 0.9일 때)



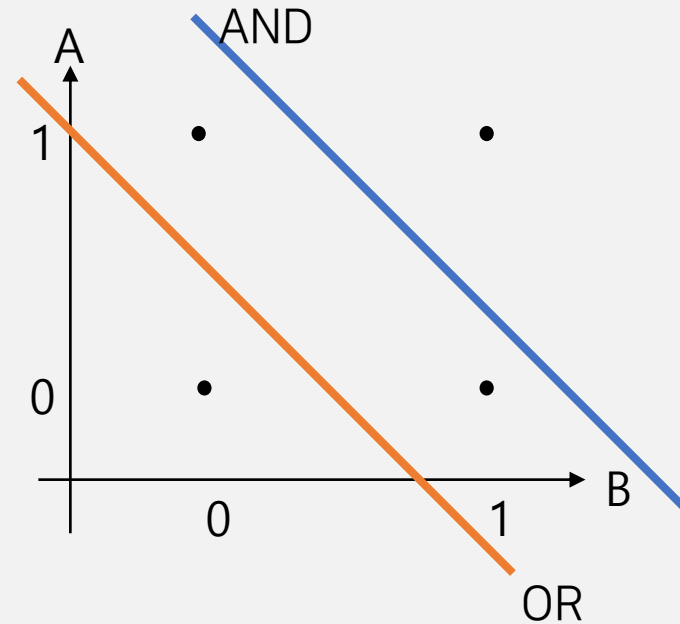
W, b,  $\theta$ 를 수정하여 여러가지 기능을 하는 퍼셉트론을 만들 수 있다!  
(퍼셉트론은 계단 함수가 있는 선형 회귀 모델이다!)

# 퍼셉트론의 한계: XOR 문제

- 퍼셉트론의 큰 한계는 XOR 문제를 풀지 못한다는 것이다.

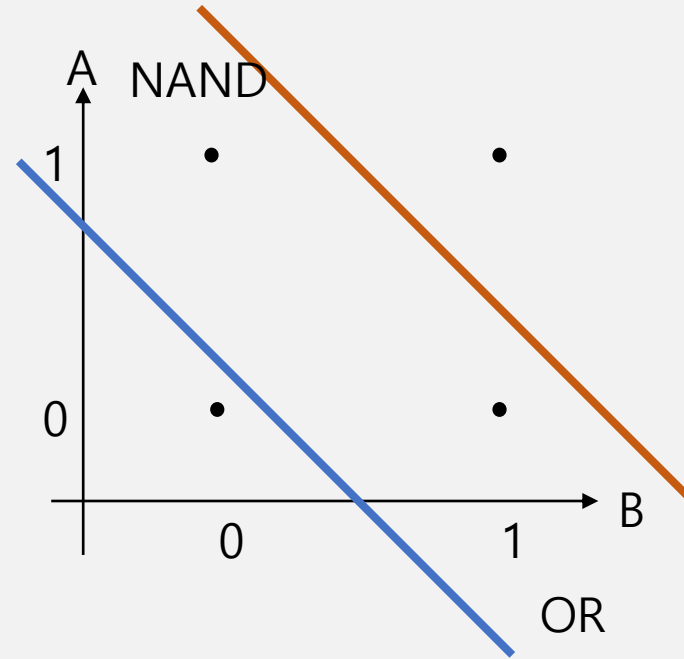
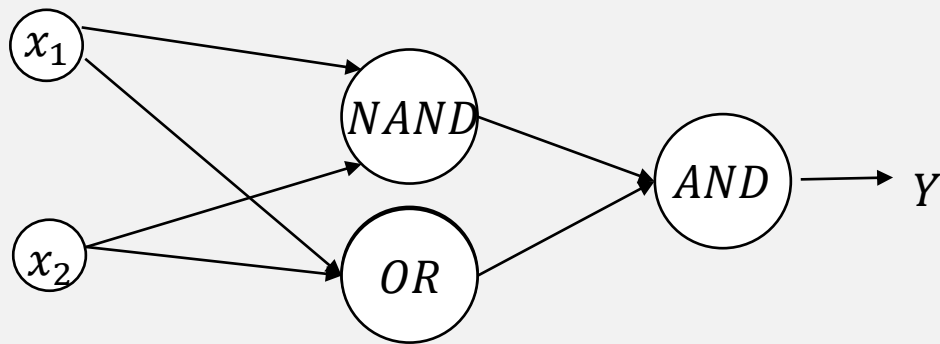
| A | B | A xor B |
|---|---|---------|
| 0 | 0 | 0       |
| 0 | 1 | 1       |
| 1 | 0 | 1       |
| 1 | 1 | 0       |

XOR 문제의 진리표



# XOR 문제의 해결: 다층 퍼셉트론(MLP)

- 실제 뇌처럼, 퍼셉트론을 여러 층으로 쌓아서 해결 가능!
- 다층 퍼셉트론 (Multi-Layer Perceptron)





# First AI Winter

- 퍼셉트론은 AND, OR, NAND 등의 문제를 풀 수 있지만, XOR과 같은 단순한 문제조차 해결하지 못 한다.
- 다층 퍼셉트론의 경우 다양한 문제를 해결 가능하지만, 이는 지도학습이 불가능하다!
- First AI Winter의 시작

# 역전파 알고리즘 (Back Propagation)

- 여러 계층으로 쌓인 퍼셉트론의 학습을 가능하게 해주는 알고리즘.
- 데이터가 한번 신경망을 통과하며 각 계층별로 Error를 산출하고, 결과에 기반하여 오차를 뒤로 전파(Propagate)시킴.
- 조건: 각 계층이 미분가능한 활성화 함수를 가져야 한다!

# 미분 가능한 활성화 함수

- 지금까지는 미분 불가능한 함수인 계단 함수를 사용했지만, 이제 미분 가능한 함수를 사용하자!
- 초기에는 계단 함수 대신 Sigmoid 함수를 사용
$$Y(x) = \text{Sigmoid}(WX + b)$$
- 그러나, 모델이 커지면 학습이 진행되지 않는 경사 소실 (Vanishing Gradient) 문제 발생 (2<sup>nd</sup> AI Winter)

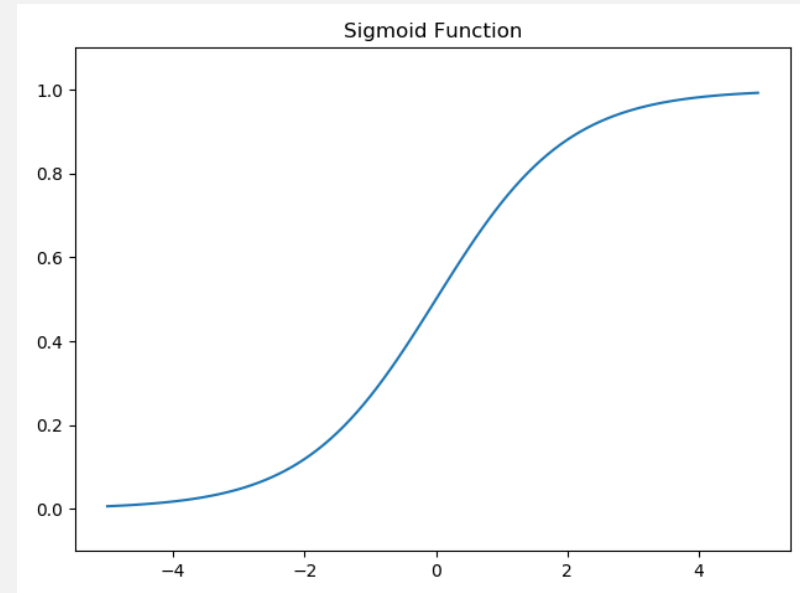
# 경사 소실 문제 (Vanishing Gradient)

- 신경망의 크기가 너무 커지자, 역전파를 이용한 학습이 이루어지지 않는 문제 발생.
- 왠지 모를 이유로, 신경망이 깊어질 수록 경사가 0에 수렴해 사라짐.
- 이로 인해 한동안 AI 투자가 줄어드는 AI Winter의 재림

# 경사 소실 문제 (Vanishing Gradient)

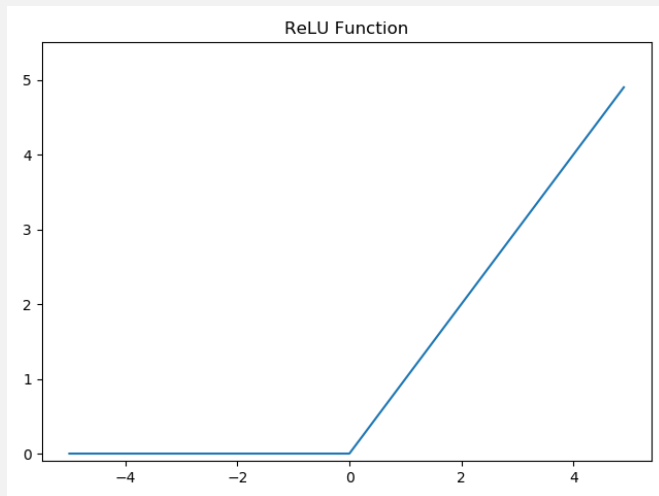
- 재프리 힌튼과 그의 동료들이 연구한 결과, 원인은 활성화 함수로 쓰인 Sigmoid 함수였음!
- 모든 값을 0과 1 사이로 축소하는 Sigmoid 함수 때문에, 경사가 점점 0.000... 으로 작아짐!
- 작은 신경망에서는 상관없지만, 신경망이 어느 정도 깊어지면 경사가 사라짐.

**새로운 활성화 함수가 필요하다!**



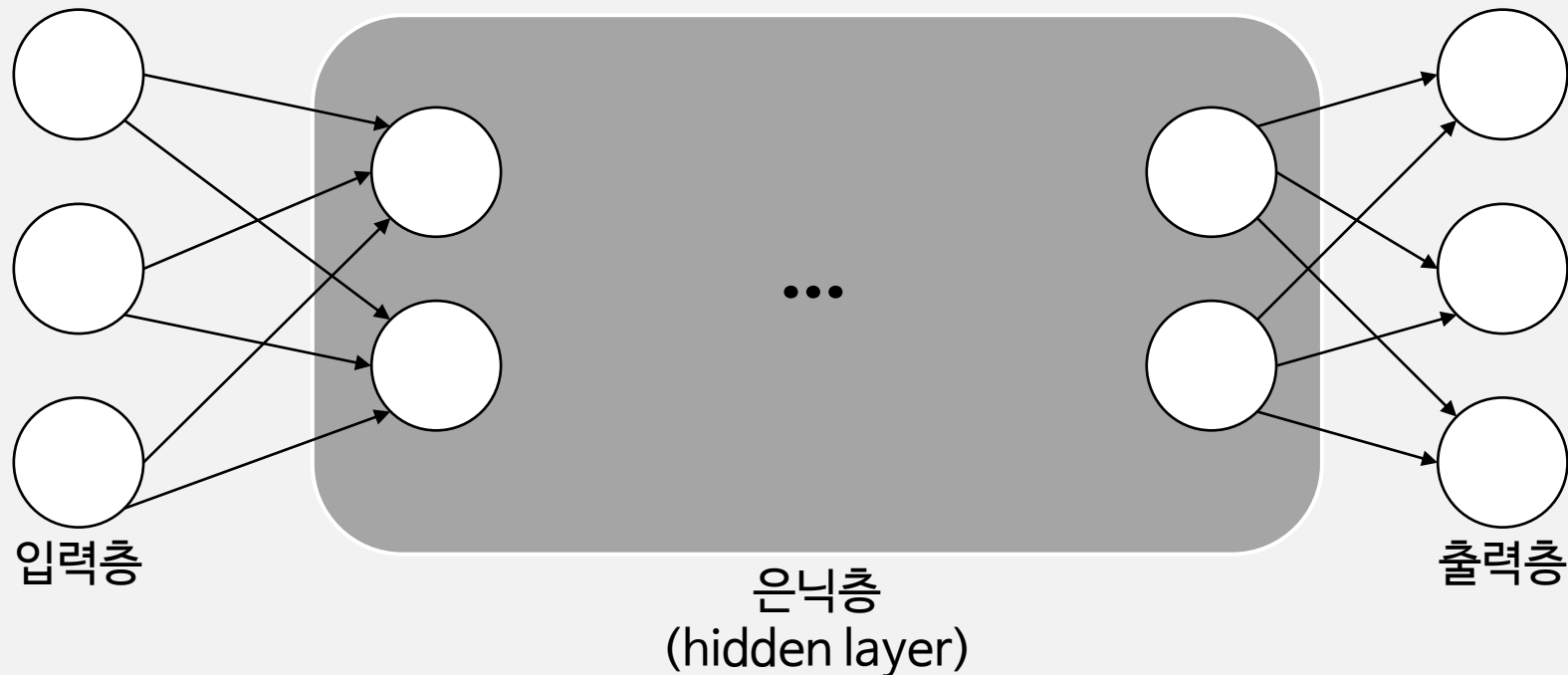
# ReLU 함수

- 정류된 선형 함수(Rectified Linear Unit)이라는 뜻.
- 0 이하의 값은 0을 리턴하고, 이외에는 그대로 리턴하는 함수
$$ReLU(x) = \max(0, x)$$
- 값의 표현 범위는 무한하고 ( $0 \sim \infty$ ), 미분 가능함.



# 심층 신경망 (Deep Neural Network)

- 비선형/미분가능한 활성화 함수를 갖는 다층 퍼셉트론(MLP)이라 볼 수 있음
- 층을 쌓을 수록 고차원적인 특징을 학습할 수 있음  
(점 → 선 → 도형 → 객체 등..)



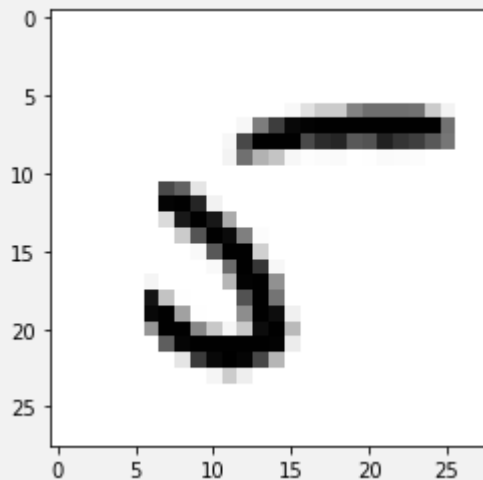
# 오늘의 키워드

- 퍼셉트론 (Perceptron)
  - 계단 함수 (Step Function)
  - XOR 문제의 해결
  - 다층 퍼셉트론 (MLP)
- 심층 신경망 (Deep Neural Network: DNNs)
  - 비선형 활성화 함수
    - ReLU 함수
  - 경사 소실 (Vanishing Gradient)
  - 입력층 / 은닉층 / 출력층



# MNIST 실습

- 손으로 쓴 정수 이미지를 분류하는 문제! (Multi-Class Classification)
- $28 * 28$  크기의 숫자 이미지가 입력으로 주어짐.
- 이를 구분하는 DNN을 만들어보자!



**감사합니다!**