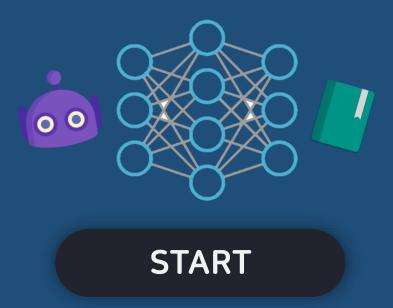
Deep Learning

Chapter 2 퍼셉트론, 다층 퍼셉트론(Perceptron, MLP)

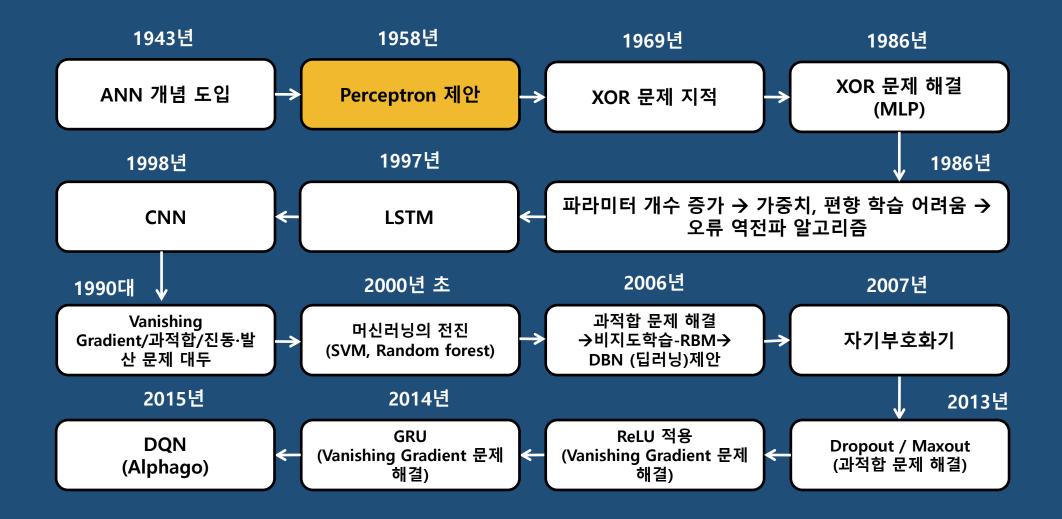




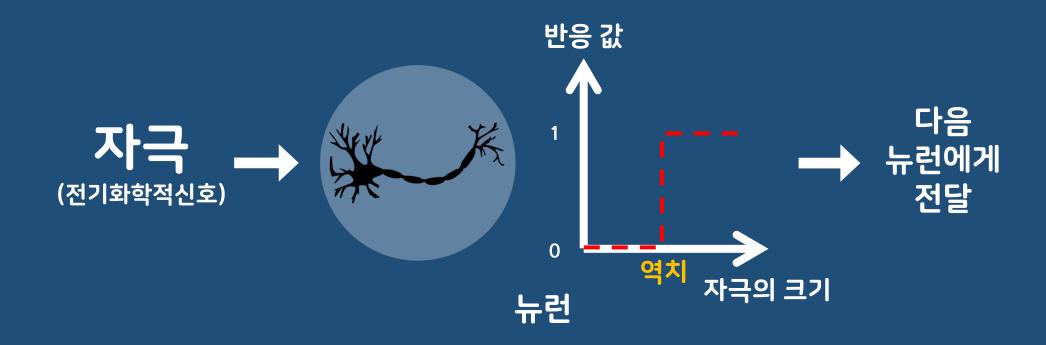
- 퍼셉트론의 개념을 이해 할 수 있다.
- 다층 퍼셉트론의 개념을 이해 할 수 있다.



딥러닝 역사 – 퍼셉트론 제안



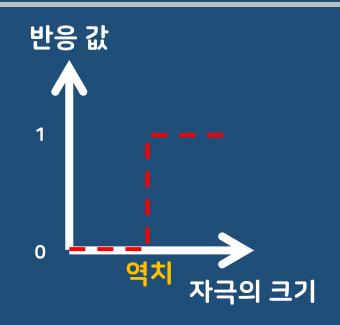




신경의 흥분이 전달되기 위해서는 뉴런에 전달되는 자극의 크기가 <mark>역치 이상</mark>이 돼야함



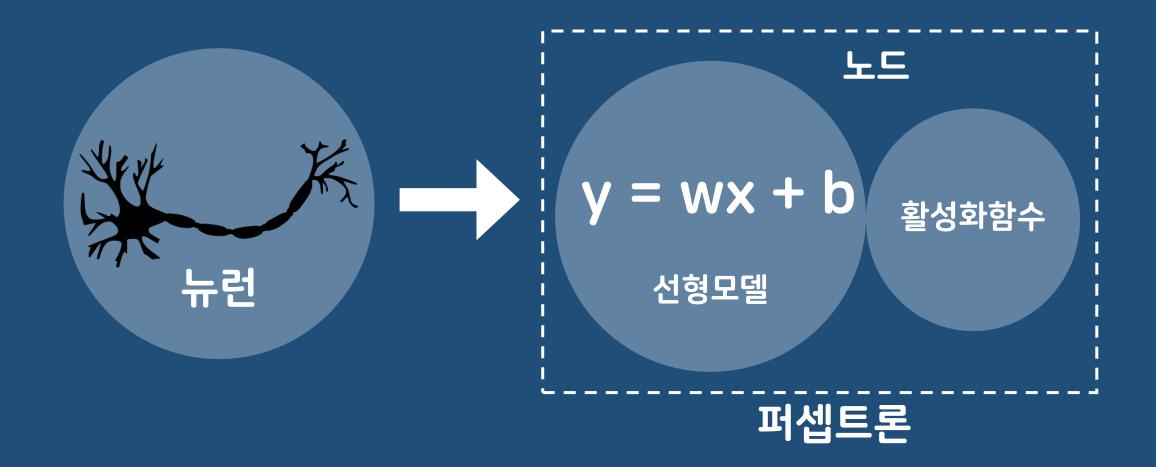














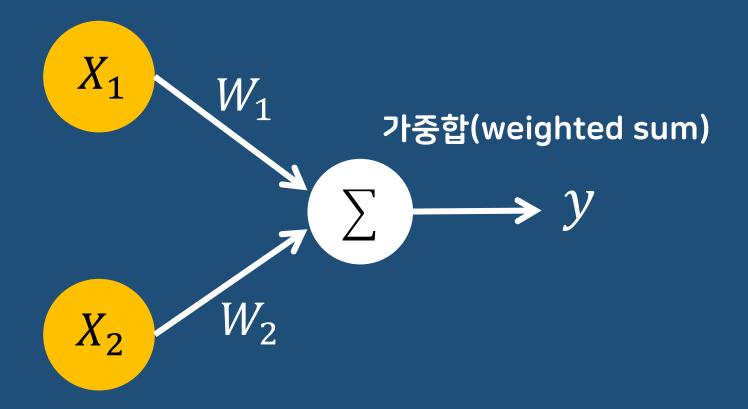
퍼셉트론 (Perceptron)

프랑크 로젠블라트가 1957년에 고안한 알고리즘

The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain

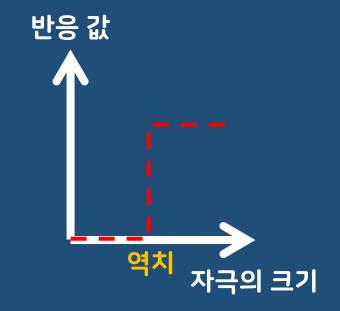


$$y = W_1 X_1 + W_2 X_2 + b$$









$$f(x)$$
 Activation function (활성화함수)

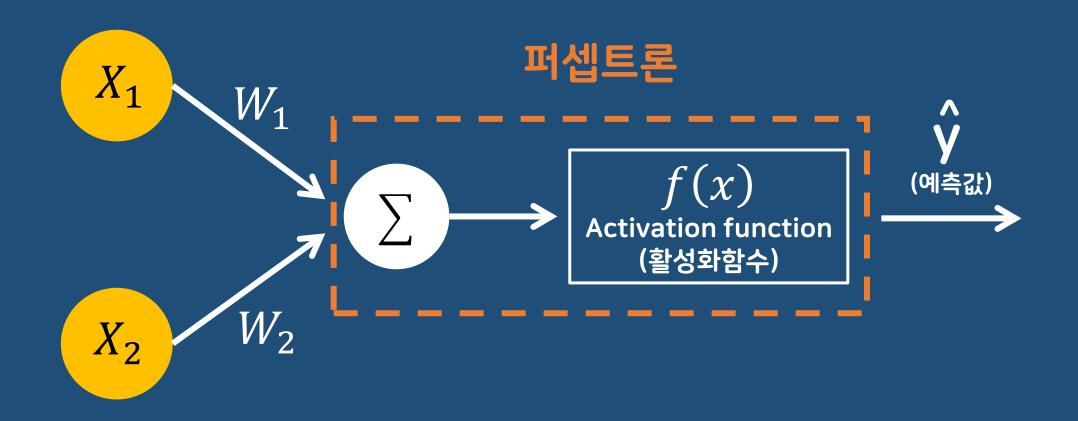


$$\mathbf{y} = \begin{cases} \mathbf{0}, & (W_1 X_1 + W_2 X_2 + b \leq 0) \\ \mathbf{1}, & (W_1 X_1 + W_2 X_2 + b > 0) \end{cases}$$

 W_1,W_2 : 가중치 (weight) – 각 입력 신호가 결과에 주는 영향력을 조절하는 매개변수

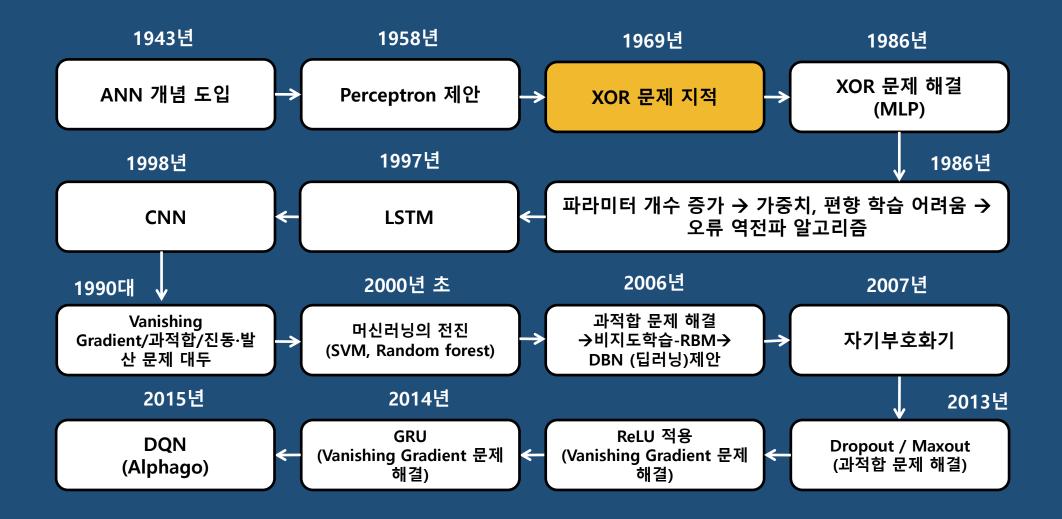
 $oldsymbol{b}$: 편향 (bias) – 뉴런이 얼마나 쉽게 활성화하느냐를 조절하는 매개변수





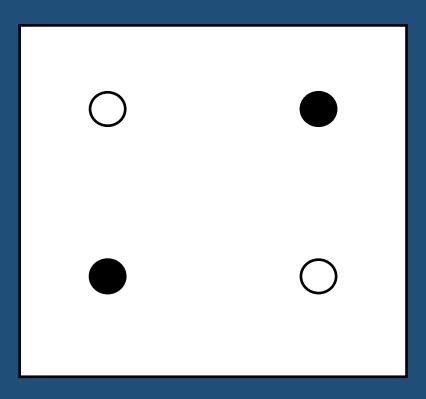


딥러닝 역사 - XOR 문제





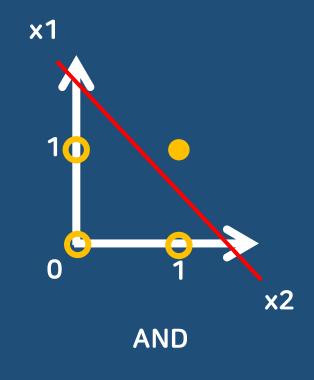
딥러닝 역사 - XOR 문제





AND 게이트

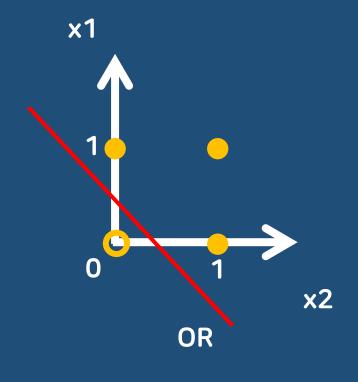
| x1 | x2 | AND |
|-----------|-----------|-----|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |





OR 게이트

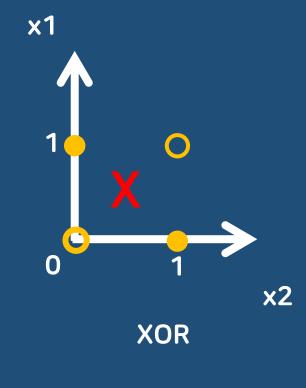
| x1 | x2 | OR |
|-----------|-----------|----|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 1 |





XOR 게이트

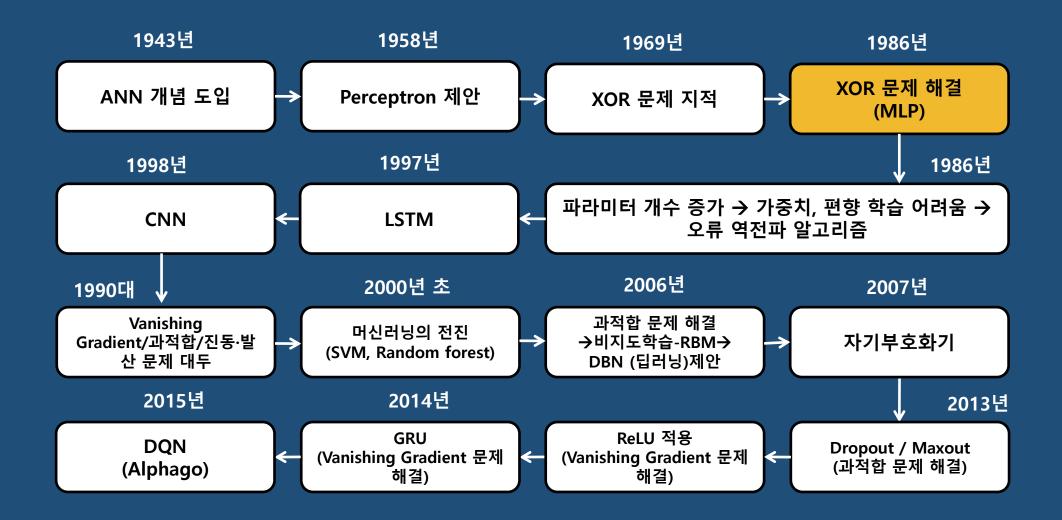
| x1 | x2 | XOR |
|-----------|-----------|-----|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |



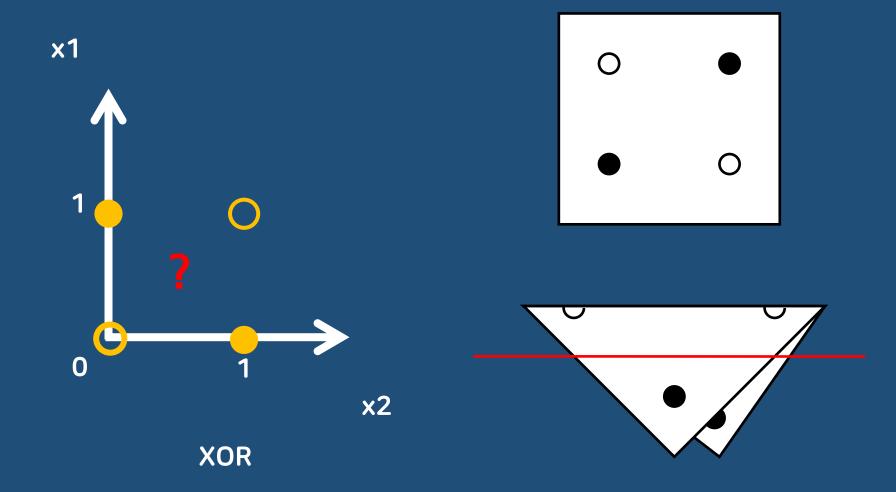


AND,OR는 해결이 가능하지만 간단한 XOR 문제를 해결 할 수 없었다..!







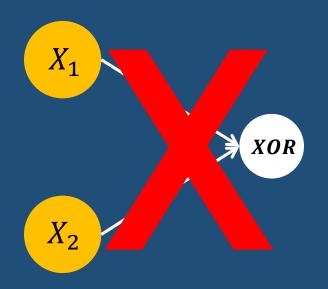




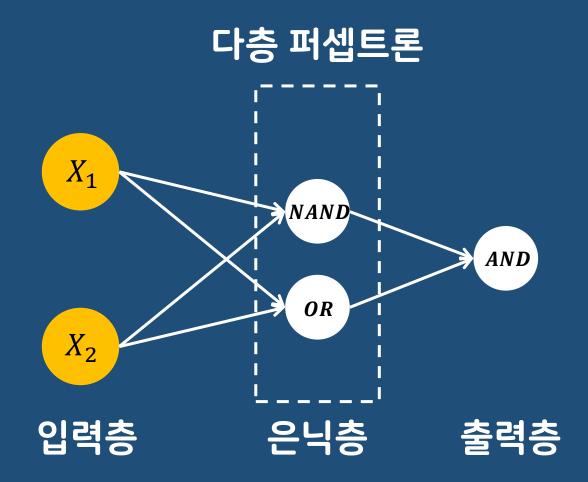
다층 퍼셉트론(Multi Layer Perceptron)

퍼셉트론을 여러 개의 층으로 구성하여 만든 신경망











NAND 게이트

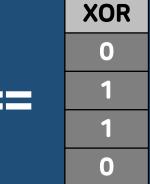
| x1 | x2 | NAND |
|-----------|-----------|------|
| 0 | 0 | 1 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |



| x1 | x2 | NAND |
|-----------|-----------|------|
| 0 | 0 | 1 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

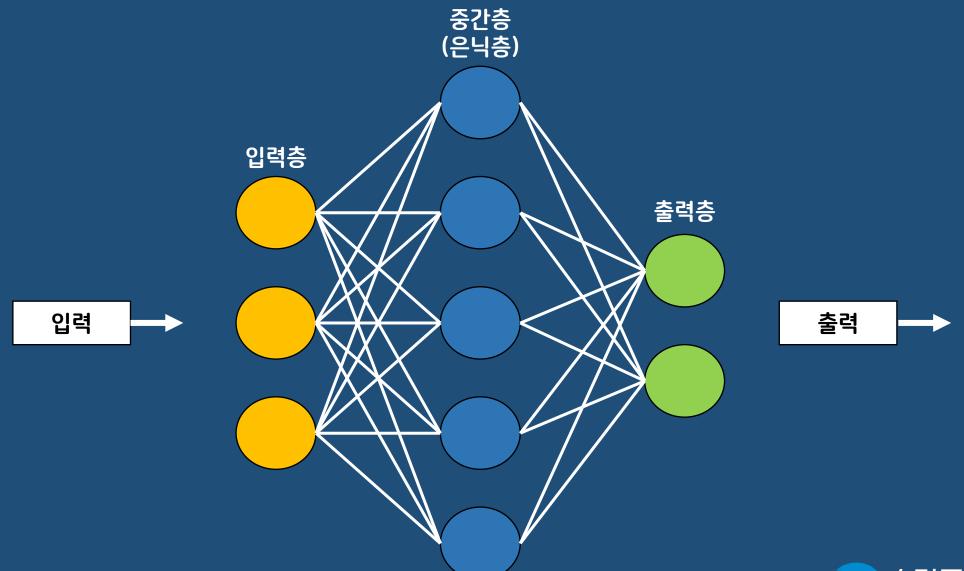
| x1 | x2 | OR |
|-----------|-----------|----|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 1 |

| | NAND | OR | AND |
|---|------|----|-----|
| | 1 | 0 | 0 |
| X | 1 | 1 | 1 |
| | 1 | 1 | 1 |
| | 0 | 1 | 0 |





다층 퍼셉트론(Multilayer Perceptron)





다층 퍼셉트론(Multilayer Perceptron)

- 비선형 데이터를 분리 할 수 있다.
- 단층에 비해 학습시간이 오래 걸린다.
- 가중치 파라미터가 많아 과대적합되기 쉽다.

