

12강

인공지능

인공 신경망(1)

컴퓨터과학과 이병래교수

학습목차

1 인공지능의 개념

2 퍼셉트론 학습





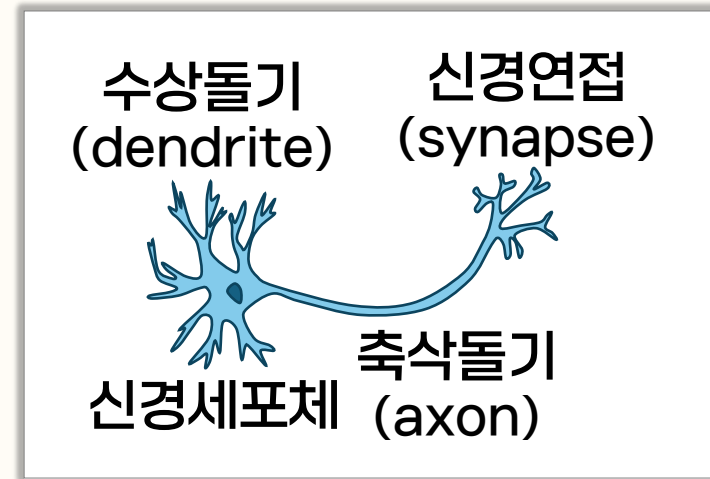
인공 신경망의 개념

1. 인공 신경망의 기본 구조

■ 인공 신경망이란?

• 두뇌 속의 신경 구조

- 약 10억~100억 개의 신경세포(뉴런, neuron)로 구성
- 신경세포는 매우 간단한 처리만 담당
- 각 신경세포는 약 1천~10만 개의 다른 신경세포와 신경연접(synapse)을 통해 연결
- 신경연접을 지나가는 신호는 신경연접의 특성에 따라 증폭되거나 감소되어 전달됨



1. 인공 신경망의 기본 구조

■ 인공 신경망이란?

- 컴퓨터와 신경계통의 처리 방법 특성 비교

	CPU	뉴런
처리속도	▪ 수 나노 초(10^{-9} 초) 단위	▪ 수 밀리 초(10^{-3} 초) 단위
처리방법	▪ 나열된 명령어를 순차적으로 수행	▪ 대단위 병렬처리

- ☑ 신경 구조를 모델링하여 지능적 처리에 응용하려고 시도함

➡ **인공 신경회로망**(Artificial Neural Network, ANN)

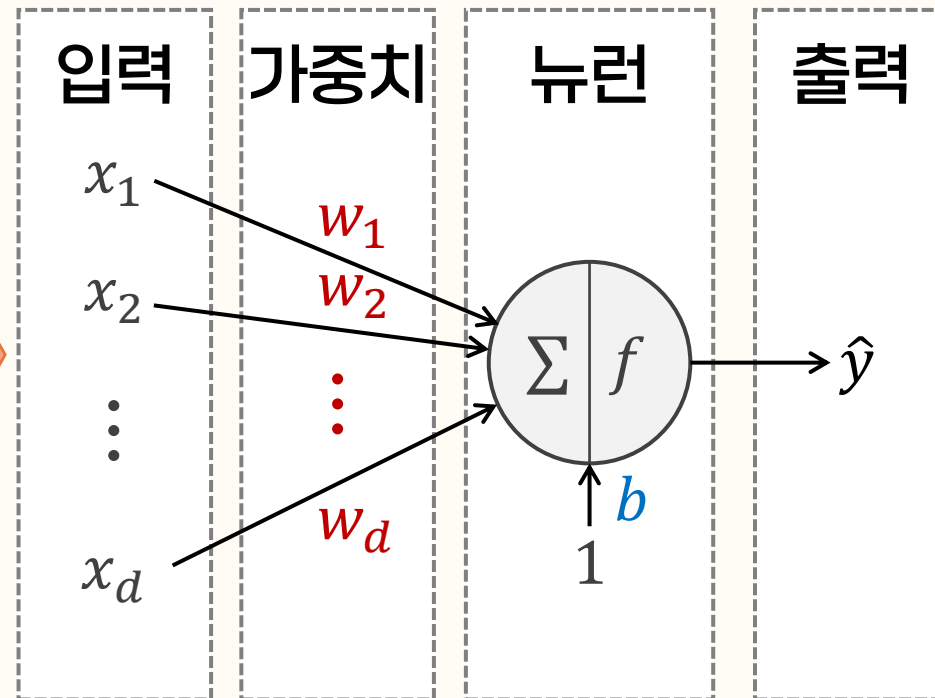
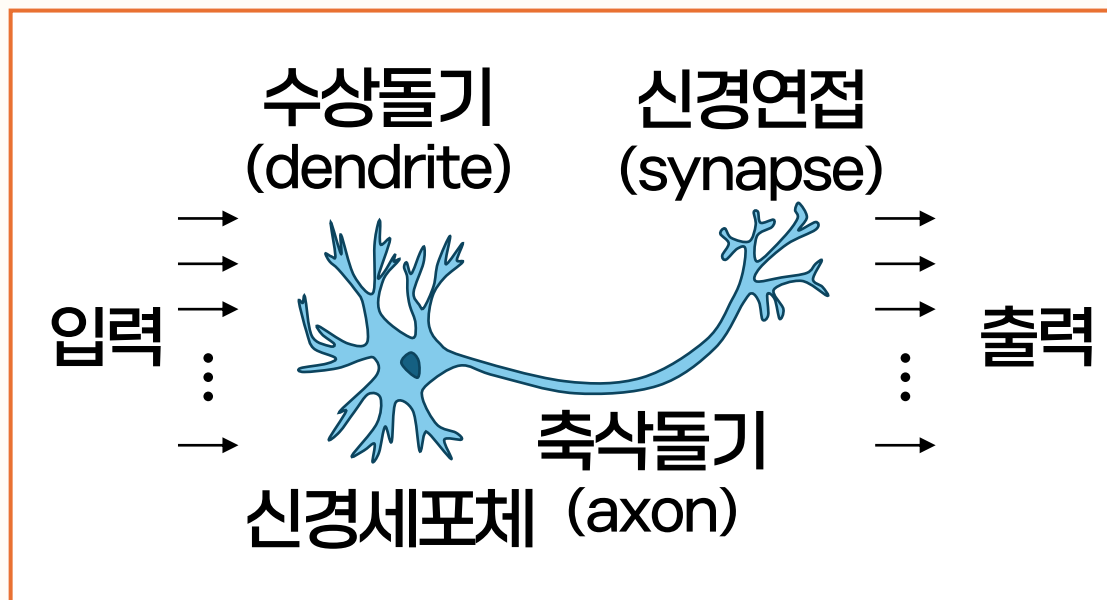
1. 인공 신경망의 기본 구조

■ 인공 신경망의 특성

- 입력된 내용을 합하고 변환을 하는 **간단한 연산기능**을 가지고 있는 **많은 수의 뉴런**으로 구성
- 각각의 뉴런은 다른 뉴런과 **방대한 연결**을 유지
- 수많은 뉴런이 동시에 동작하는 **대단위 병렬처리**
- 정보의 저장: **신경연접의 연결 가중치 벡터**를 통해 저장
 - ➡ 수많은 뉴런이 서로 연결되는 신경연접에 **분산 저장**됨
- **학습능력**: 학습 데이터에 따라 자동적으로 연결 가중치 조정
- **결함내성**: 일부 뉴런에 고장이 발생해도 전체 시스템의 성능이 급격하게 저하되지 않음

1. 인공 신경망의 기본 구조

■ 뉴런의 기본 구조



$$\hat{y} = f \left(\sum_{i=1}^d x_i w_i + b \right), \quad b: \text{바이어스(임계치)}, \quad f: \text{활성함수}$$

1. 인공 신경망의 기본 구조

■ 활성화함수(activation function)

- 비선형 특성을 갖는 함수를 사용함

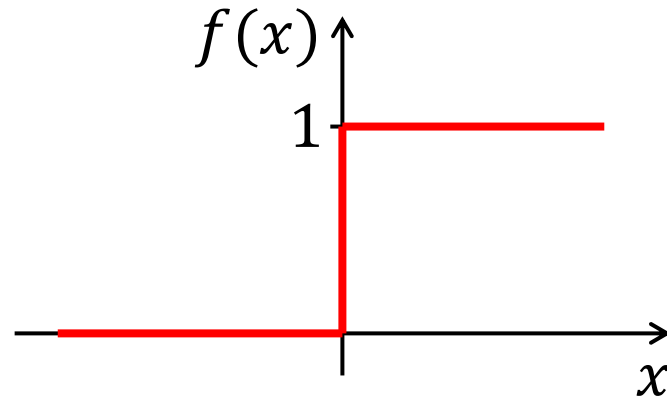
단위 계단함수

시그모이드

TanH

ReLU

$$f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



함수의 입력이 0보다 작으면 0, 그렇지 않으면 1을 출력함
(Unit step function, Heaviside step function)

1. 인공 신경망의 기본 구조

■ 활성화함수(activation function)

- 비선형 특성을 갖는 함수를 사용함

단위 계단함수

시그모이드

TanH

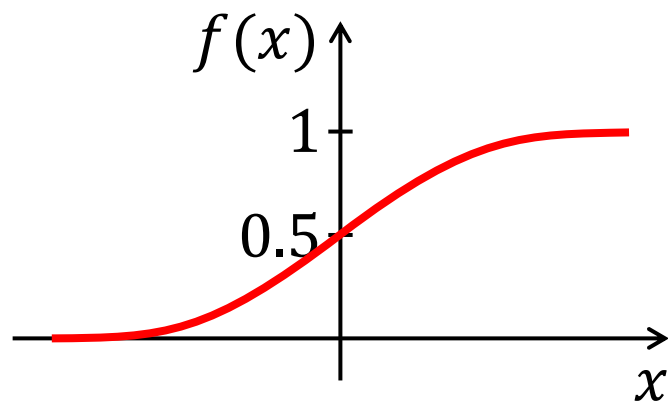
ReLU

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



전 구간에서 미분 가능

$$f'(x) = f(x)\{1 - f(x)\}$$



자동이득조절: 작은 입력에 대해서는 이득(gain)이 크고,
큰 입력에 대해서는 이득이 작음

1. 인공 신경망의 기본 구조

■ 활성화함수(activation function)

- 비선형 특성을 갖는 함수를 사용함

단위 계단함수

시그모이드

TanH

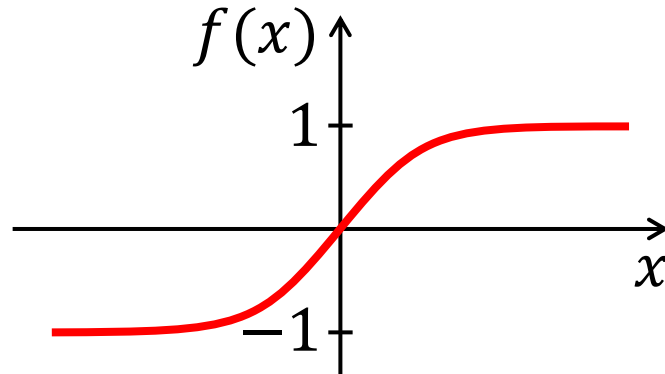
ReLU

$$\begin{aligned} f(x) &= \tanh(x) \\ &= \frac{e^x - e^{-x}}{e^x + e^{-x}} \end{aligned}$$



전 구간에서 미분 가능

$$f'(x) = 1 - f(x)^2$$



1. 인공 신경망의 기본 구조

■ 활성화함수(activation function)

- 비선형 특성을 갖는 함수를 사용함

단위 계단함수

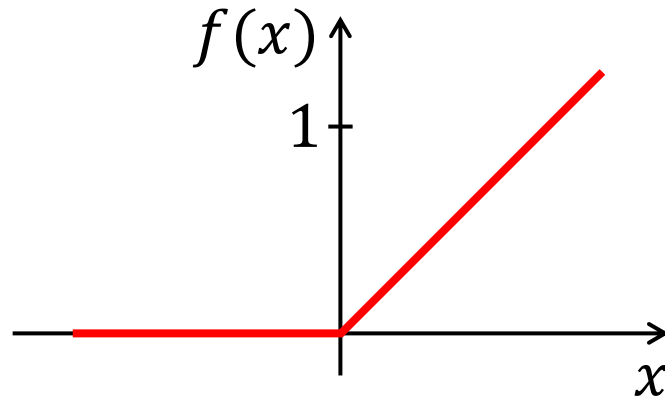
시그모이드

TanH

ReLU

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

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



Rectified Linear Unit

1. 인공 신경망의 기본 구조

■ 연결 형태

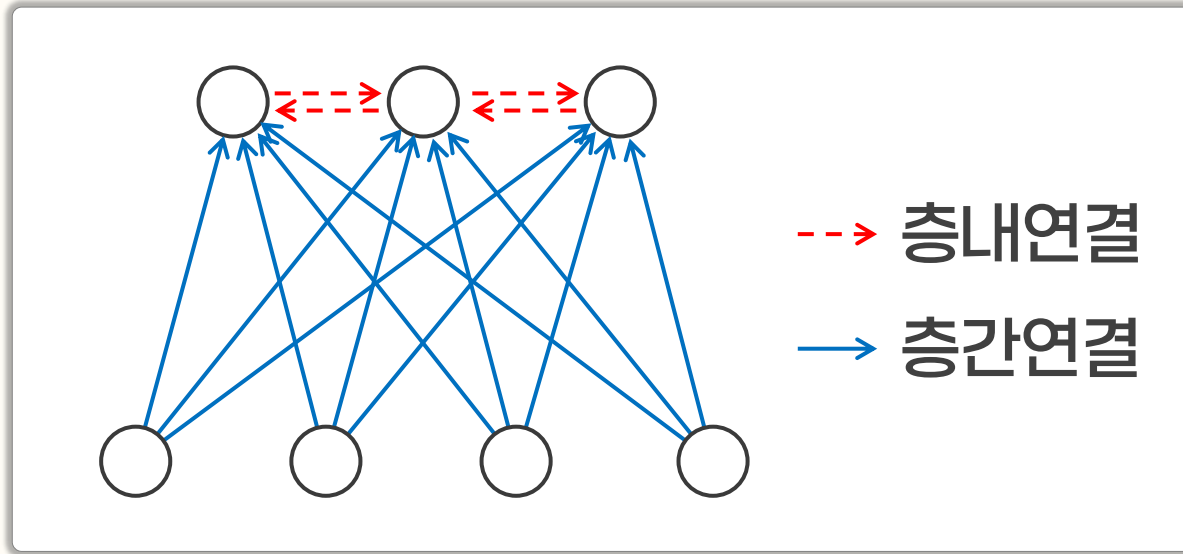
- 흥분성 연결 및 금지 연결

흥분성 연결	금지 연결
<ul style="list-style-type: none"> ■ 연결된 뉴런의 활성을 높임 	<ul style="list-style-type: none"> ■ 연결된 뉴런의 활성을 낮춤 

1. 인공 신경망의 기본 구조

■ 연결 형태

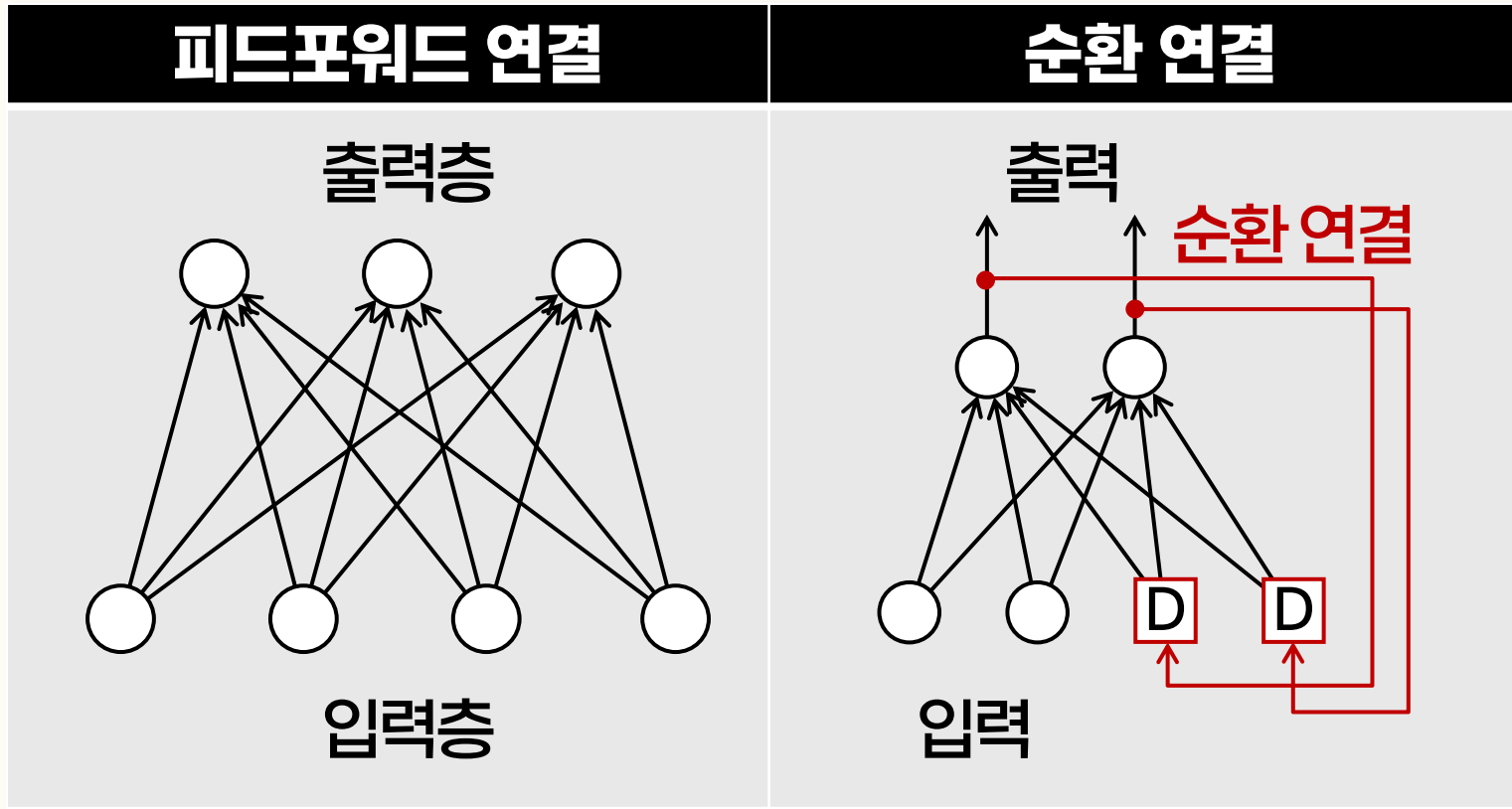
- 층내연결 및 층간연결



1. 인공 신경망의 기본 구조

■ 연결 형태

- 피드포워드 연결 및 순환 연결



2. 신경망의 학습

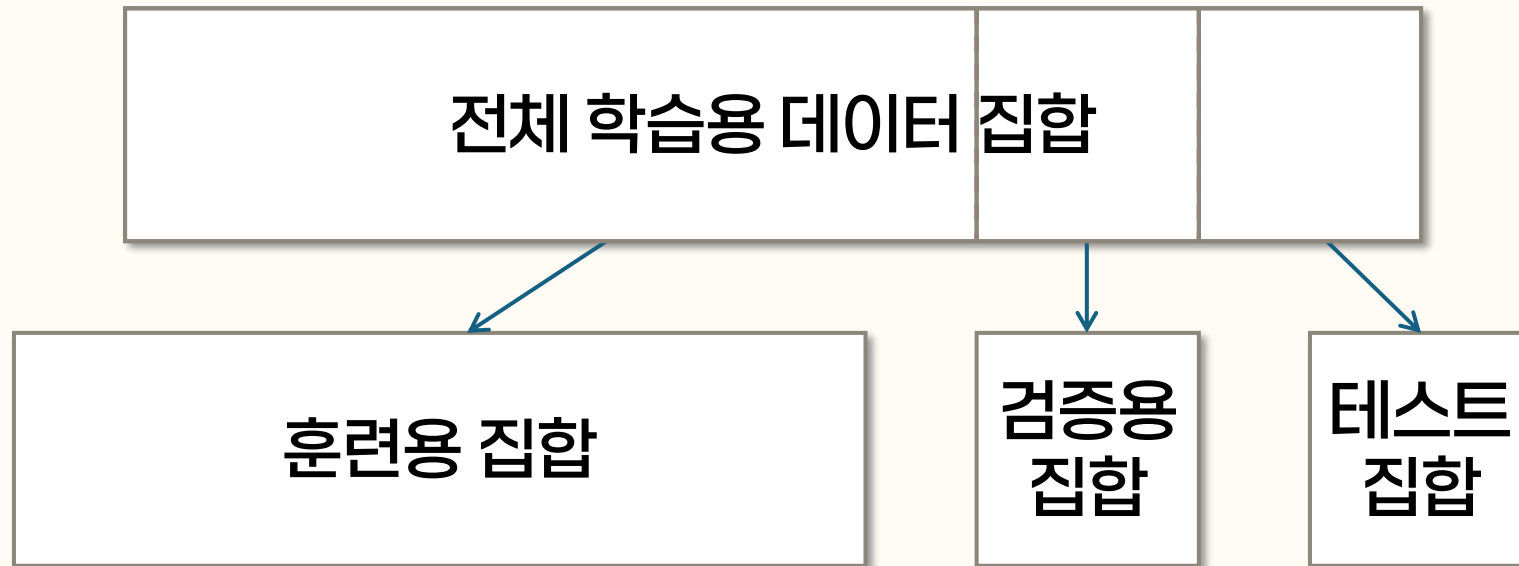
■ 신경망 학습의 개념

- 학습(learning): 신경망이 주어진 목적에 맞게 동작할 수 있도록 연결 가중치 w 의 값을 결정하는 제반 과정을 의미함
 - 손실함수(loss function) 정의
 - 하이퍼파라미터(hyperparameter) 설정
 - 신경망의 초기화
 - 훈련(training): 학습 대상 파라미터의 반복적 업데이트
 - 훈련 결과의 검증(validation)
 - 모델의 테스트(test)

2. 신경망의 학습

■ 학습 데이터 집합의 활용

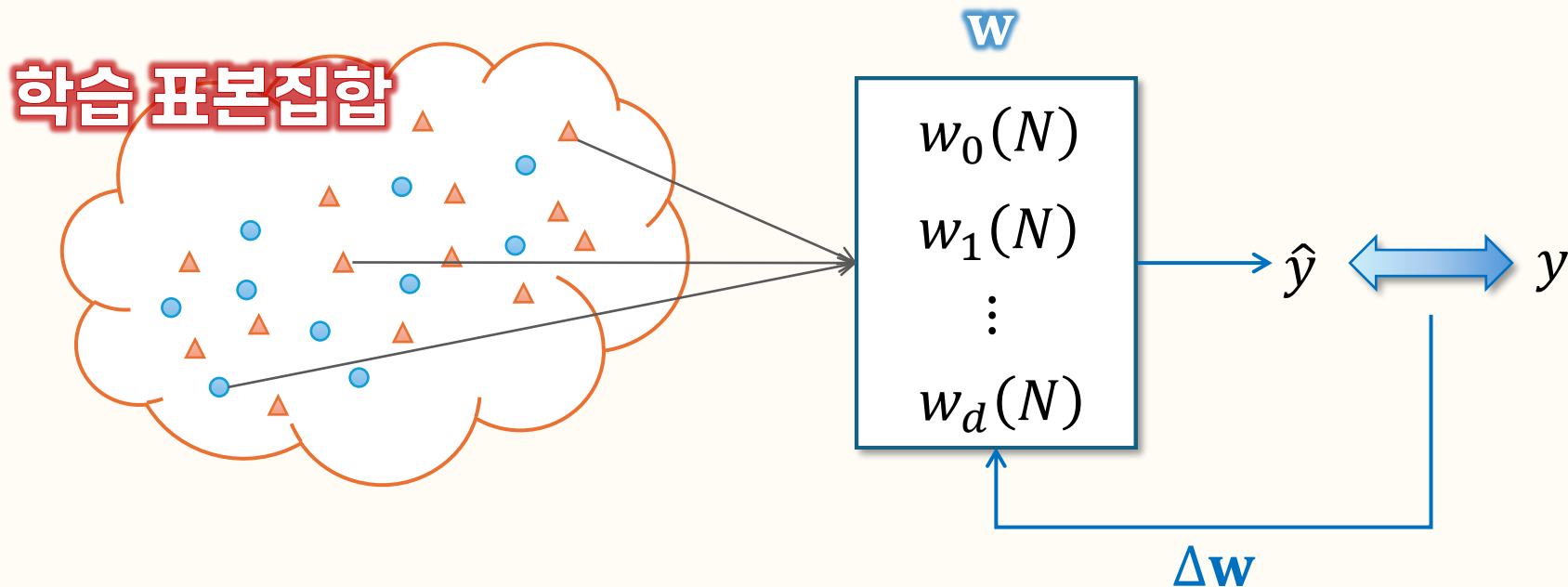
- 학습 데이터의 구성: 입력 데이터와 레이블(지도학습)
- 훈련, 검증, 테스트를 위한 데이터 활용



2. 신경망의 학습

■ 훈련의 진행 단위

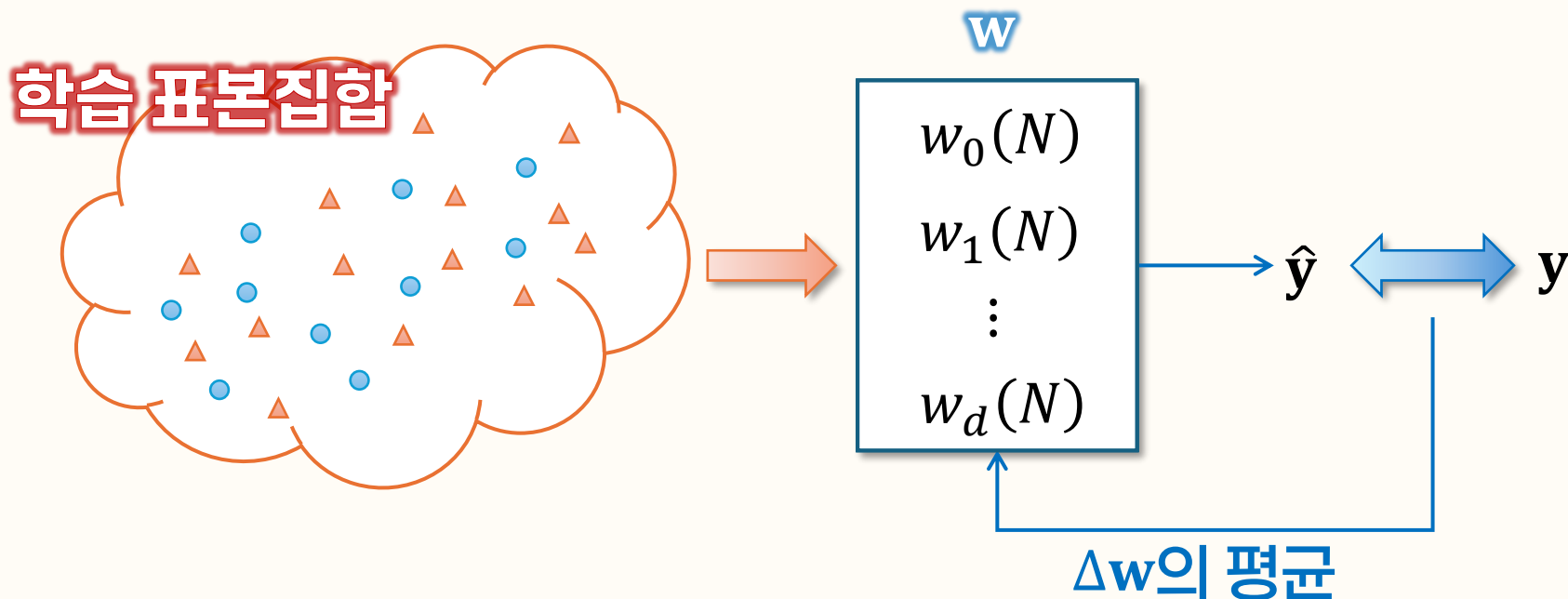
- 개별 학습표본 단위의 파라미터 업데이트
 - 확률적 경사하강법(stochastic gradient descent, SGD):
매 학습 단계에서 학습표본을 무작위로 섞어 훈련을 진행함



2. 신경망의 학습

■ 훈련의 진행 단위

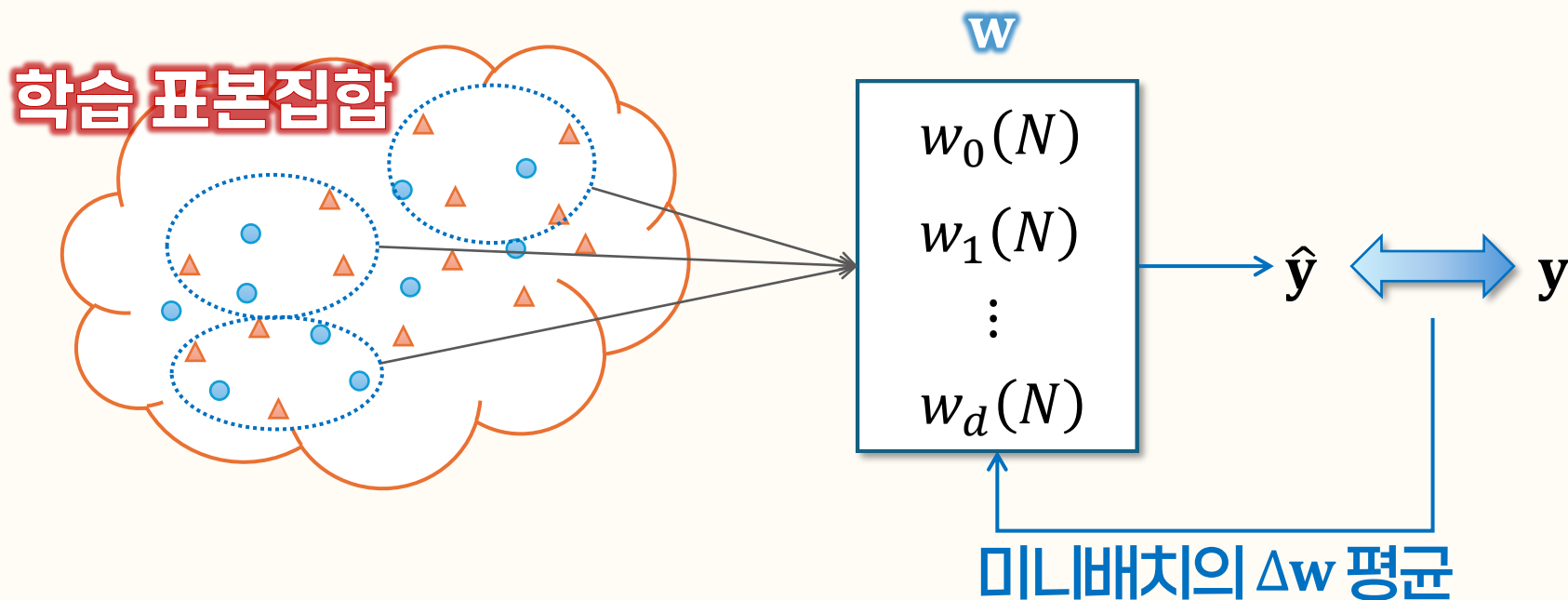
- 배치 학습(batch learning)
 - 학습표본 집합 내의 각각의 표본에 의한 파라미터 변화량을 누적함
 - 전체 학습표본에 대한 변화량의 평균으로 파라미터를 업데이트함



2. 신경망의 학습

■ 훈련의 진행 단위

- 미니배치(mini-batch) SGD
 - 전체 학습표본 집합을 작은 부분집합(미니배치)으로 분할
 - 미니배치 단위의 파라미터 변화량 평균으로 파라미터를 업데이트함





퍼셉트론 학습

1. 퍼셉트론 모델

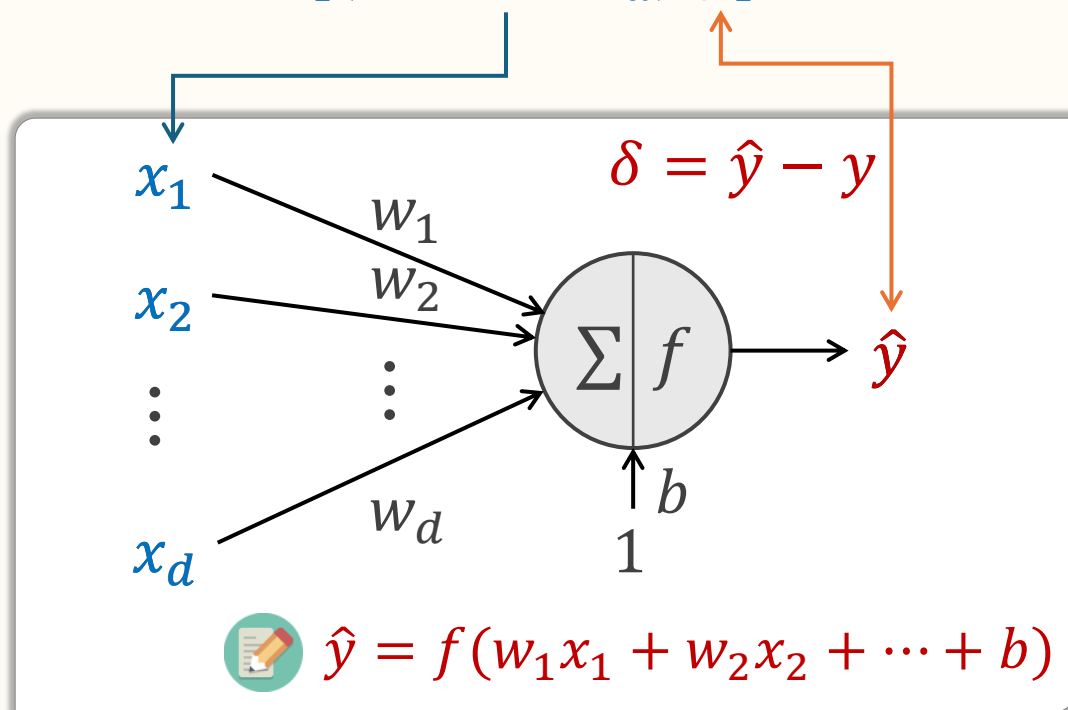
■ 퍼셉트론(Perceptron) 모델 개요

- Frank Rosenblatt – Cornell 대학
 - 1957년 퍼셉트론 학습이론 제시
 - 선형분리가 가능한 입력벡터의 집합에 대한 **선형 결정경계**를 학습할 수 있음을 입증
- 뉴런의 기능
 - 입력벡터의 각각의 요소는 연결 가중치를 거쳐 합산되며, 그 결과에 활성화함수를 가하여 출력
 - 활성화함수: **단위 계단 함수**(unit step function)
- 지도학습 방식 적용

1. 퍼셉트론 모델

■ 퍼셉트론의 학습

 학습 데이터: $[(x_1, x_2, \dots, x_d), y]$

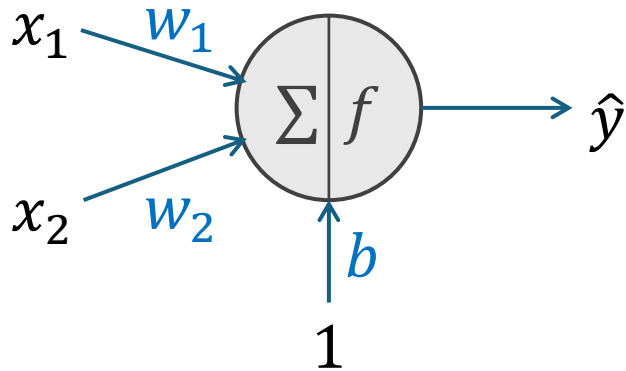


- $n + 1$ 번째 훈련 단계에서 연결 가중치의 업데이트

➡ $w_i(n + 1) = w_i(n) - \eta \delta x_i, i = 0, 1, 2, \dots, d$

2. 퍼셉트론 학습의 예

■ 학습 데이터



$$\hat{y} = f(w_1x_1 + w_2x_2 + b)$$

학습 데이터

1(0.25, 0.75) **0**(1.25, 1.75)

1(0.5, 0.5) **0**(1.75, 1.25)

1(0.75, 0.25) **0**(1.5, 1.5)

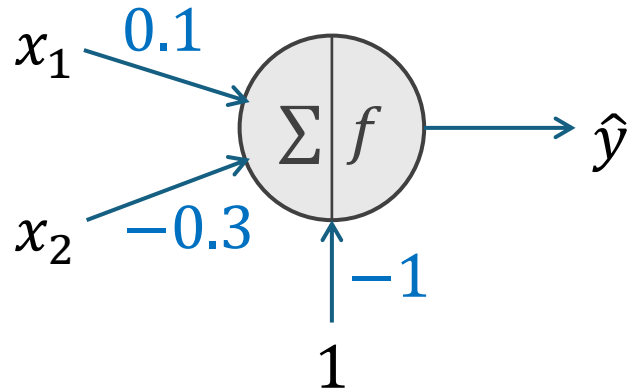
1(0.0, 1.5) **0**(2.5, 0.0)

1(1.5, 0.0) **0**(0.0, 2.5)

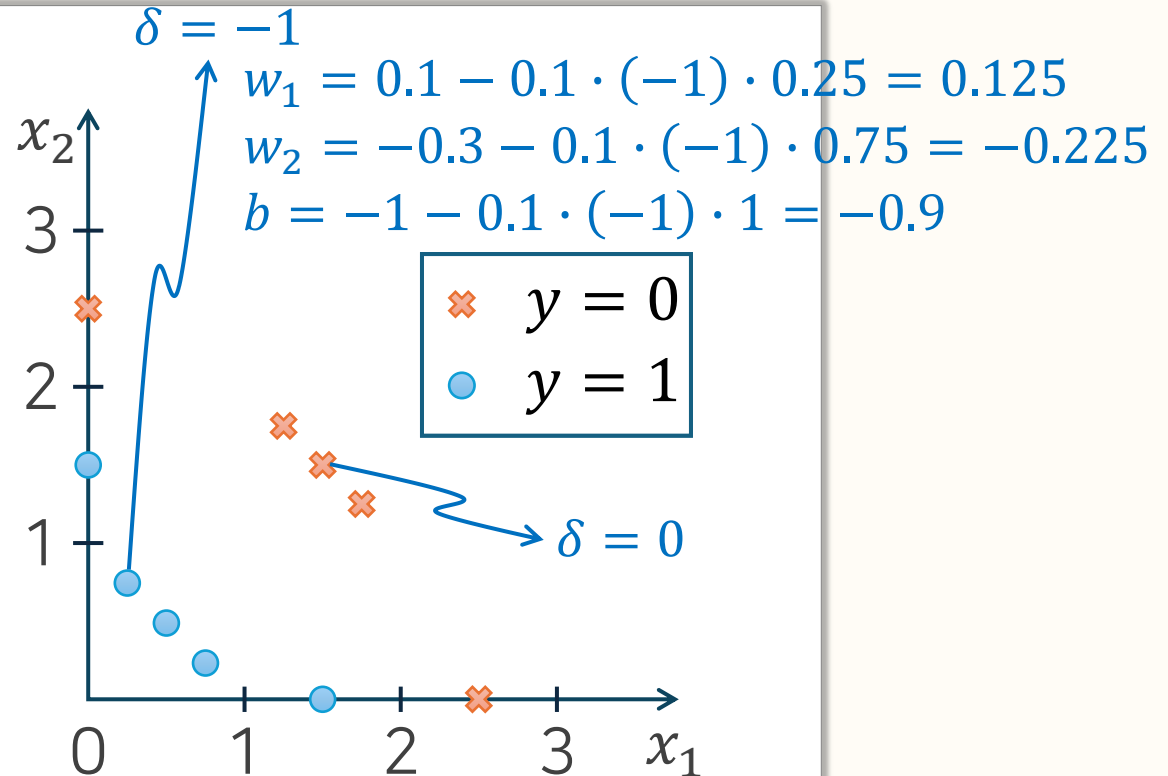
 학습률: $\eta = 0.1$

2. 퍼셉트론 학습의 예

■ 초기상태: $n = 0 \rightarrow w_1 = 0.1, w_2 = -0.3, b = -1$



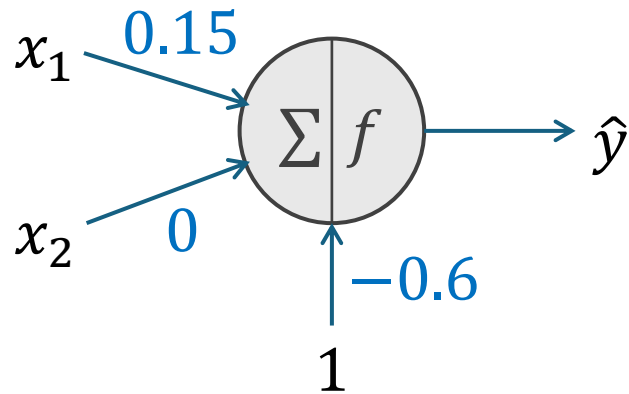
$$\hat{y} = f(0.1x_1 - 0.3x_2 - 1)$$



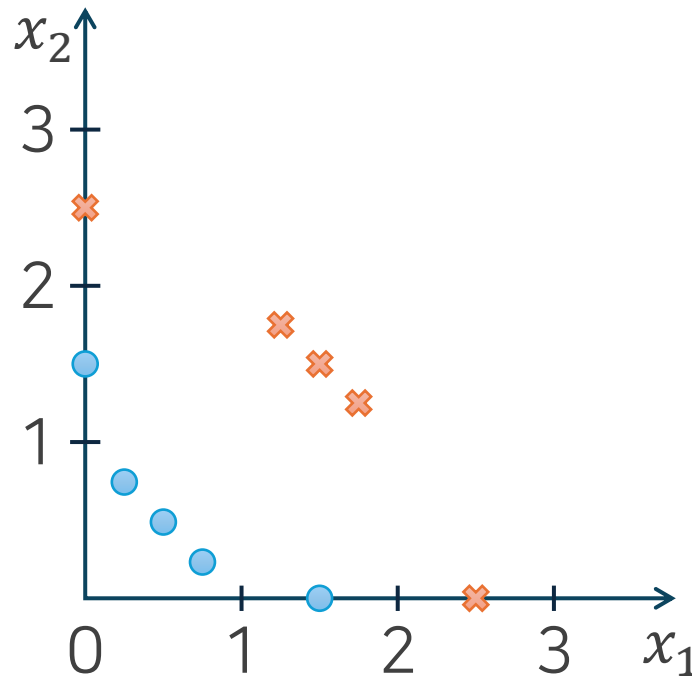
학습률: $\eta = 0.1$

2. 퍼셉트론 학습의 예

■ $n = 1 \rightarrow w_1 = 0.15, w_2 = 0, b = -0.6$



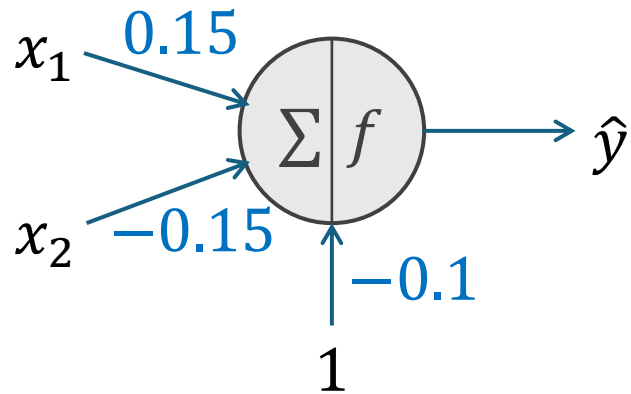
$$\hat{y} = f(1.5x_1 - 0.6)$$



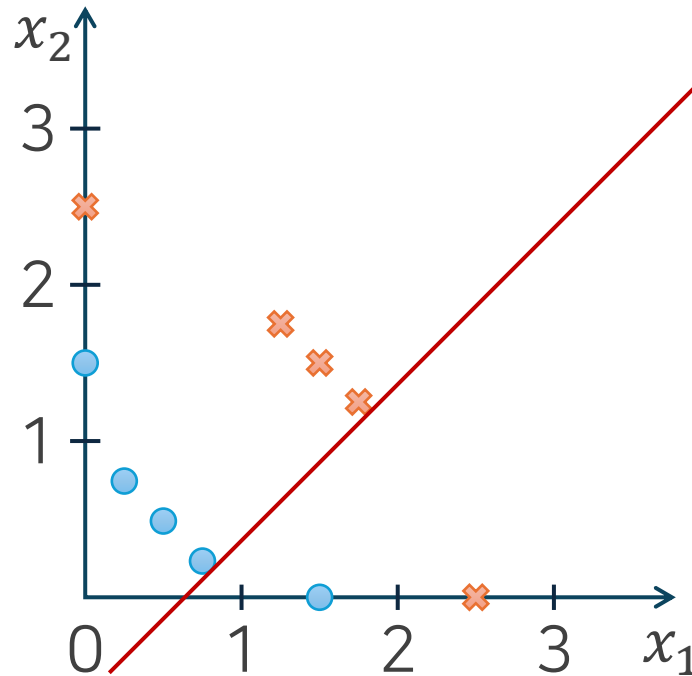
학습률: $\eta = 0.1$

2. 퍼셉트론 학습의 예

■ $n = 3 \rightarrow w_1 = 0.15, w_2 = -0.15, b = -0.1$



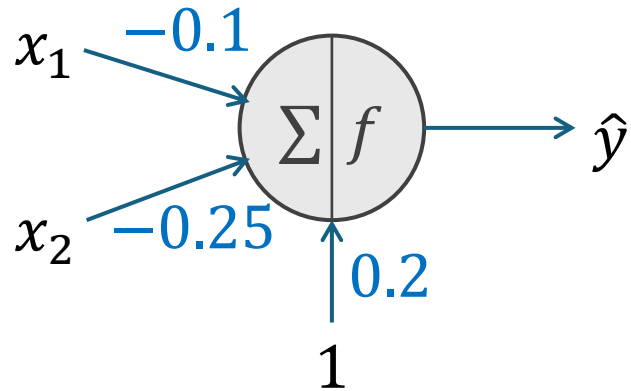
$$\hat{y} = f(0.15x_1 - 0.15x_2 - 0.1)$$



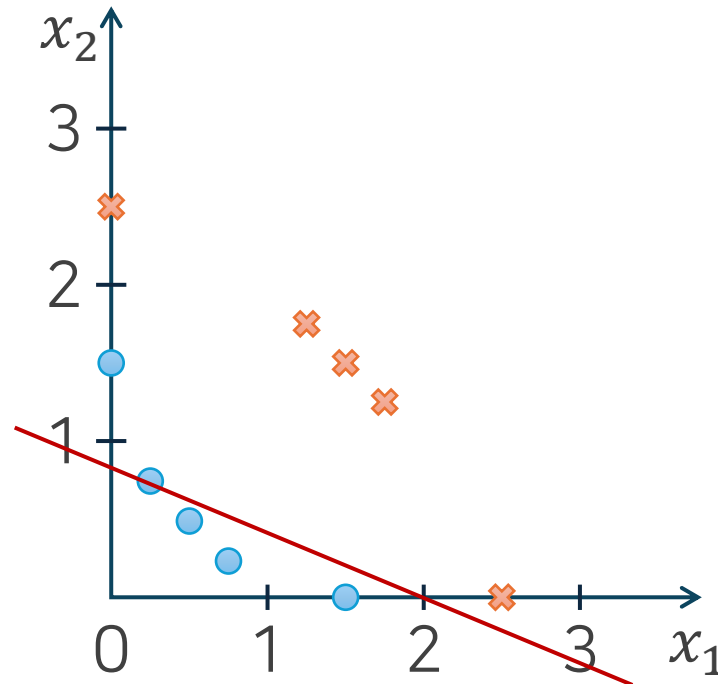
학습률: $\eta = 0.1$

2. 퍼셉트론 학습의 예

■ $n = 5 \rightarrow w_1 = -0.1, w_2 = -0.25, b = 0.2$



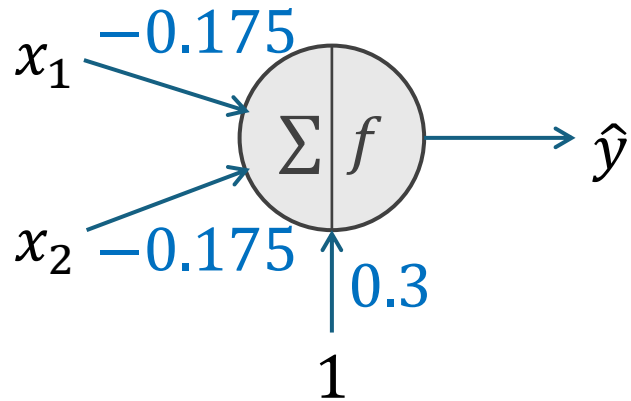
$$\hat{y} = f(-0.1x_1 - 0.25x_2 + 0.2)$$



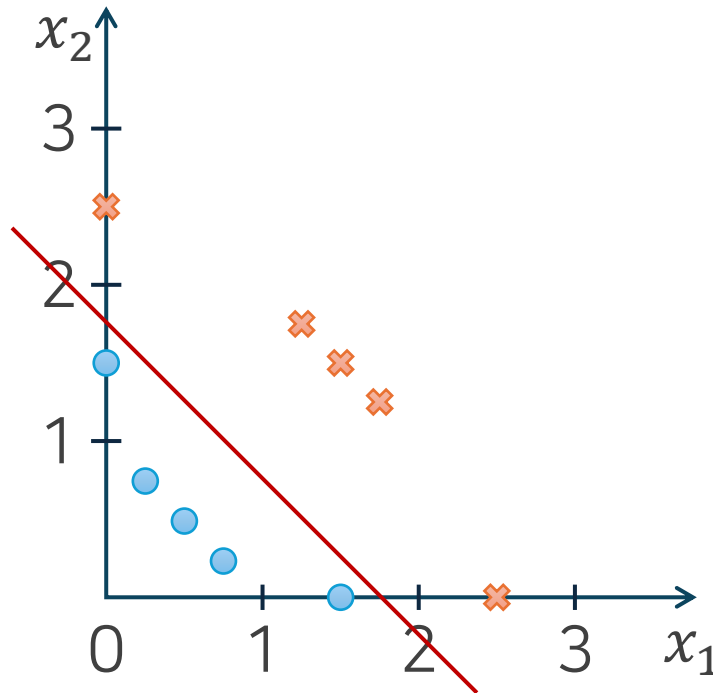
학습률: $\eta = 0.1$

2. 퍼셉트론 학습의 예

■ $n = 6 \rightarrow w_1 = -0.175, w_2 = -0.175, b = 0.3$



$$\hat{y} = f(-0.175x_1 - 0.175x_2 + 0.3)$$



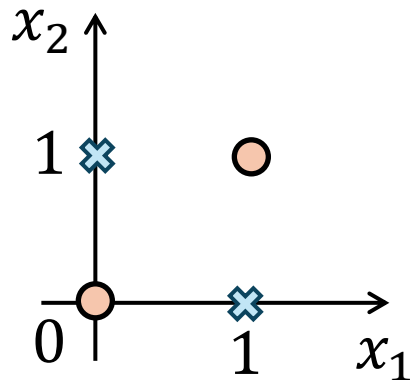
학습률: $\eta = 0.1$

3. 퍼셉트론 학습의 특성

■ 퍼셉트론 학습모델의 성과와 한계

- **성과**: 학습 데이터를 이용하여 반복 학습하면 이를 분리할 수 있는 **선형 결정경계**를 형성
- **한계**: 선형분리가 되지 않는 문제는 해결 불가

예 XOR 문제

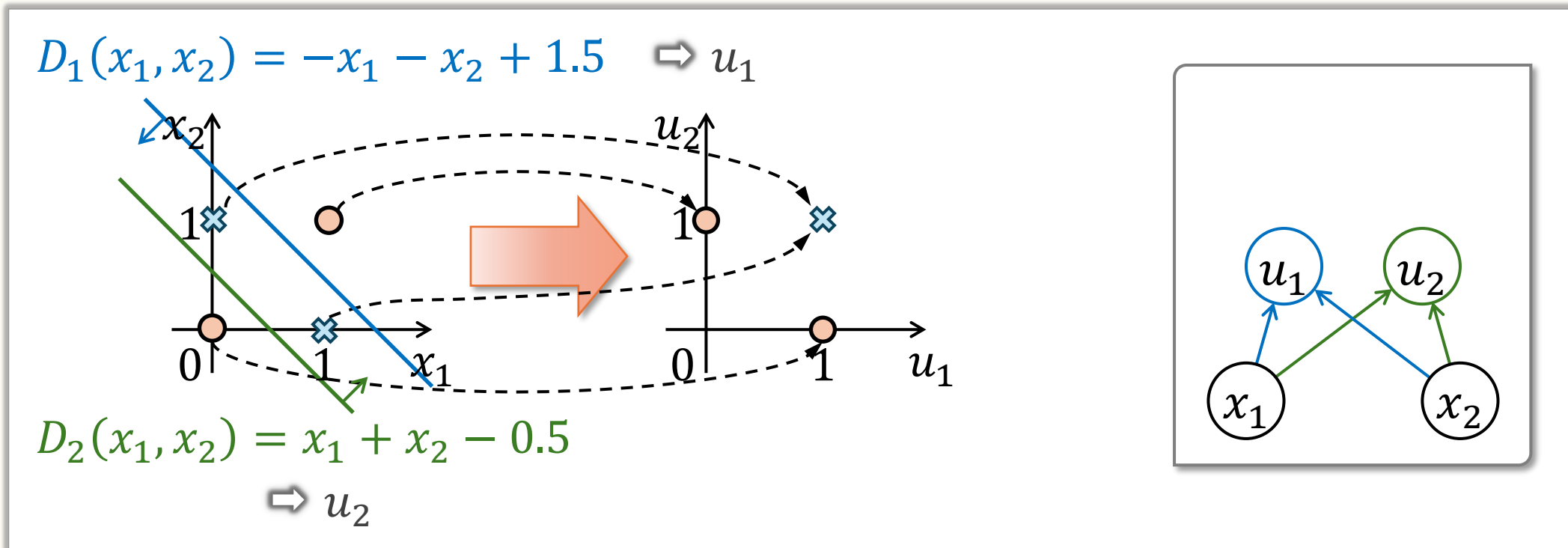


○ : 출력 = 0
× : 출력 = 1

3. 퍼셉트론 학습의 특성

❑ XOR 문제의 해법: 2단계 처리

- **1단계:** 동일한 입력에 2개의 퍼셉트론을 구성하여 각각의 출력 u_1 과 u_2 로 새로운 2차원 공간 형성

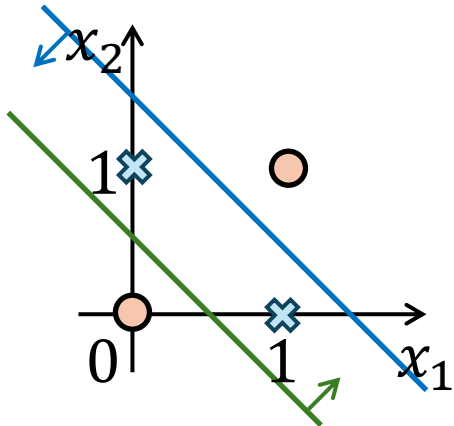


3. 퍼셉트론 학습의 특성

■ XOR 문제의 해법: 2단계 처리

- **2단계:** u_1-u_2 공간에 사상된 결과를 u_1 과 u_2 를 입력으로 하는 퍼셉트론에 의해 결정경계 형성

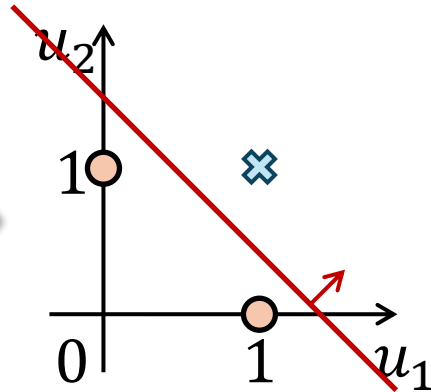
$$D_1(x_1, x_2) = -x_1 - x_2 + 1.5 \Rightarrow u_1$$



$$D_2(x_1, x_2) = x_1 + x_2 - 0.5$$

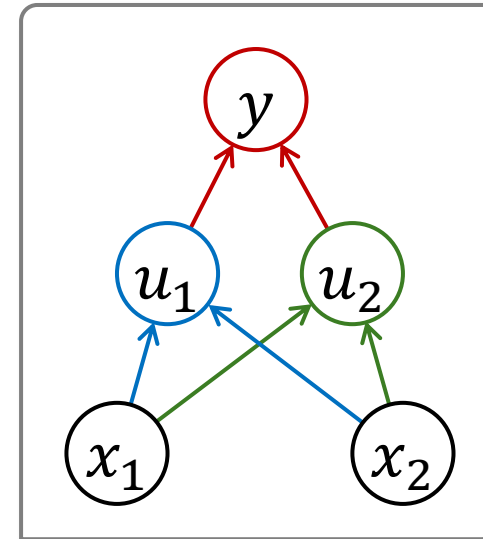
$$\Rightarrow u_2$$

How?



$$D(u_1, u_2)$$

$$= u_1 + u_2 - 1.5$$



정리하기

- ✓ 인공 신경회로망은 사람의 중추신경 구조를 모델링하여 지능적 처리에 응용하기 위한 모델로서, 단순한 연산기능을 수행하는 많은 수의 뉴런들이 서로 방대한 연결을 유지하며, 이들 사이의 연결 가중치를 조정함으로써 학습을 수행한다.
- ✓ 기본적인 뉴런의 기능은 각각의 입력에 가중치를 곱하여 전달된 값에 바이어스를 더한 후 활성화함수를 거쳐 출력하는 것이다.
- ✓ 뉴런의 연결 유형은 흥분성 연결과 금지 연결, 층내 연결과 층간 연결, 피드포워드 연결과 순환연결 등 모델에 따라 여러 가지 형태를 취한다.

정리하기

- ✓ 신경망의 학습은 신경망이 주어진 목적에 맞게 동작할 수 있도록 연결 가중치의 값을 결정하는 제반 과정을 의미하는데, 여기에는 모델의 하이퍼파라미터 설정, 신경망의 초기화, 훈련, 검증 등이 포함된다.
- ✓ 퍼셉트론 모델은 연결 가중치를 통해 전달되는 입력과 바이어스의 합에 단위 계단 함수 형태의 활성화함수를 가하여 출력하는 뉴런으로 구성된다. 단층 퍼셉트론은 지도학습을 통해 선형 결정경계를 형성한다.
- ✓ 선형 결정경계로 해결되지 않는 문제(예: XOR 문제)는 단층 퍼셉트론으로는 학습할 수 없다.

13강

다음시간안내 ▶▶▶

인공 신경망(2)