

Machine Learning

9강

# 신경망 (1)

컴퓨터과학과 이관용 교수

# 학습목차

01 신경망 개요

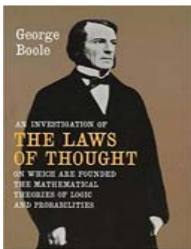
02 다층 퍼셉트론

1

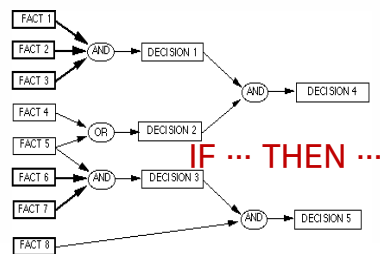
# 신경망 개요

## 인공지능 접근 방법

부울 논리

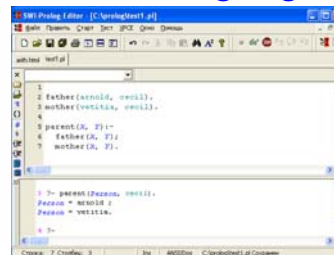


규칙 기반 지식 표현

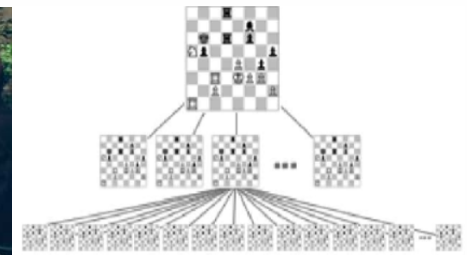


### SYMBOLIC AI

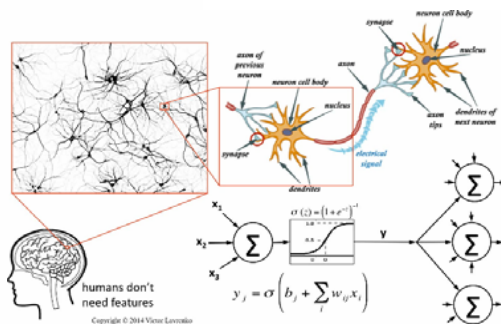
PROLOG: Language for AI



IBM Deep Blue, Chess AI

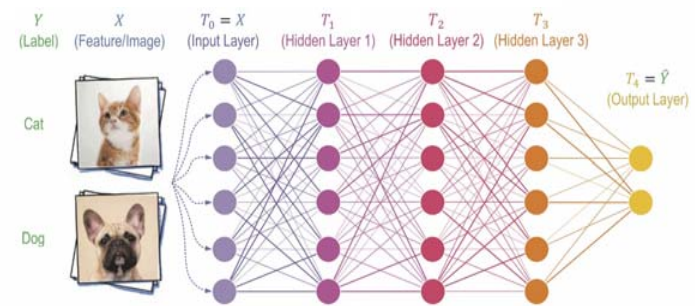
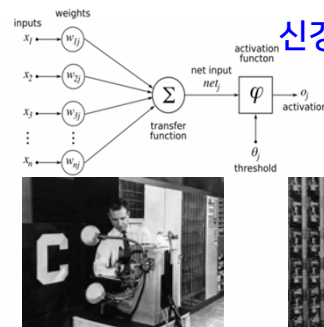


# ARTIFICIAL INTELLIGENCE



뇌에서 영감을 받은 계산 모형

신경망: 퍼셉트론



딥러닝

## CONNECTIONIST AI : Artificial Neural Networks

# 신경망과 딥러닝

## ○ 신경망 신경회로망, neural networks

- 생물학적 신경회로망을 모델링한 수학적 함수
- 원하는 입출력 매핑 함수의 형태를 스스로 찾는 학습 능력을 가짐
- 데이터를 이용하여 학습이 수행되므로 데이터 분석 툴로 사용
- 학습 방식(데이터 분석 용도)에 따라 다양한 모델이 존재

## ○ 심층 신경망 deep networks

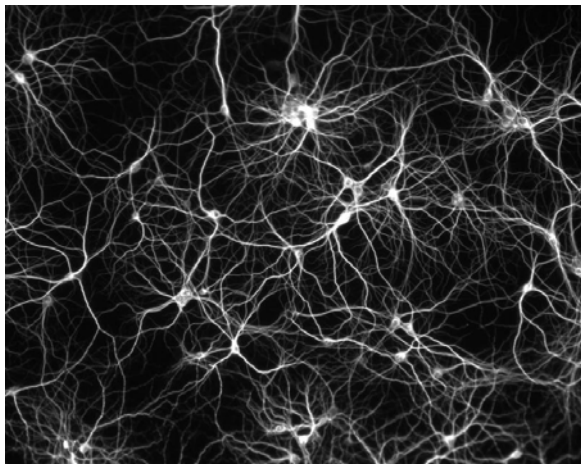
- 가장 발전된 형태의 신경망 모델들

## ○ 딥러닝 deep learning

- 심층 신경망을 이용하여 데이터를 분석하는 머신러닝 기술

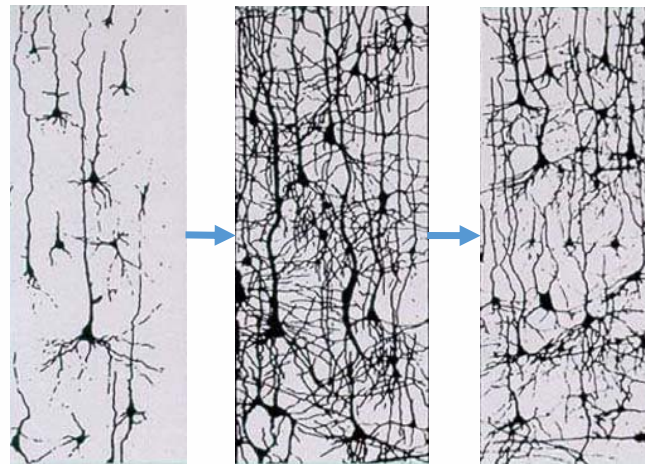
# 신경망?

- 인간 뇌의 구조와 뇌에서 수행되는 정보처리 방식을 모방함으로써 인간이 지능적으로 처리하는 복잡한 정보처리 능력을 기계를 통해 실현하고자 하는 연구



- **신경세포** → 100억 개 이상
- **세포간 연결** → 60조 이상

<https://images.app.goo.gl/yQTpAR8hRbJrwd3b7>



출생 직후

6년

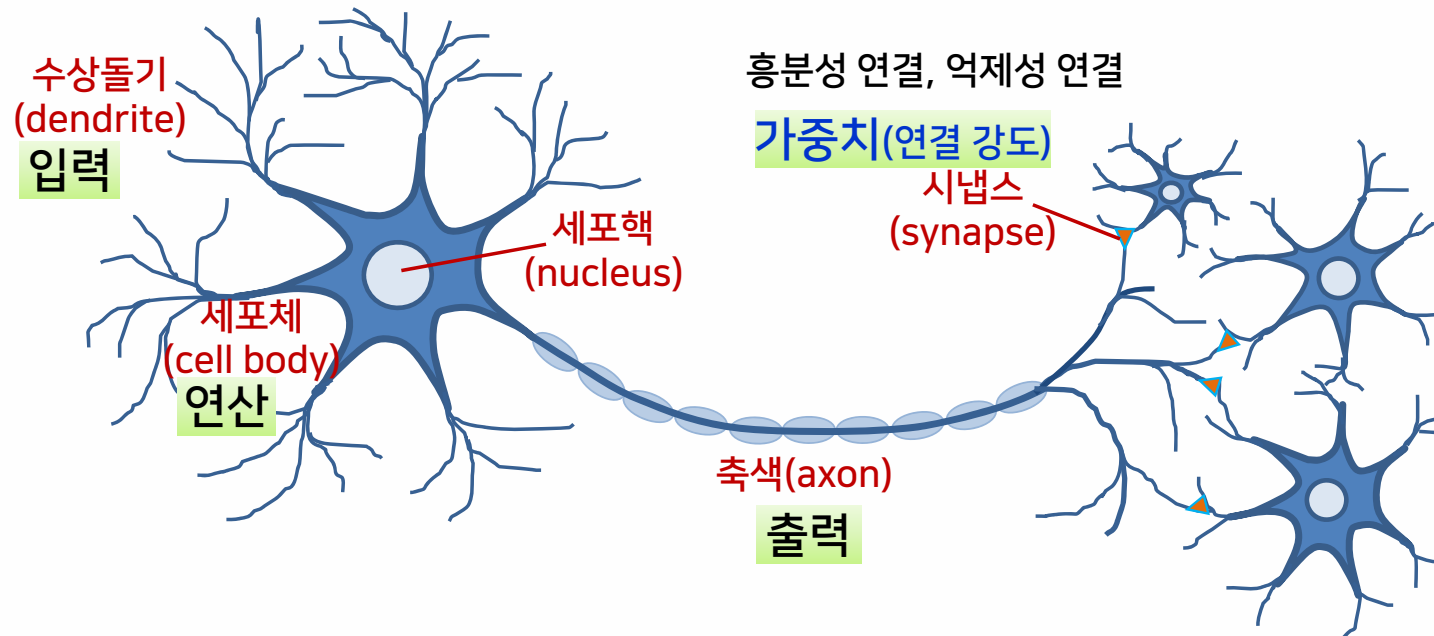
14년

Rethinking the Brain: New Insights into Early Development  
by Rima Shore (NY: Families and Work Institute, 1997)

- ✓ **적응성** adaptation
- ✓ **학습** learning

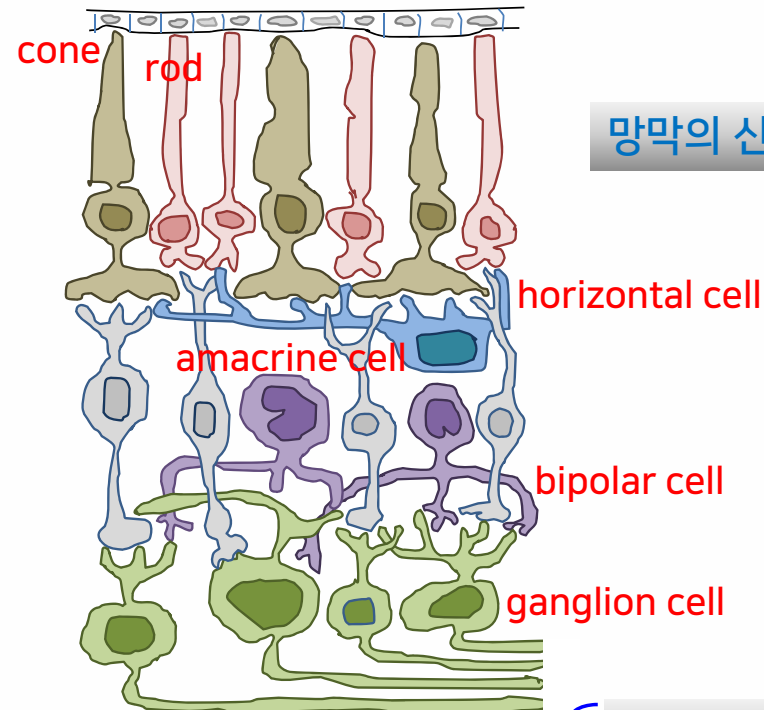
# 생물학적 신경망

## 신경세포 neuron의 구조와 연결



# 생물학적 신경망

신경세포의 연결 구조



망막의 신경세포들의 계층 구조

○ 인공 신경망 artificial neural networks

□ 인간 뇌의 정보처리 방식을 **모델링**하는 방법

1. 신경세포
2. 신경망의 구조
3. 학습 메커니즘



## 생물학적 신경망에서 인공 신경망으로

### ○ 인공신경망의 구성 요소

#### □ 신경세포 neuron, node, unit

- ✓ 하나의 신경세포가 수행하는 기능을 수학적 함수로 정의

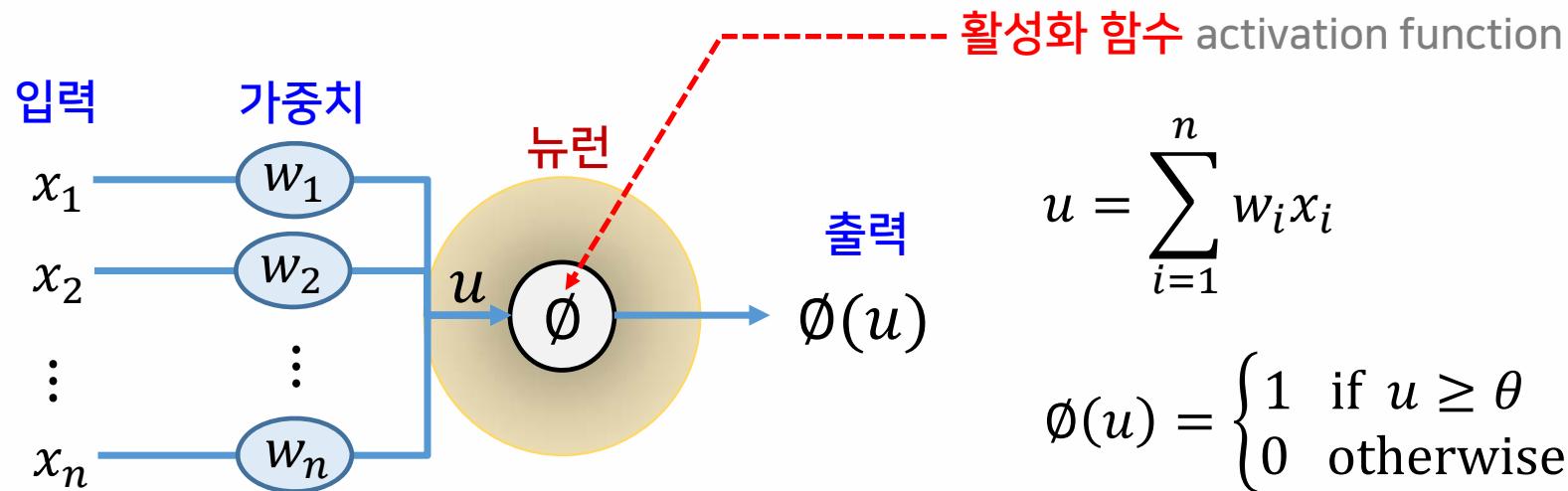
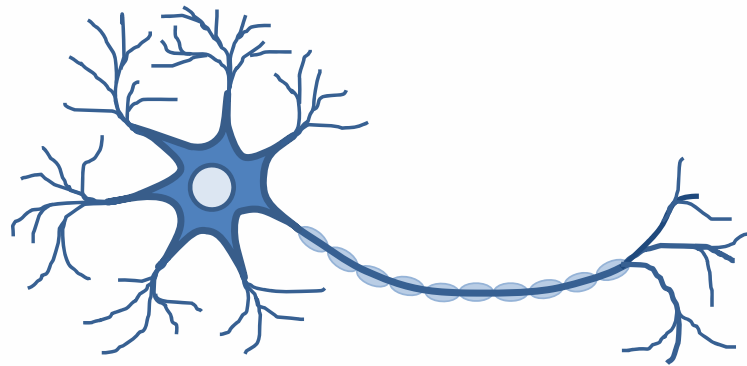
#### □ 신경망 구조 network structure

- ✓ 신경세포들이 서로 정보를 전달하는 연결 구조

#### □ 학습 알고리즘 learning algorithm

- ✓ 신경망이 원하는 기능을 수행할 수 있도록  
신경세포들 간의 연결 강도를 조정하는 방법

# 신경망의 구성 요소: **인공 신경세포**

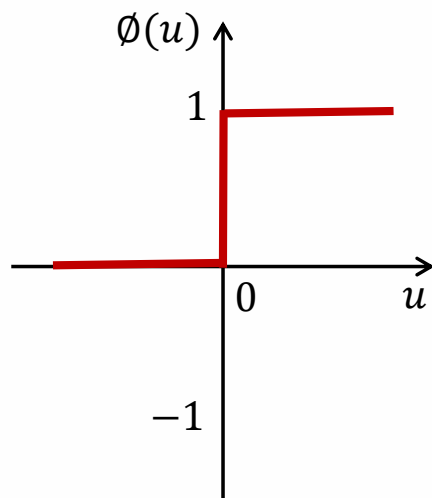


# 신경망의 구성 요소: **인공 신경세포**

## ○ 활성화 함수

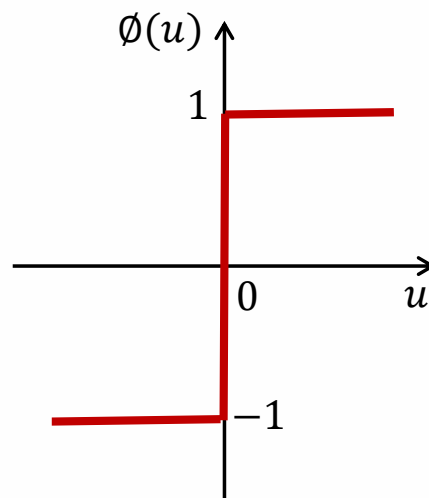
□ 뉴런의 핵심 → 하나의 뉴런의 특성을 결정하는 역할

계단함수



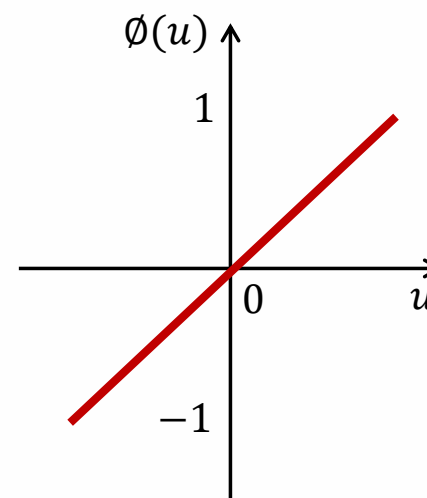
$$\phi_{step}(u) = \begin{cases} 1 & \text{if } u \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

부호함수



$$\phi_{sign}(u) = \begin{cases} 1 & \text{if } u \geq 0 \\ -1 & \text{otherwise} \end{cases}$$

선형함수

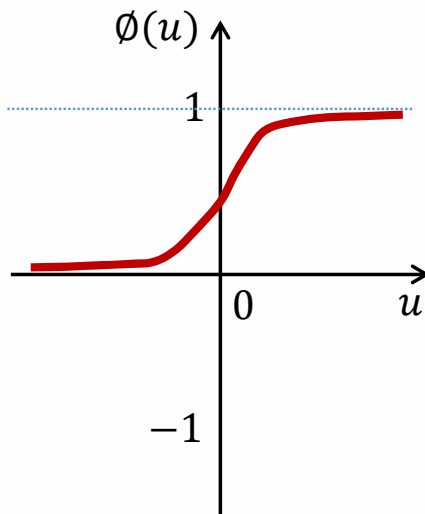


$$\phi_{linear}(u) = u$$

# 신경망의 구성 요소: **인공 신경세포**

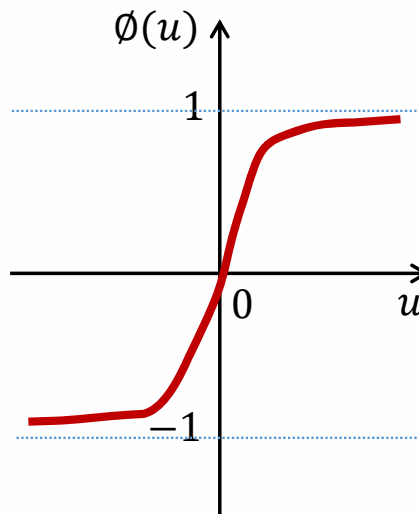
## ○ 활성화 함수

시그모이드 함수



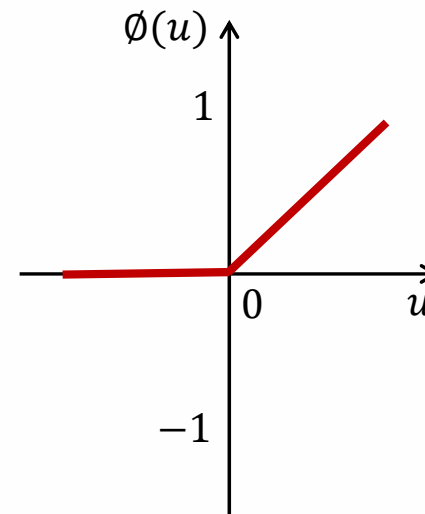
$$\phi_{\text{sigmoid}}(u) = \frac{1}{1 + e^{-u}}$$

하이퍼탄젠트 함수



$$\phi_{\text{tanh}}(u) = \frac{1 - e^{-2u}}{1 + e^{-2u}}$$

ReLU 함수

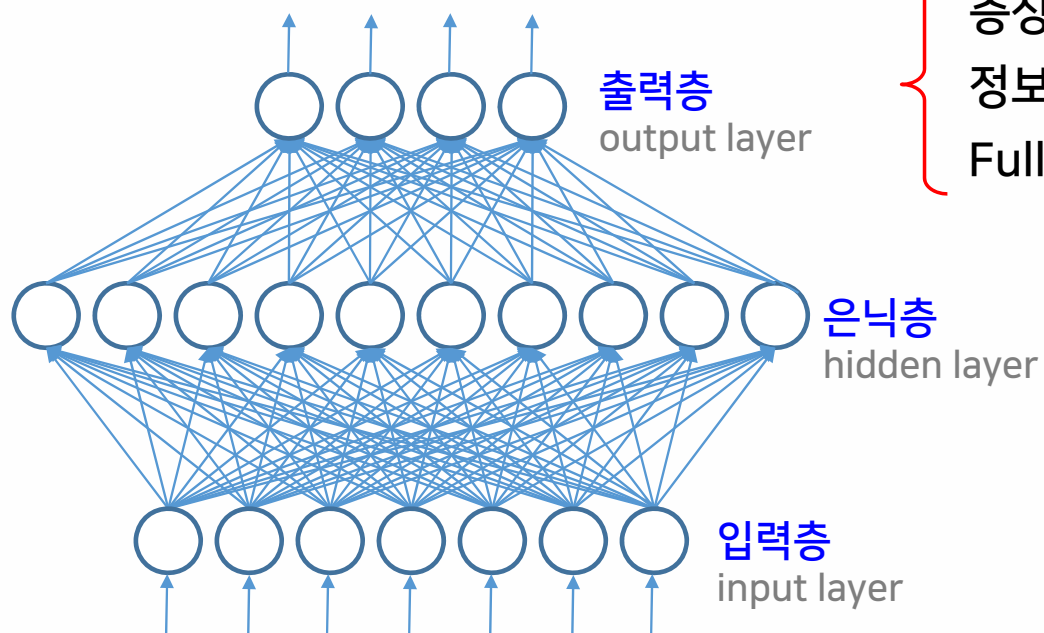


$$\phi_{\text{relu}}(u) = \max(0, u)$$

# 신경망의 구성 요소: 연결 구조

## ○ 신경세포들의 대표적 연결 방식

### □ 다층 전방향 신경망 multi-layer feed forward neural network



층상 구조

정보의 흐름: 한쪽 방향(입력층→출력층)

Fully connected network, dense network

#### 은닉층의 존재 여부

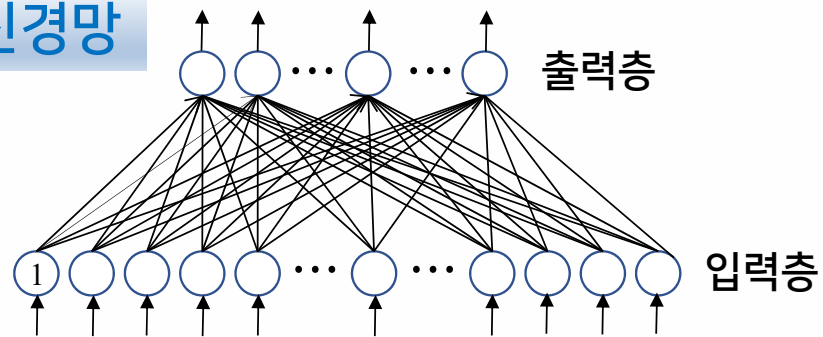
- 단층 single layer 신경망
- 다층 multilayer 신경망 → **심층 신경망**

#### 정보 흐름의 방향

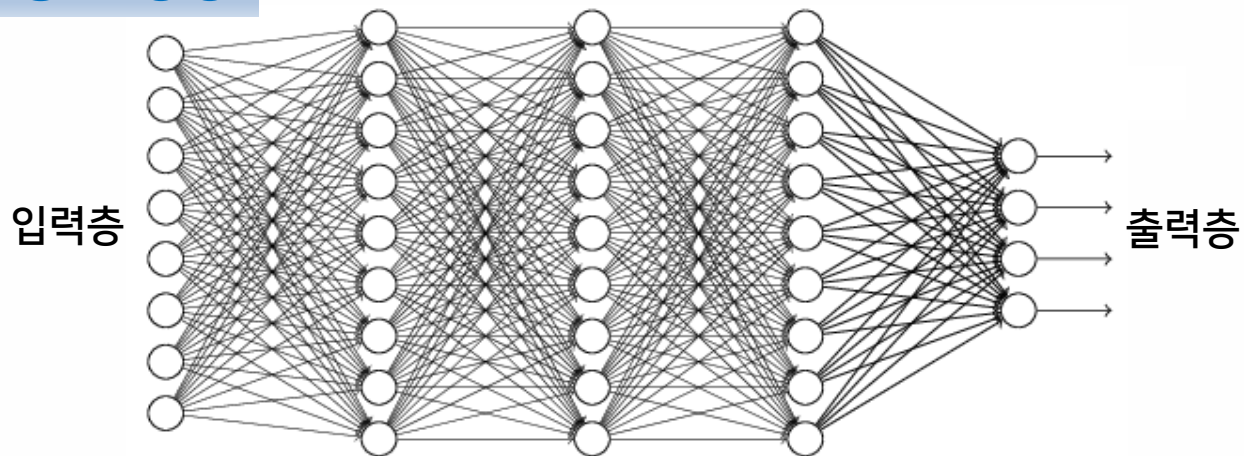
- 전방향 feed-forward 신경망
- 회귀 recurrent 신경망 (RNN)

## 다양한 연결 구조 - 층수의 변화

### 단층 신경망

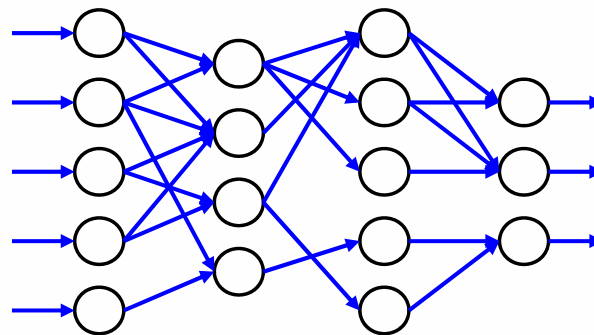


### 심층 신경망

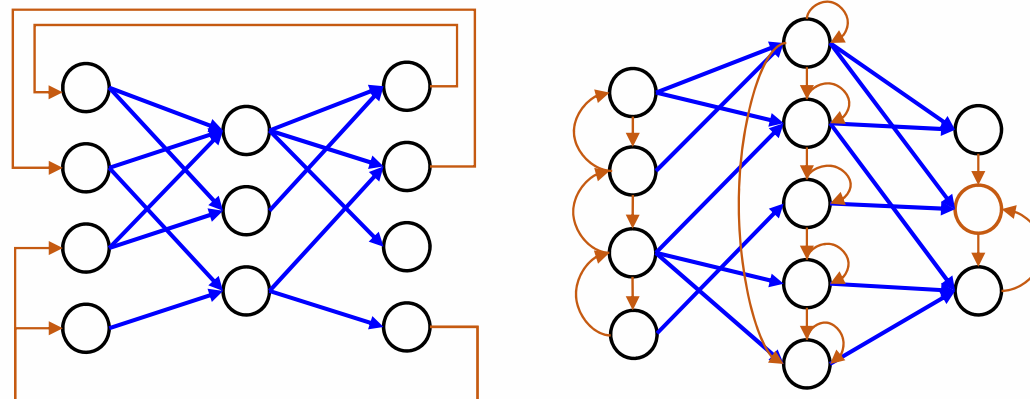


## 다양한 연결 구조 - 정보 흐름의 방향

### 전방향 신경망



### 회귀 신경망

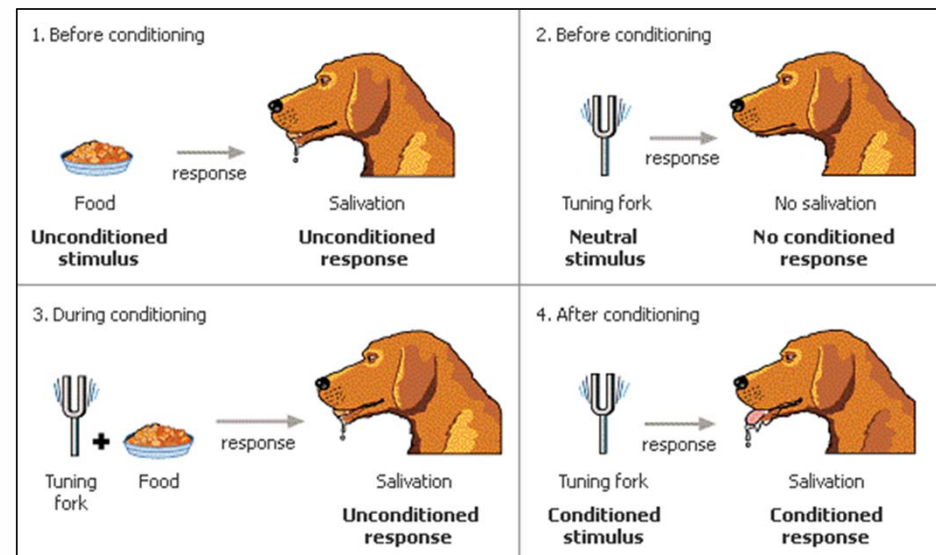


# 신경망의 구성 요소: 학습

## ○ 인간 뇌의 학습

- 성장하면서 뇌 세포들 간의 연결이 형성되어  
여러가지 기능을 수행하게 되는 과정
- 세포들 간의 연결 형성 규칙
  - ✓ 연결된 두 신경세포가 동시에  
활성화되면 연결 강도는 강해짐

## 파블로프의 개 실험



<http://www.age-of-the-sage.org/psychology/pavlov.html>



## 신경망의 구성 요소: 학습

### ○ 인공신경망의 학습이란?

- 신경망이 원하는 기능을 수행할 수 있도록 만드는 것
- 신경망에 어떤 입력  $x$ 가 주어졌을 때, 최종적을 내는 출력  $y$ 가 원하는 값이 되도록 가중치  $w$ 를 조정하는 것

#### □ 가중치 조정식

$$w^{(\tau+1)} = w^{(\tau)} + \Delta w^{(\tau)}$$

가중치 변화량

학습 후 가중치 ← 현재 가중치

#### □ 가중치 변화량을 결정하는 방법

✓ 학습 데이터 사용

- 반복적인 가중치 수정을 통해 점점 원하는 기능에 근접해 감

## 학습의 종류

### ○ 지도학습 supervised learning

- ☐ 입력값  $x$ 에 대한 목표 출력값  $t$ 이 함께 주어짐
- ☐ 주어진 입력  $x$ 에 대한 신경망의 출력값  $y$ 가 원하는 목표값  $t$ 에 가까워지도록 가중치  $w$ 를 수정
- ☐ 오류 역전파 학습 error backpropagation learning 알고리즘

### ○ 비지도학습 unsupervised learning

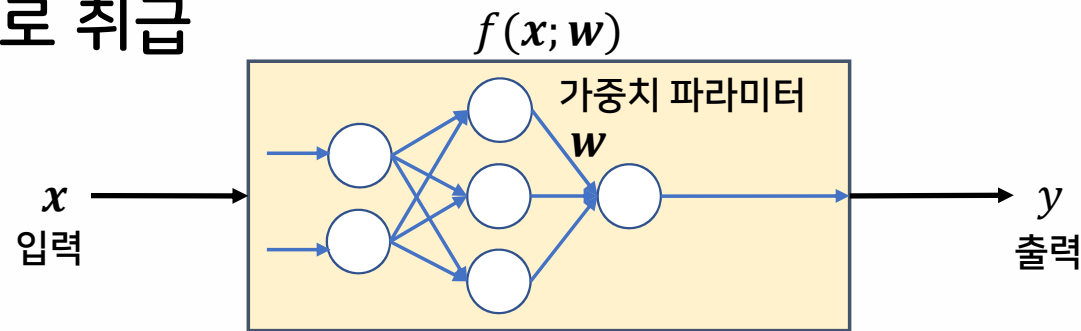
- ☐ 입력값  $x$ 만 주어짐 → 비슷한 입력에 대해 비슷한 출력을 내도록 학습
- ☐ self-organizing feature map, Boltzmann machine

### ○ 강화학습 reinforcement learning

- ☐ 입력  $x$ 에 대한 신경망의 출력값  $y$ 의 보상이 최대가 되도록 가중치를 수정

## 응용 관점에서의 신경망에 대한 이해

### ○ 하나의 함수로 취급

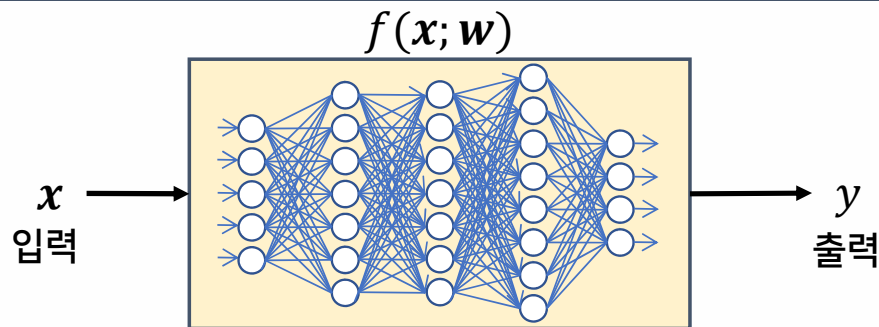


- ☐ 기능 → 입력  $x$ 를 받아 출력  $y$ 를 계산
- ☐ 수학적 정의 →  $y = f(x)$

### ○ 함수 $f$ 를 결정하는 요소

- ☐ 신경세포의 활성화 함수와 연결 구조 → 고정된 형태
- ☐ 연결 가중치 → 학습을 통해 함수  $f$ 의 형태를 조정
  - ✓ 신경망의 학습 → 원하는 함수  $f$ 를 찾는 과정

# 왜 신경망인가?



## ○ 표현 representation 능력

- ☐ 신경망은 어떤 형태의 함수도 표현할 수 있음

## ○ 학습 learning 능력

- ☐ 데이터에 대한 학습을 통해 최적의 함수를 찾을 수 있음

## ○ 일반화 generalization 능력

- ☐ 데이터에 대한 단순한 암기가 아닌 데이터에서 일반화된 규칙을 찾음  
→ 새로운 데이터에 대해서도 처리 가능

2

## 다층 퍼셉트론

## M-P 뉴런

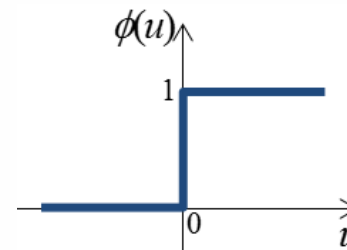
## ○ 1943. McCulloch &amp; Pitts

## □ 단일 신경세포에 대한 첫 번째 모델

$$y_j = \phi_{step} \left( \sum_{i=1}^n w_{ij} x_i + w_{0j} \right)$$

$\xrightarrow{\text{바이어스 bias}}$   $w_{0j}$   
 $\xrightarrow{u}$   $u$

$$\phi_{step}(u) = \begin{cases} 1 & \text{if } u \geq 0 \\ 0 & \text{otherwise} \end{cases}$$



# 퍼셉트론 Perceptron

## ○ 1958. Rosenblatt

- M-P 뉴런을 여러 개 결합하여 네트워크 형태를 갖춘 신경망
- ✓ 패턴인식을 수행하는 최초의 신경망 → 단층 전방향 신경망



## ○ 뉴런

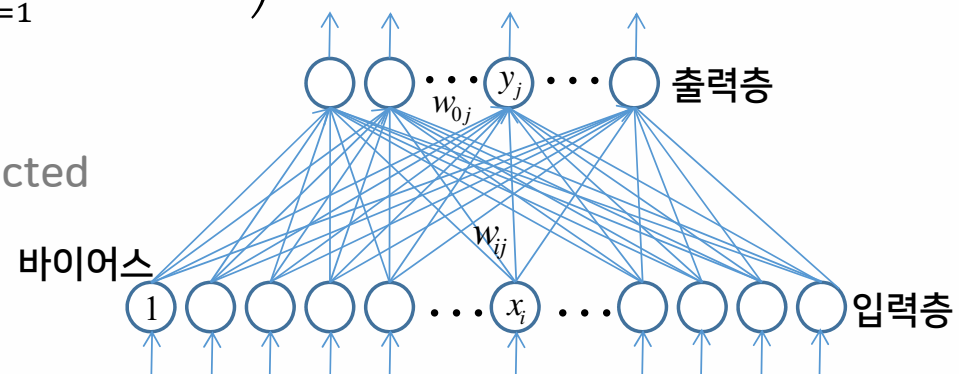
- M-P 뉴런 → 계단함수  $y_j = \phi_{step} \left( \sum_{i=1}^n w_{ij} x_i + w_{0j} \right)$

## ○ 연결 구조

- 단층, 전방향, 완전 연결 fully-connected

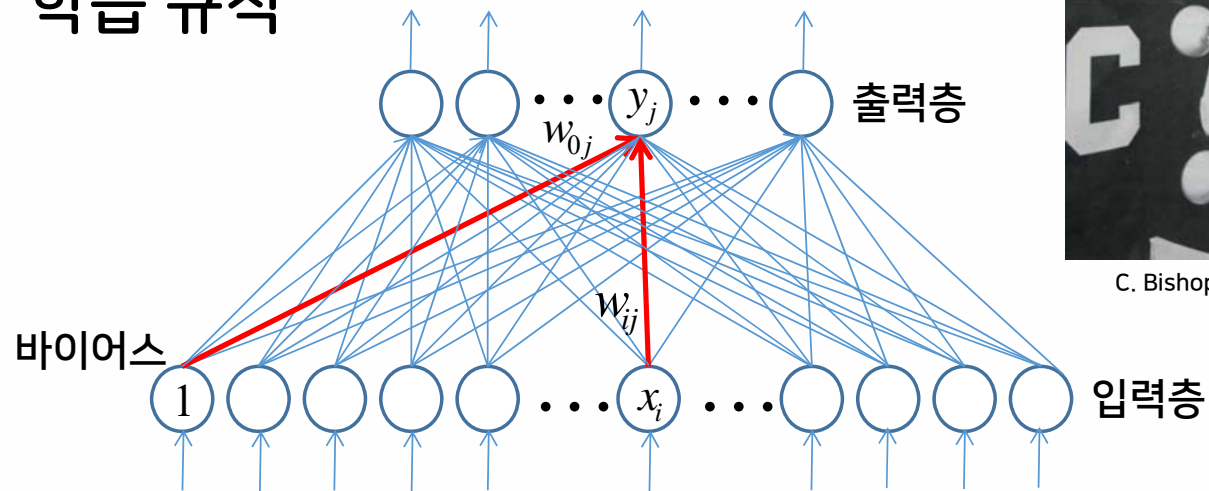
## ○ 학습 규칙

- 이진 입출력을 사용한 지도학습



# 퍼셉트론

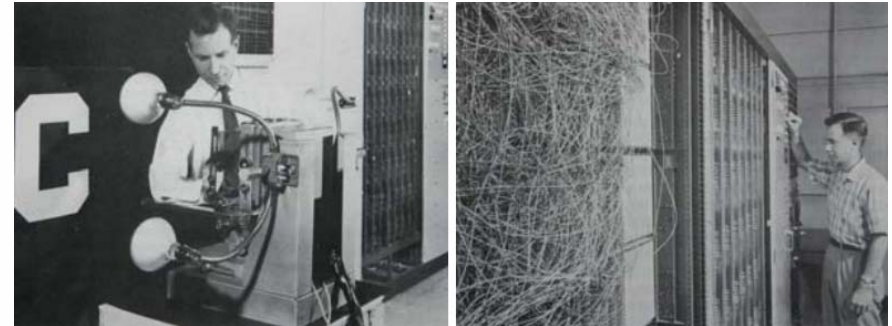
## ○ 학습 규칙



$$w_{ij}^{(\tau+1)} = w_{ij}^{(\tau)} + \boxed{\eta}(t_j - y_j)x_i$$

$\eta$ : 학습률 (learning rate)  
 $t_j$ : 목표 출력 (target output)  
 $y_j$ : 실제 출력 (actual output)

Mark 1 perceptron 하드웨어



C. Bishop, Pattern Recognition and Machine Learning. Springer (2006)



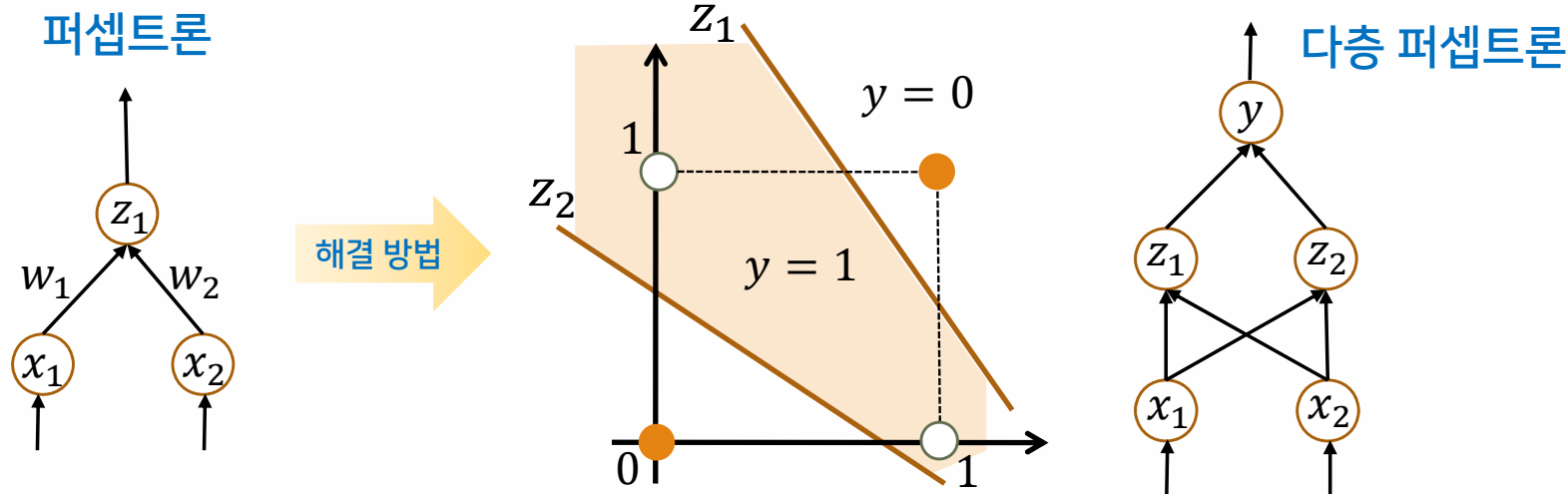
# 퍼셉트론의 한계

## ○ 선형 판별함수

→ 비선형 결정경계를 표현할 수 없음

## ○ XOR 문제

□ Minsky & Papert → XOR 문제의 해결 불가능을 지적



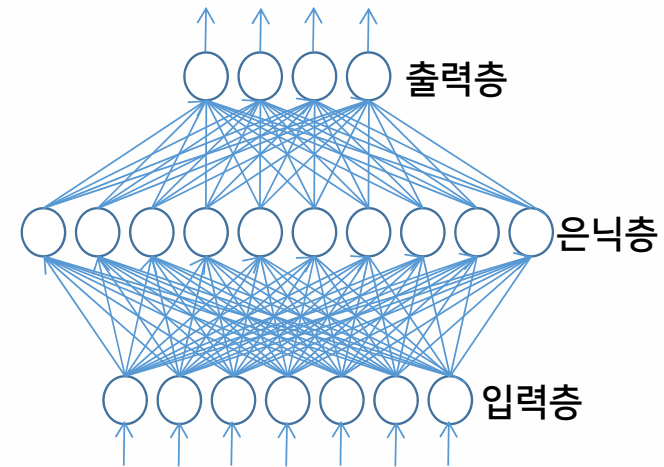
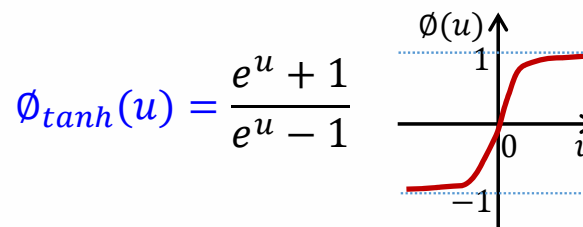
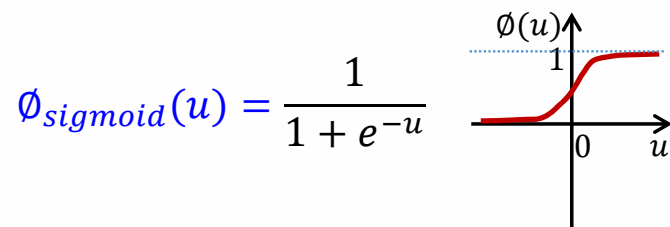
# 다층 퍼셉트론

## ○ MLP, Multi-Layer Perceptron

- 1개 이상의 은닉층을 가짐

## ○ 뉴런

- 출력은 입력에 대한 비선형 매핑

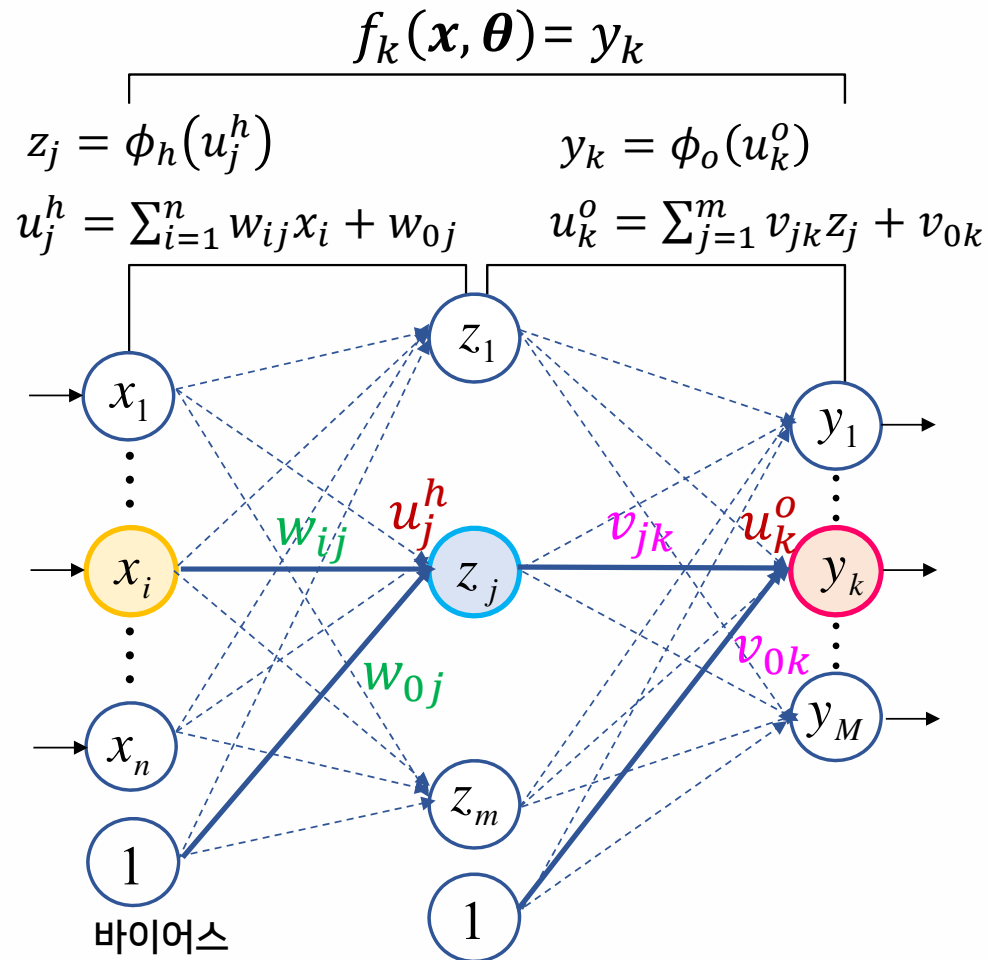


## ○ 연결 구조 → 다층, 전방향, 완전연결

## ○ 학습 알고리즘

- 오류 역전파 error backpropagation 알고리즘 → 지도학습

# 다층 퍼셉트론의 입출력 관계

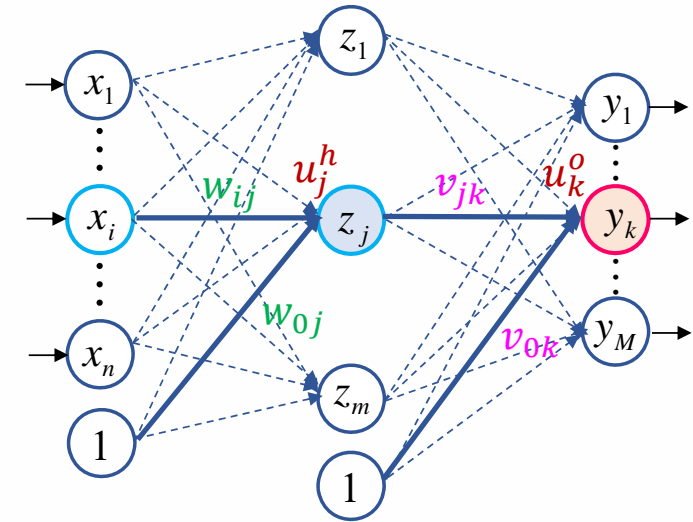


# 다층 퍼셉트론의 입출력 관계

$k$ 번째 출력 노드의 출력

$$\begin{aligned}
 y_k &= f_k(\mathbf{x}, \boldsymbol{\theta}) = \phi_o \left( \sum_{j=1}^m v_{jk} z_j + v_{0k} \right) \\
 &= \phi_o \left( \sum_{j=1}^m v_{jk} \phi_h \left( \sum_{i=1}^n w_{ij} x_i + w_{0j} \right) + v_{0k} \right)
 \end{aligned}$$

$u_k^o$   
 $u_j^h$



$k$ 번째 출력 노드의 출력

$$y_k = \phi_o(u_k^o)$$

$$u_k^o = \sum_{j=1}^m v_{jk} z_j + v_{0k}$$

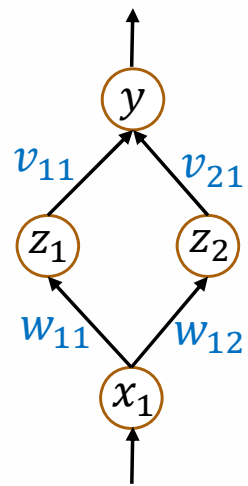
$j$ 번째 은닉 노드의 출력

$$z_j = \phi_h(u_j^h)$$

$$u_j^h = \sum_{i=1}^n w_{ij} x_i + w_{0j}$$

하나의 은닉층을 가진 MLP는 임의의 정확도로 모든 연속 함수의 근사가 가능

# 다층 퍼셉트론의 표현 능력



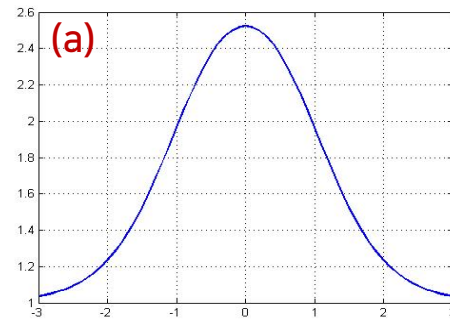
출력 뉴런  $\rightarrow \phi_{linear}(u)$

은닉 뉴런  $\rightarrow \phi_{tanh}(u)$

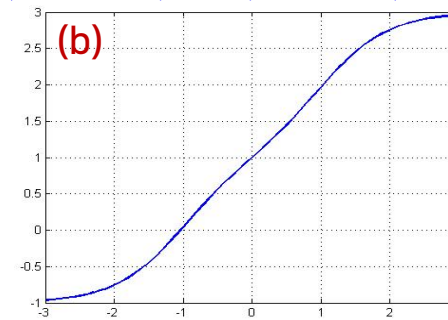
$$y = v_{11}z_1 + v_{21}z_2 + v_{01} = v_{11}\tanh(w_{11}x_1 + w_{01}) + v_{21}\tanh(w_{12}x_1 + w_{02}) + v_{01}$$

가중치의 변화에 따른 다양한 함수의 형태

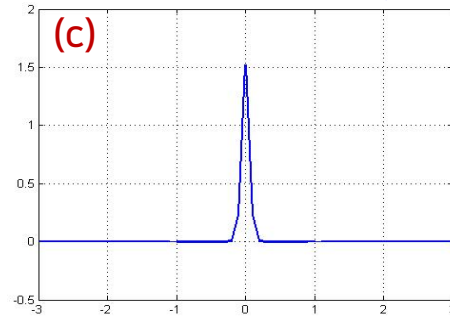
$$f(x) = \tanh(x + 1) - \tanh(x - 1) + 1$$



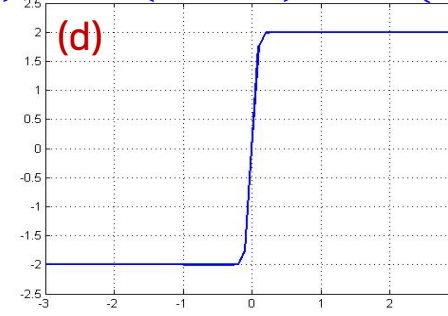
$$f(x) = \tanh(x + 1) + \tanh(x - 1) + 1$$



$$f(x) = \tanh(x + 1) + \tanh(-20x + 1)$$



$$f(x) = \tanh(20x + 1) + \tanh(20x - 1)$$



# 다층 퍼셉트론의 표현 능력

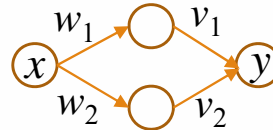
## ○ 입력 1개 - 출력 1개

하나의 은닉층을 가진 MLP는  
임의의 정확도로  
모든 연속 함수의 근사가 가능

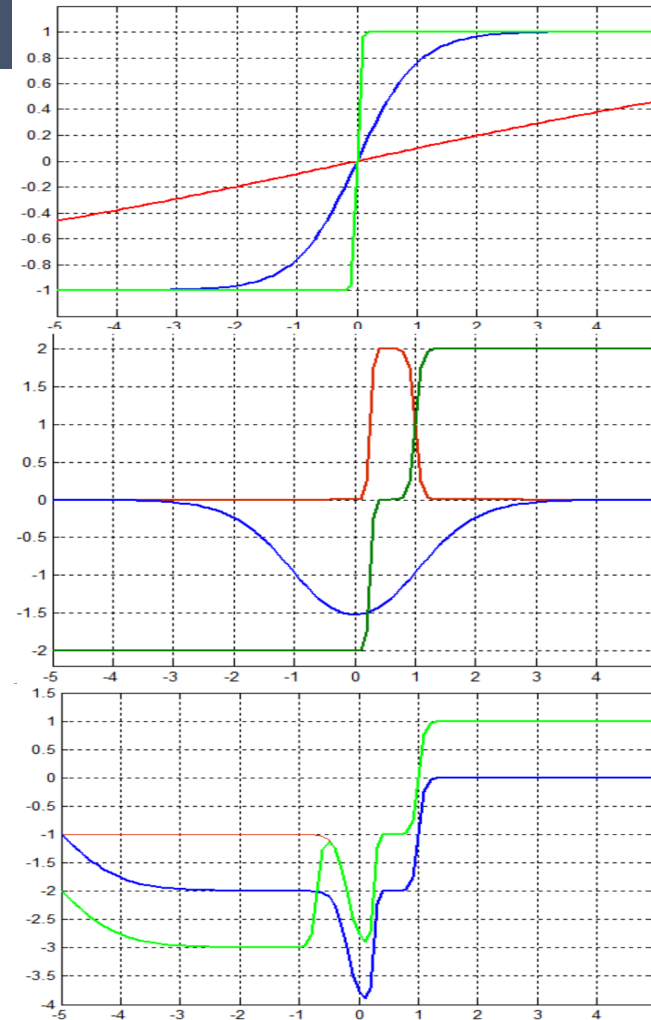
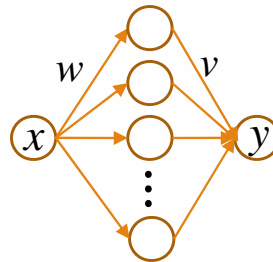
0개 은닉 뉴런

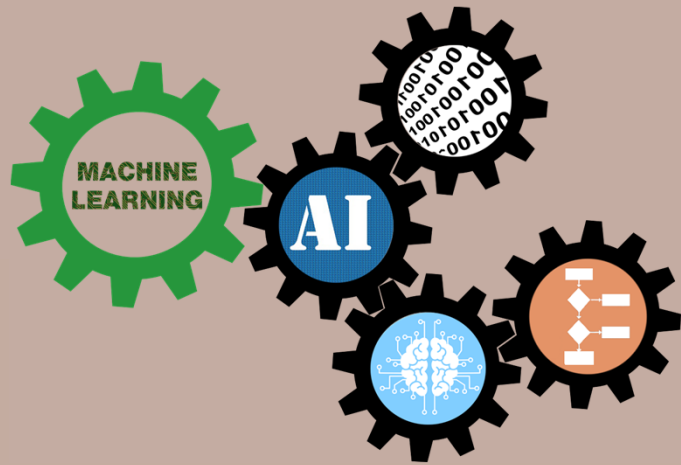


2개 은닉 뉴런



많은 은닉 뉴런





다음시간안내

제10강

## 신경망 (2)