

인공지능을 위한 머신러닝 알고리즘

6. 신경망

CONTENTS

1

뇌와 컴퓨터

2

퍼셉트론: 뇌를 모사한 신경망 알고리즘

3

다층 신경망의 분류 원리

학습 목표

- 컴퓨터와 뇌의 동작 방식을 비교할 수 있다.
- 퍼셉트론의 구성 요소를 이해할 수 있다.
- 다층 신경망의 분류 원리를 이해할 수 있다.



1. 뇌와 컴퓨터


■ 뇌와 컴퓨터의 비교



1. 100억 개의 뉴런
2. 60조 개의 시냅스
3. 분산 처리 방식
4. 비선형 연산
5. 병렬 처리



1. 뉴런보다 빠른 연산(10^{-9} 초)
뉴런: 10^{-3} 초
2. 중앙 처리 방식
3. 산술 연산 (선형)
4. 순차 처리



2. 퍼셉트론: 뇌를 모사한 신경망 알고리즘

■ 신경회로망에 대한 연구

1940년대 ~ 1960년대

- ◉ 뇌의 구조에 대한 모방이 두드러진 시기
- ◉ **McCulloch and Pitts**, 뉴런 모델화(1943)
- ◉ **Weiner**, 사이버네틱스(1948)
- ◉ **Rosenblatt** - 퍼셉트론(1957)

생물학적 뉴런의 구조 vs. 인공 뉴런

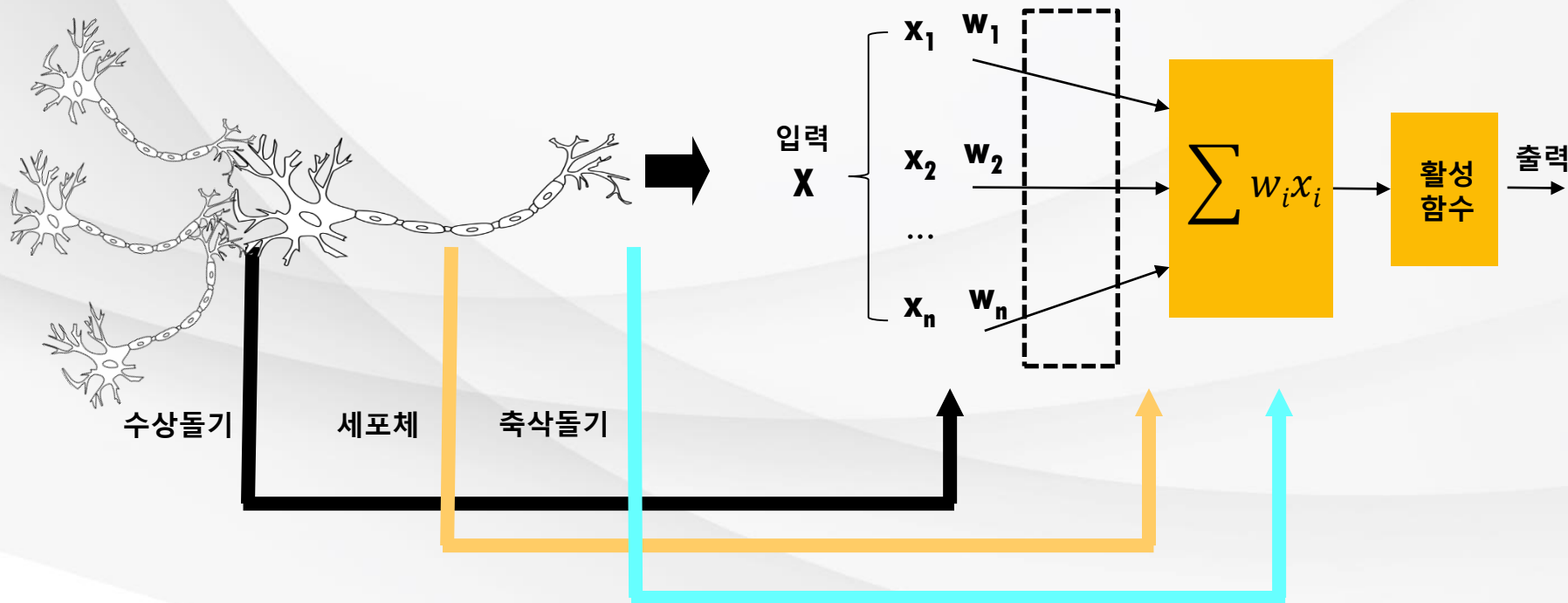
- ◉ 생물학적인 뉴런과의 유사성

병렬 계산(parallel computing)

분산 표현(distributed representation)

2. 퍼셉트론: 뇌를 모사한 신경망 알고리즘

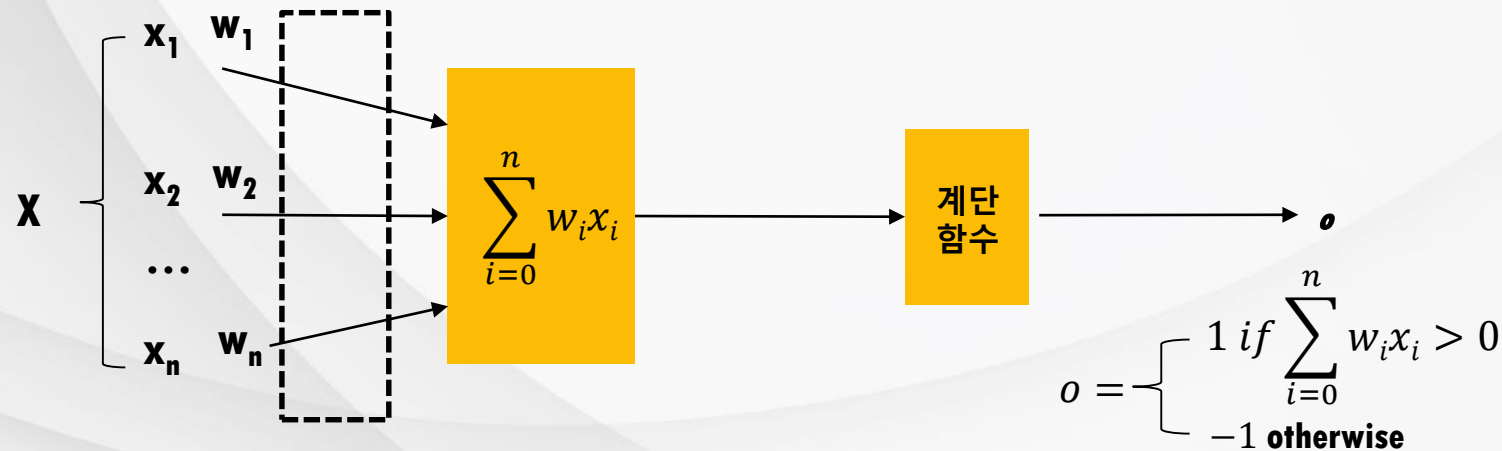
■ 생물학적 뉴런에서 인공 뉴런으로



■ 신경망(인공 뉴런) 학습에 적합한 문제

- ◉ 학습해야 하는 현상이 여러 가지 속성에 의해 표현되는 경우
- ◉ 학습 예제에 에러(**noise**)가 존재할 가능성
- ◉ 긴 학습 시간
- ◉ 학습된 결과를 사람이 이해하는 것이 필요 없는 경우

■ 퍼셉트론 (1957)

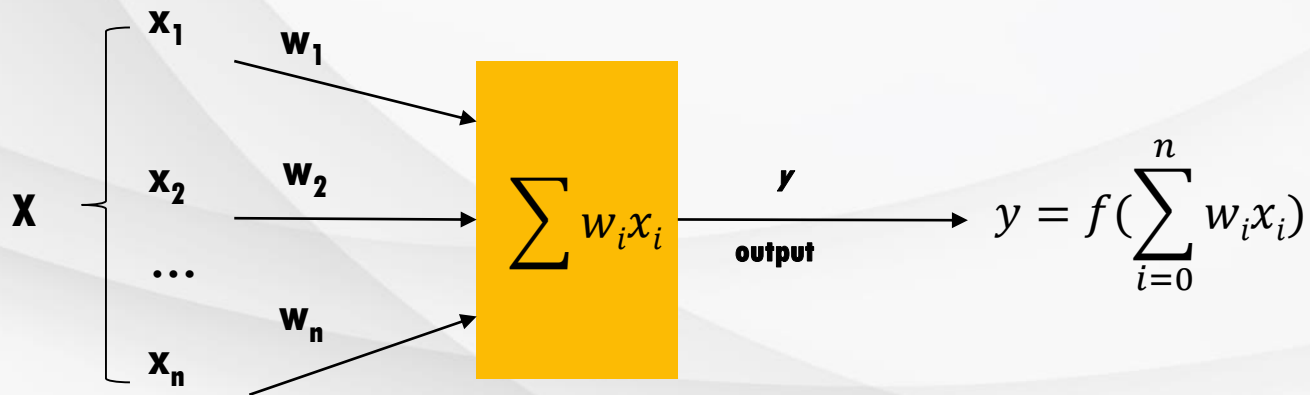


$$o(x_1, \dots, x_n) = \begin{cases} 1 & \text{if } w_0 + w_1 x_1 + \dots + w_n x_n > 0 \\ -1 & \text{otherwise} \end{cases}$$

- 입력: 실수 값을 갖는 벡터
- 모델 구성: 연결 가중치 & 임계치
- 출력: 1 또는 -1
- 학습: 알맞은 연결 가중치를 탐색

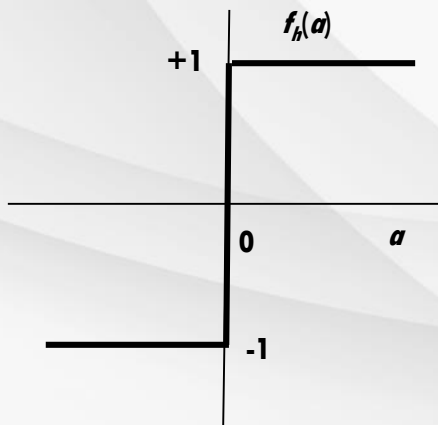
■ 퍼셉트론에서 비선형 함수

기본적인 노드에서의 입출력

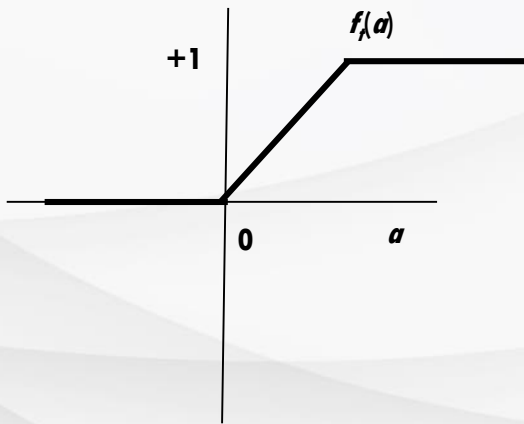


■ 퍼셉트론에서 비선형 함수

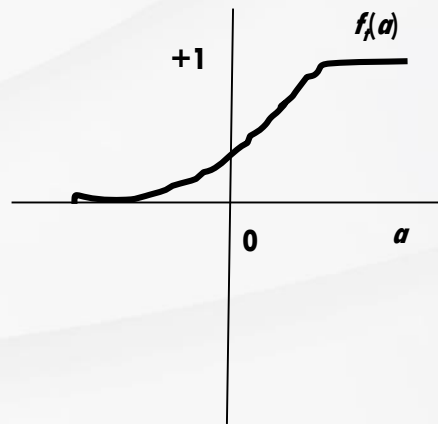
3가지 대표적인 비선형 함수



계단 함수



임계논리 함수



시그모이드 함수

■ 퍼셉트론 학습 룰

$$w_i \longleftarrow w_i + \Delta w_i$$

where $\Delta w_i = \eta(t - o)x_i$

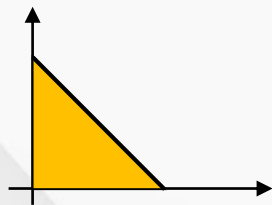
- t 는 타겟 값
- o 는 퍼셉트론의 출력
- η 는 학습률 (작은 상수)

학습 후 올바른 가중치를 찾아내려면 충족되어야 할 사항

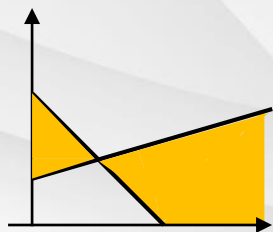
- ◉ 훈련 데이터가 선형 분리 문제이어야 함
- ◉ 충분히 작은 학습률(learning rate)

2. 퍼셉트론: 뇌를 모사한 신경망 알고리즘

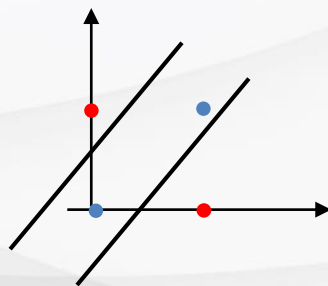
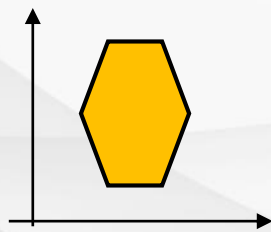
■ 퍼셉트론의 한계



선형 분리 가능 지역



선형 분리 불가능 지역



선형 분리 가능한 일반적인 개념

입력		XOR
0	0	0
0	1	1
1	0	1
1	1	0

XOR 함수

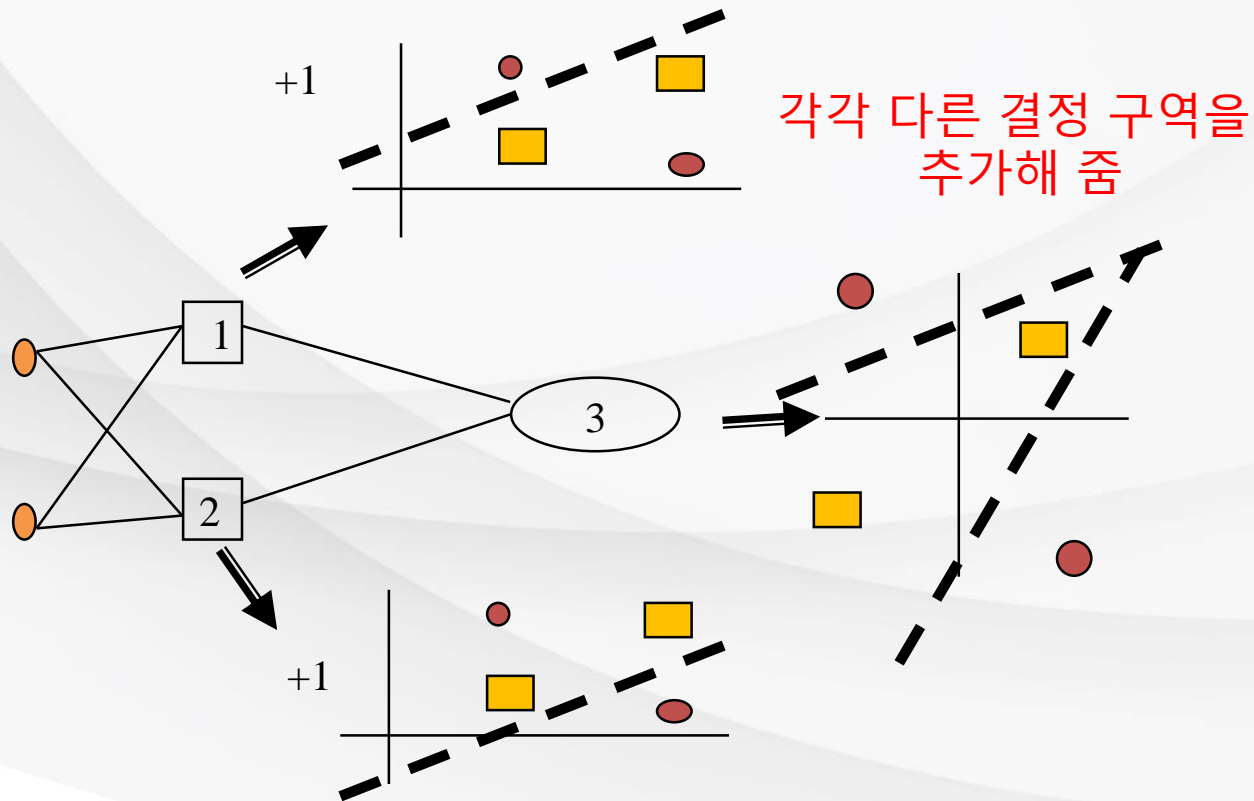
문제점: **XOR** 문제에서 선형 분리 불가능



3. 다층 신경망의 분류 원리


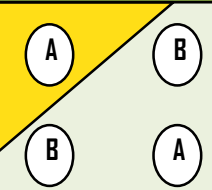
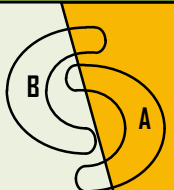
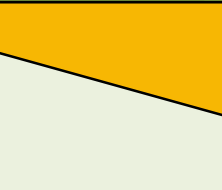
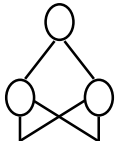
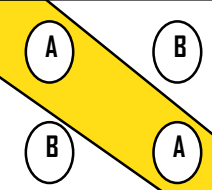
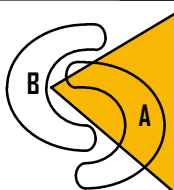
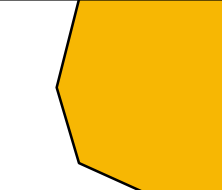
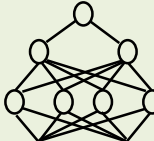
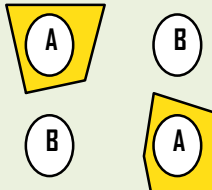

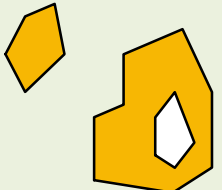
XOR 문제의 해결

모델에 은닉층을 추가하면 분류를 위한 더 좋은 결정 영역을 얻을 수 있음



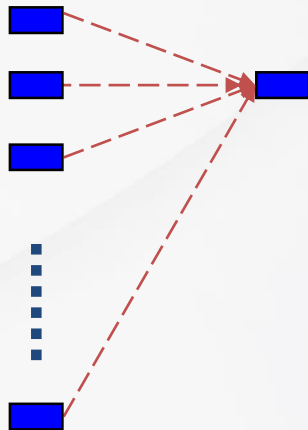
3. 다층 신경망의 분류 원리

■ 여러 패턴들의 결정 형태

구조	결정 영역 형태	Exclusive-OR 문제	얹힌 결정 영역을 갖는 클래스들	대부분의 결정 영역들
단층 	초평면에 의해 나뉘어지는 반평면 (Hyper plane)			
두 개층 	볼록한 모양 또는 닫힌 영역			
세 개층 	임의의 형태 (노드들의 개수에 따라 복잡도가 결정됨)			

■ 다층 퍼셉트론 (다층 신경망)의 특징

- ◉ 같은 층 안에서는 연결이 존재하지 않음
- ◉ 입력층과 출력층 사이 직접적인 연결이 존재하지 않음
- ◉ 각 층 사이는 완전 연결(Fully Connected) 되어 있음
- ◉ 입력층 - 은닉층 - 출력층 구조를 많이 가짐
- ◉ 출력층의 유닛 개수는 입력층의 유닛 개수와 같을 필요가 없음
- ◉ 은닉층의 유닛 개수는 입력층 또는 출력층 보다 많거나 적어도 됨



*각 유닛은 퍼셉트론

$$y_i = f\left(\sum_{j=1}^m w_{ij}x_j + b_i\right)$$

Bias를 추가적인 가중치로 갖기도 함

■ 다중 퍼셉트론 가중치 학습

- ◉ 오류 역전파 알고리즘(**Error Back Propagation**) 이용함
- ◉ **Generalized Delta Rule** 이용

$$O_{ip} = \sum_i w_{ij} \times i_{ip}$$

$$E_p = \frac{1}{2} \sum_j (t_{jp} - o_{jp})^2$$

$$E = \sum_p E_p$$

$$\frac{\delta E_p}{\delta w_{ij}} = \frac{\delta E_p}{\delta o_{jp}} \times \frac{\delta o_{jp}}{\delta w_{ij}}$$

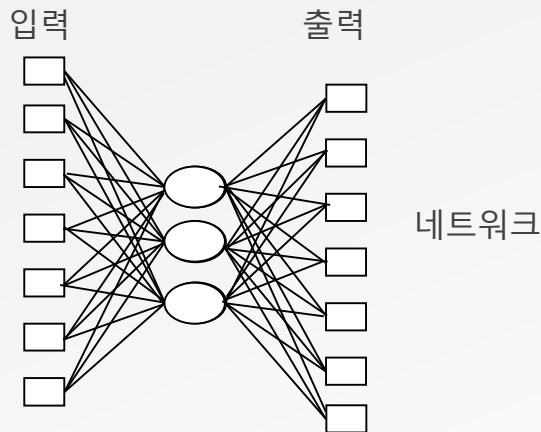
$$\frac{\delta E_p}{\delta w_{ij}} = -2 \frac{(t_{jp} - o_{jp})}{2} \times i_{ip} = -(t_{jp} - o_{jp}) \times i_{ip}$$

$$w_{ij} = -\alpha \frac{\delta E_p}{\delta w_{ij}} = -\alpha (t_{jp} - o_{jp}) \times i_{ip}$$

■ 은닉 유닛의 표현

- ◉ 입력값 들의 특성을 스스로 파악해서 **Hidden Layer**에 표현하는 능력이 있음
- ◉ 사람이 미리 정해 준 **Feature**만을 사용하는 경우보다 유연하며 미리 알 수 없는 특성을 파악하는데 유용

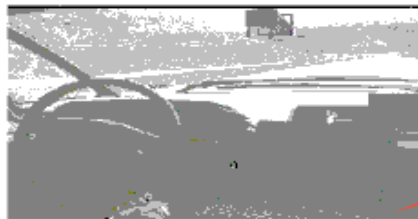
입력 - 학습된 은닉 유닛 - 출력 유닛의 예



입력	은닉값			출력
10000000	→ .89	.04	.02	→ 10000000
01000000	→ .48	.38	.25	→ 01000000
00100000	→ .69	.68	.82	→ 00100000
00010000	→ .84	.14	.57	→ 00010000
00001000	→ .70	.79	.12	→ 00001000
00000100	→ .25	.45	.09	→ 00000100
00000010	→ .07	.58	.96	→ 00000010
00000001	→ .84	.83	.74	→ 00000001

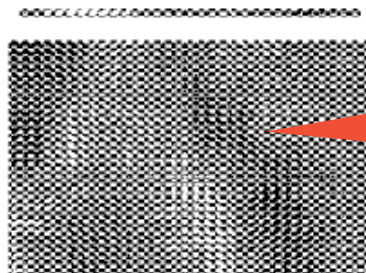
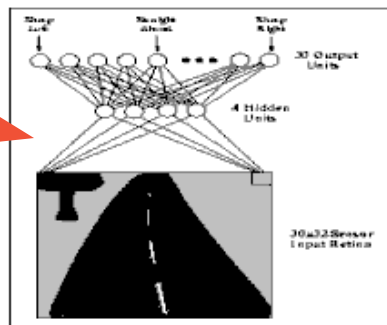
■ 응용 예제: Autonomous Land Vehicle (ALV)

- ◉ 신경망이 무인 자동차의 운전대를 움직이는 방법을 학습
- ◉ **960**개의 입력 유닛, **4**개의 은닉 유닛, **30**개의 출력 유닛
- ◉ **70 miles/h**의 속도로 주행



CMU의
ALVINN 시스템

차량 앞에
장착된
카메라로부터
받은 이미지



한 개 은닉 유닛의
가중치 값



학습정리

지금까지 [신경망]에 대해서 살펴보았습니다.

뇌와 컴퓨터

분산 처리(뇌) <-> 중앙 처리(컴퓨터)
병렬 처리(뇌) <-> 순차 처리(컴퓨터)
처리속도: 10^{-3} 초 (뇌) <-> 10^{-9} 초(컴퓨터)

퍼셉트론: 뇌를 모사한 신경망 알고리즘

퍼셉트론은 입력 유닛들의 가중치의 합과 임계치를 사용하여 뉴런의 발화 과정을 모사
선형 분리 문제를 해결하지 못하는 문제 (예> XOR)

다층 신경망의 분류 원리

다층 신경망은 입력층 - 은닉층 - 출력층 구조를 가지며 각 층의 유닛들은 퍼셉트론임
레이어를 쌓을수록 다양한 결정 영역이 추가로 생김으로써 선형 분리 문제 극복