01

인공지능, 머신러닝, 딥러닝

ARTIFICIAL INTELLIGENCE

IS NOT NEW

ARTIFICIAL INTELLIGENCE

Any technique which enables computers to mimic human behavior

MACHINE LEARNING

AI techniques that give computers the ability to learn without being explicitly programmed to do so

DEEP LEARNING

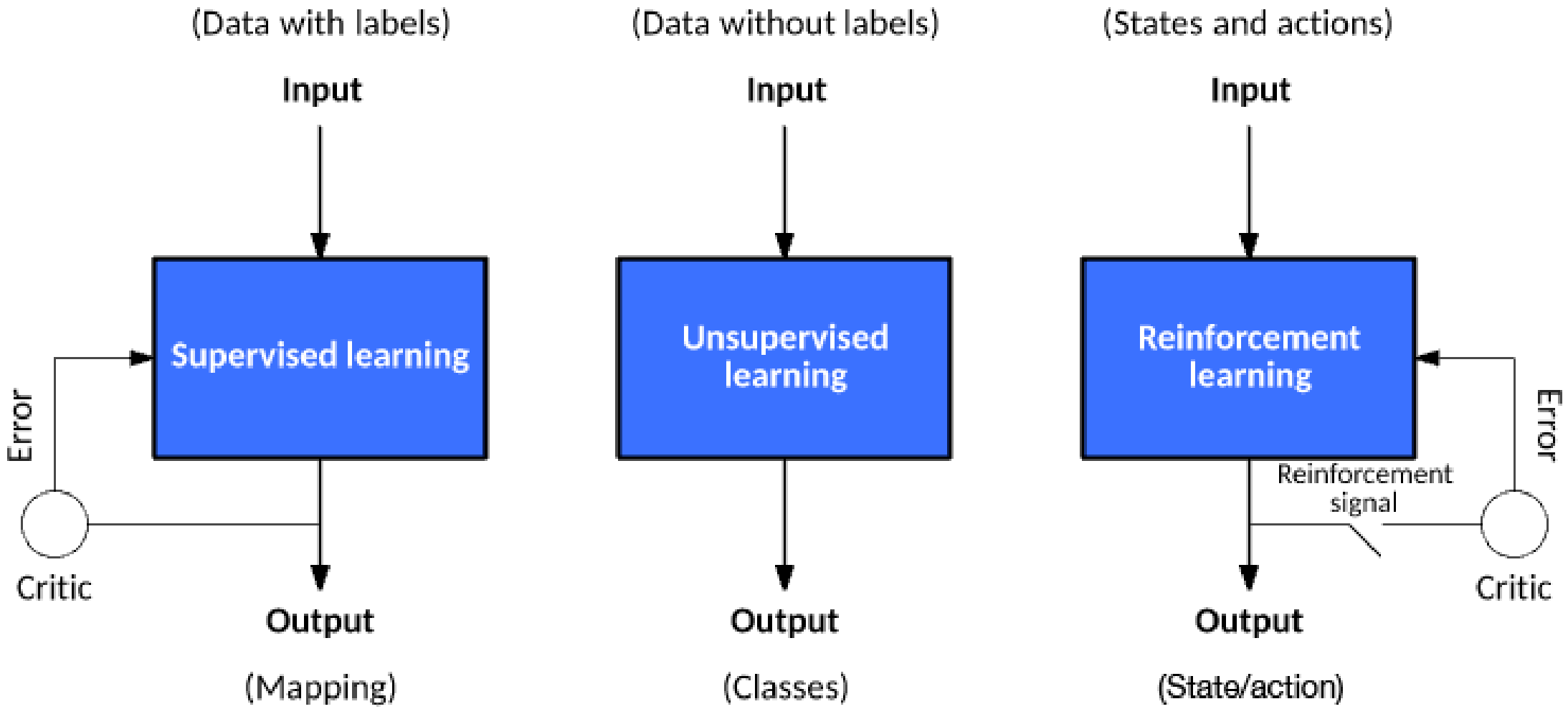
A subset of ML which make the computation of multi-layer neural networks feasible

1950's 1960's 1970's 1980's 1990's 2000's 2010s

ORACLE

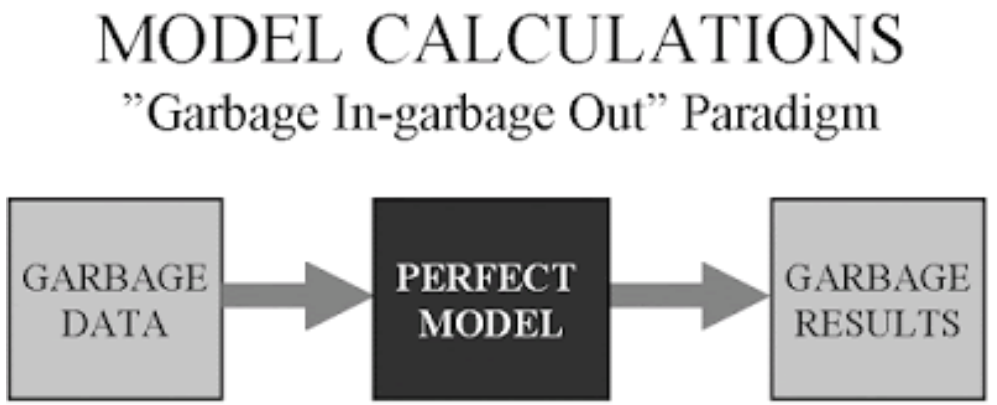
https://blogs.oracle.com/bigdata/difference-ai-machine-learning-deep-learning

머신러닝의 종류(지도, 비지도, 강화학습)



https://developer.ibm.com/articles/cc-models-machine-learning/

Garbage In Garbage Out



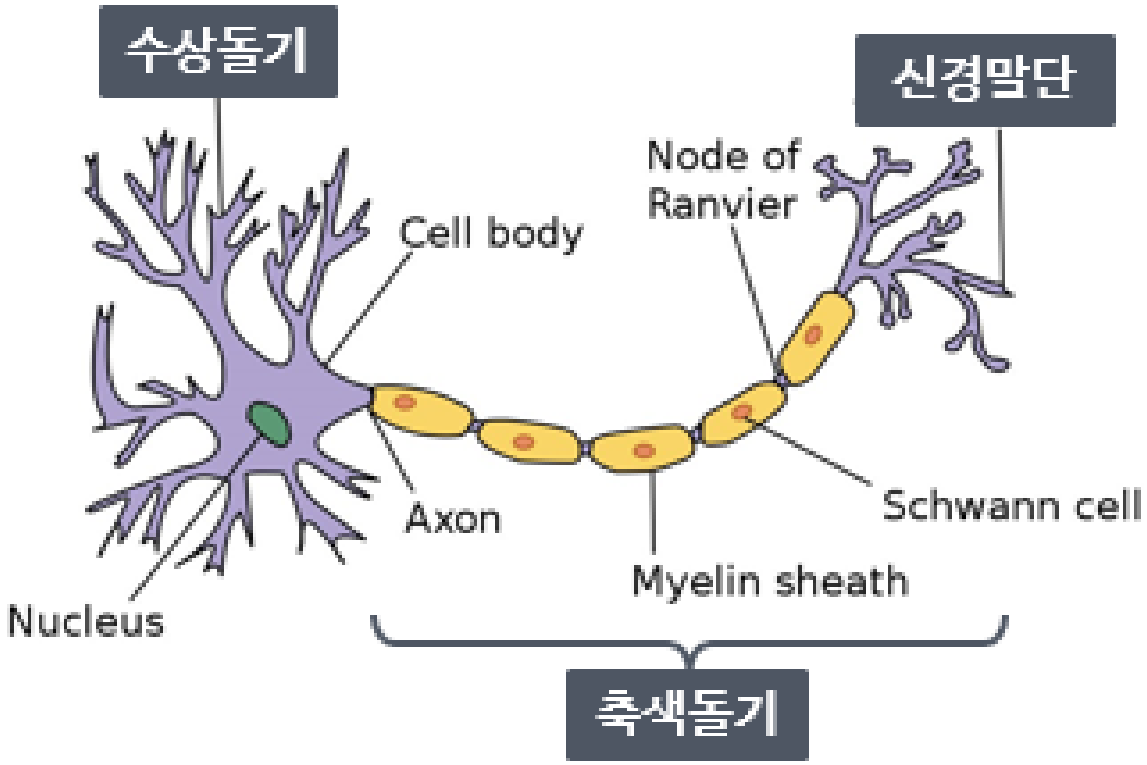
https://thedailyomnivore.net/2015/12/02/garbage-in-garbage-out/

http://redcliffsofdawlish.blogspot.com/2016/01/journalism-garbage-in-garbage-out-gigo.html

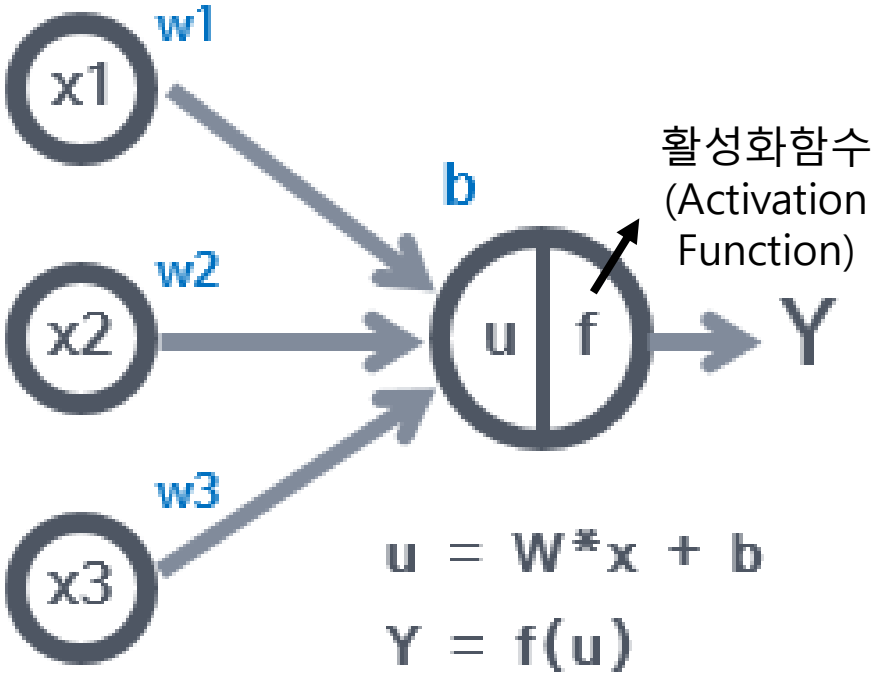
02

신경망과 퍼셉트론

딥러닝의 주요 아이디어 – 뉴런과 인공 뉴런



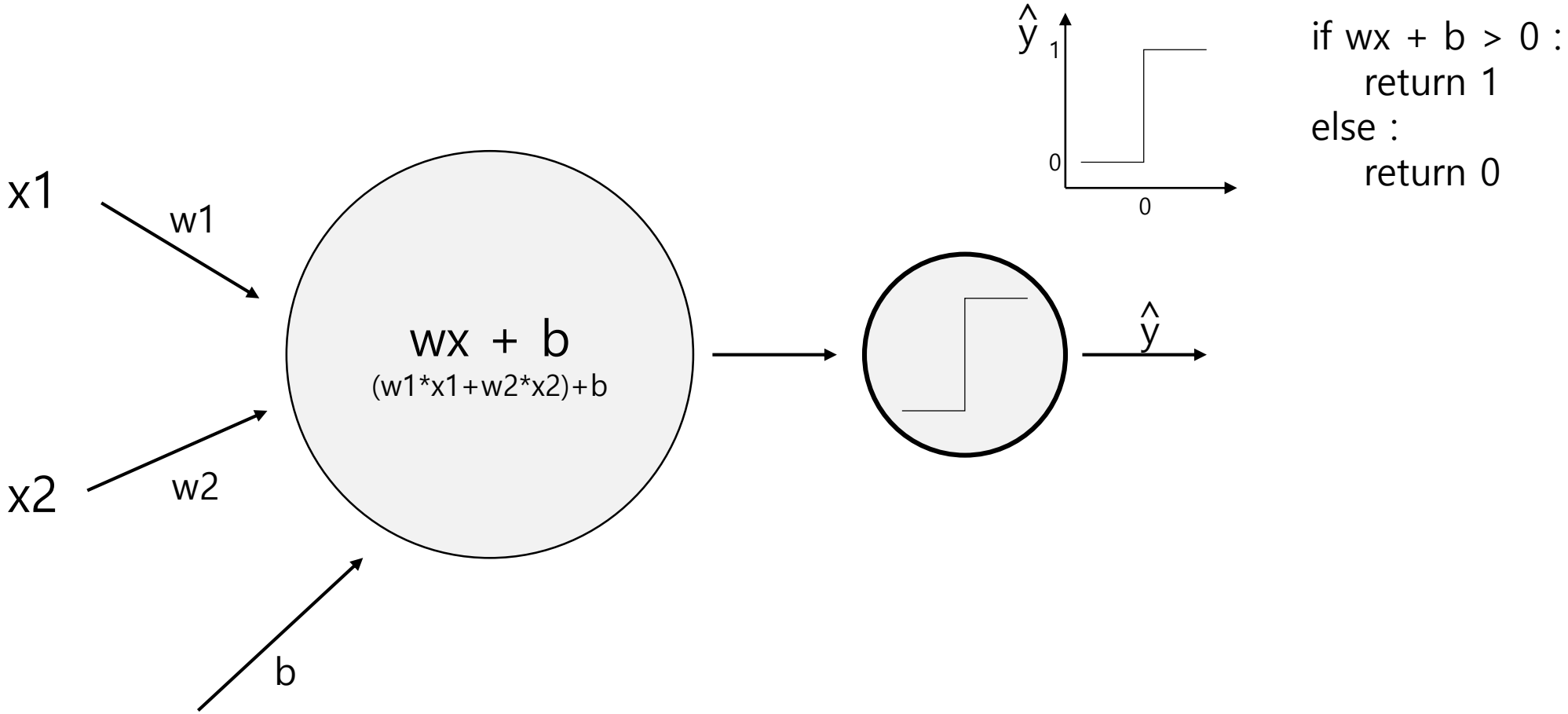
뉴런의 구조



인공 신경의 구조

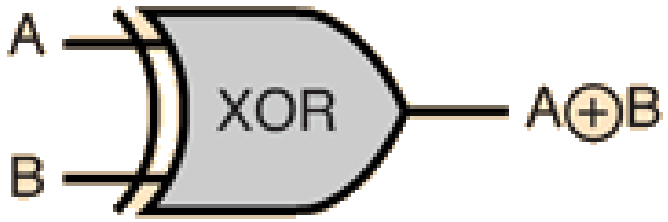
https://blog.naver.com/PostView.nhn?blogId=windowsub0406&logNo=220878782311

인공 뉴런의 초기 아이디어 – 퍼셉트론

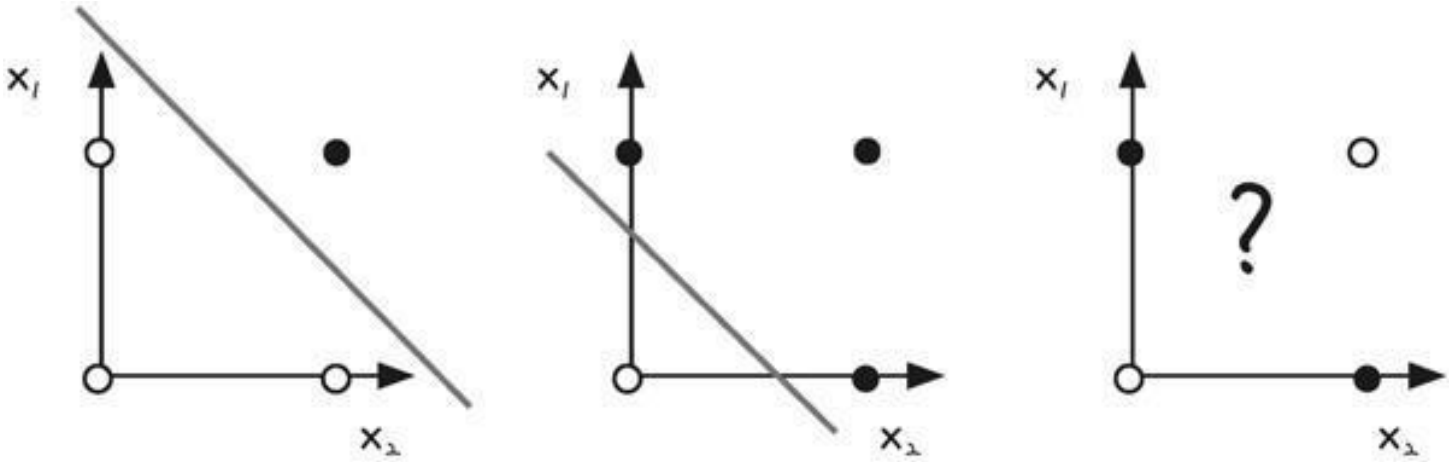



```
if wx + b > 0 :
    return 1
else :
    return 0
```

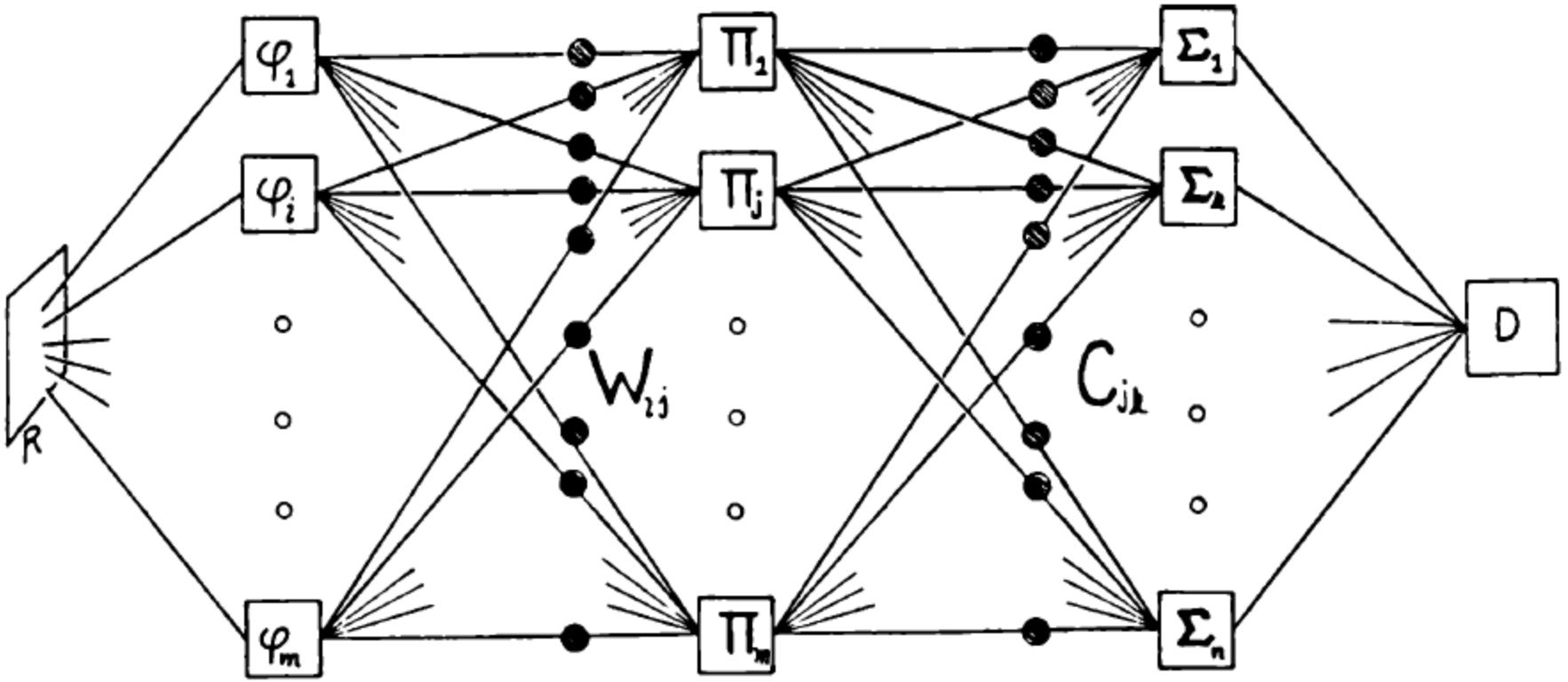
퍼셉트론의 한계 : XOR 문제



A	B	Out
0	0	0
0	1	1
1	0	1
1	1	0



퍼셉트론의 한계(1969, Marvin Minsky & Seymour Papert)



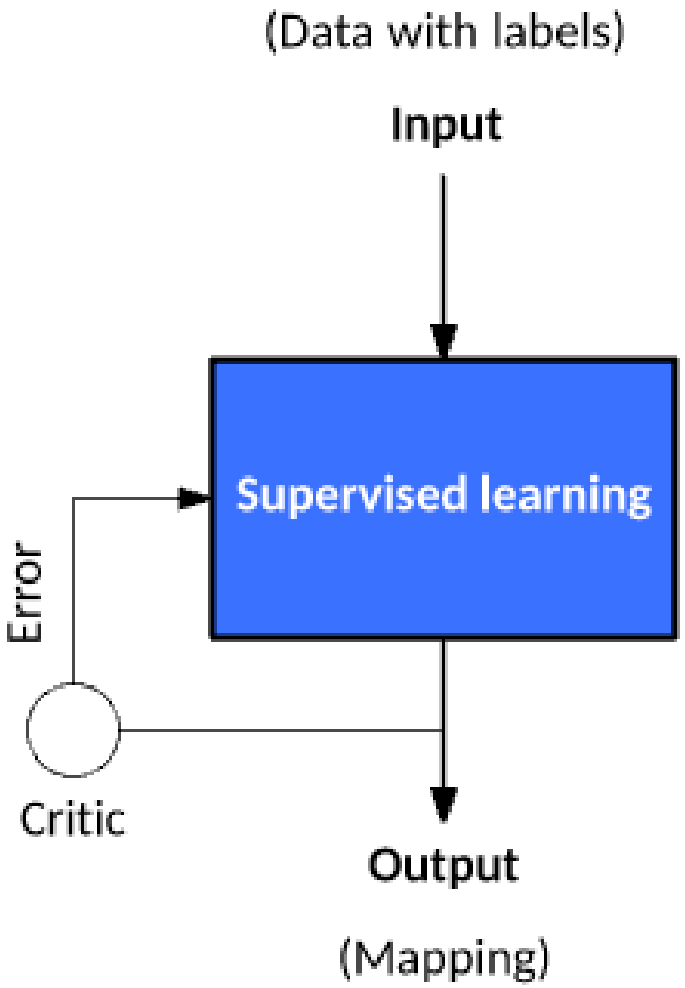
It ought to be possible to devise a training algorithm to optimize the weights in this using, say, the magnitude of a reinforcement signal to communicate to the net the cost of an error. We have not investigated this.

예를 들어, 오류의 비용을 그물에 전달하기 위해 보강 신호의 크기를 사용하여
가중치를 최적화하는 훈련 알고리즘을 고안 할 수 있어야한다.

03

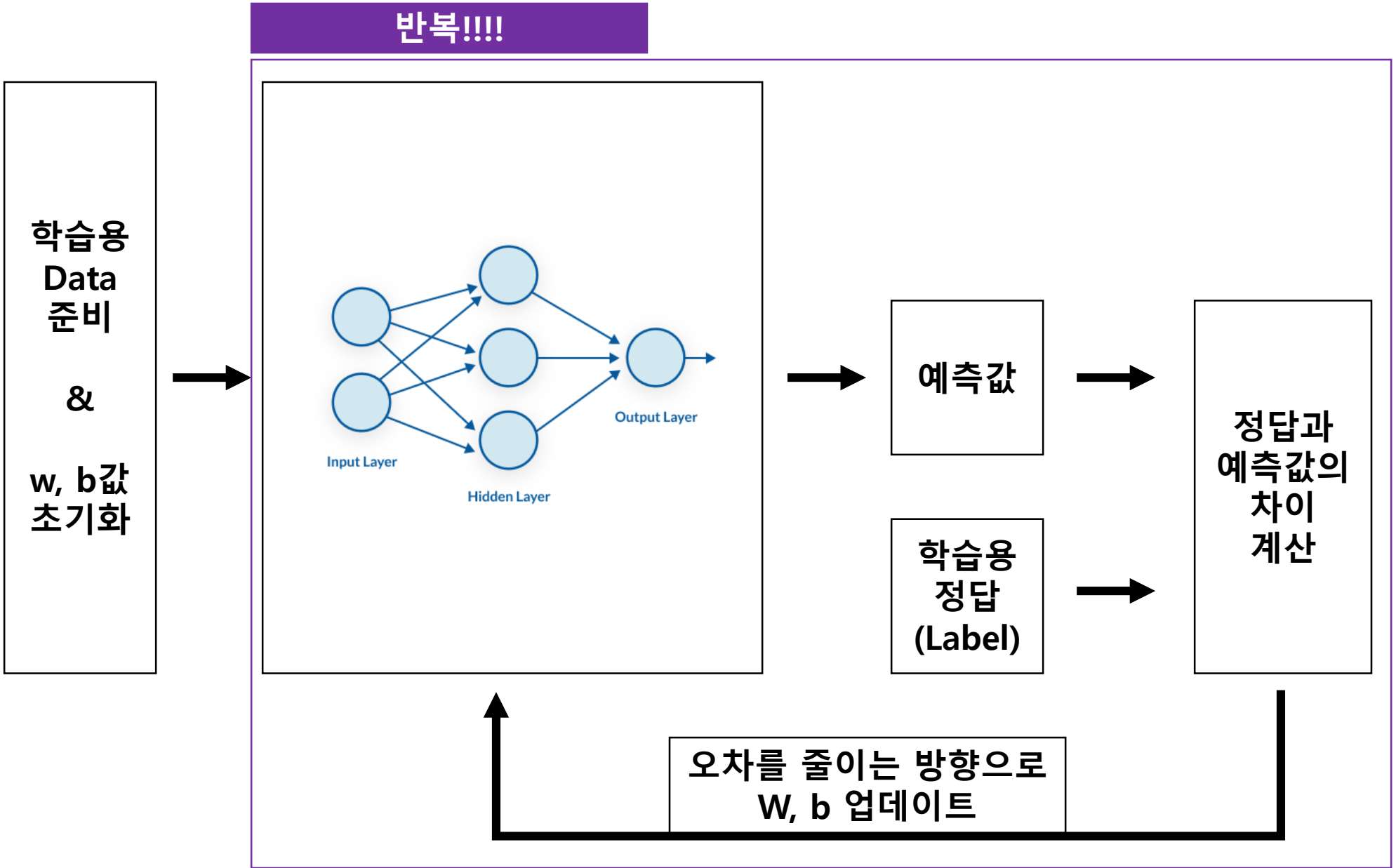
딥러닝 개요

지도학습의 개념



https://developer.ibm.com/articles/cc-models-machine-learning/

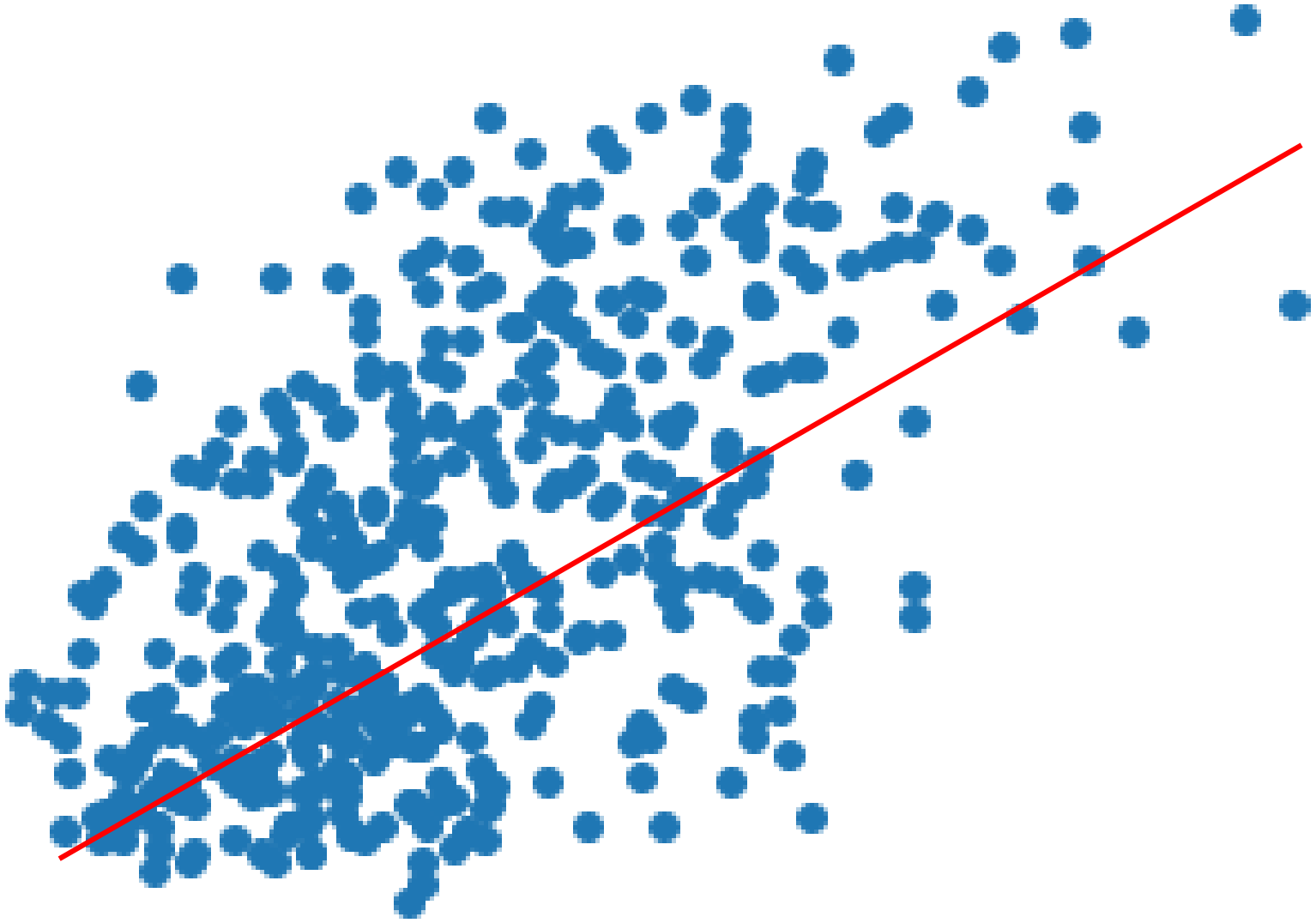
딥러닝 - 전체적인 학습 과정



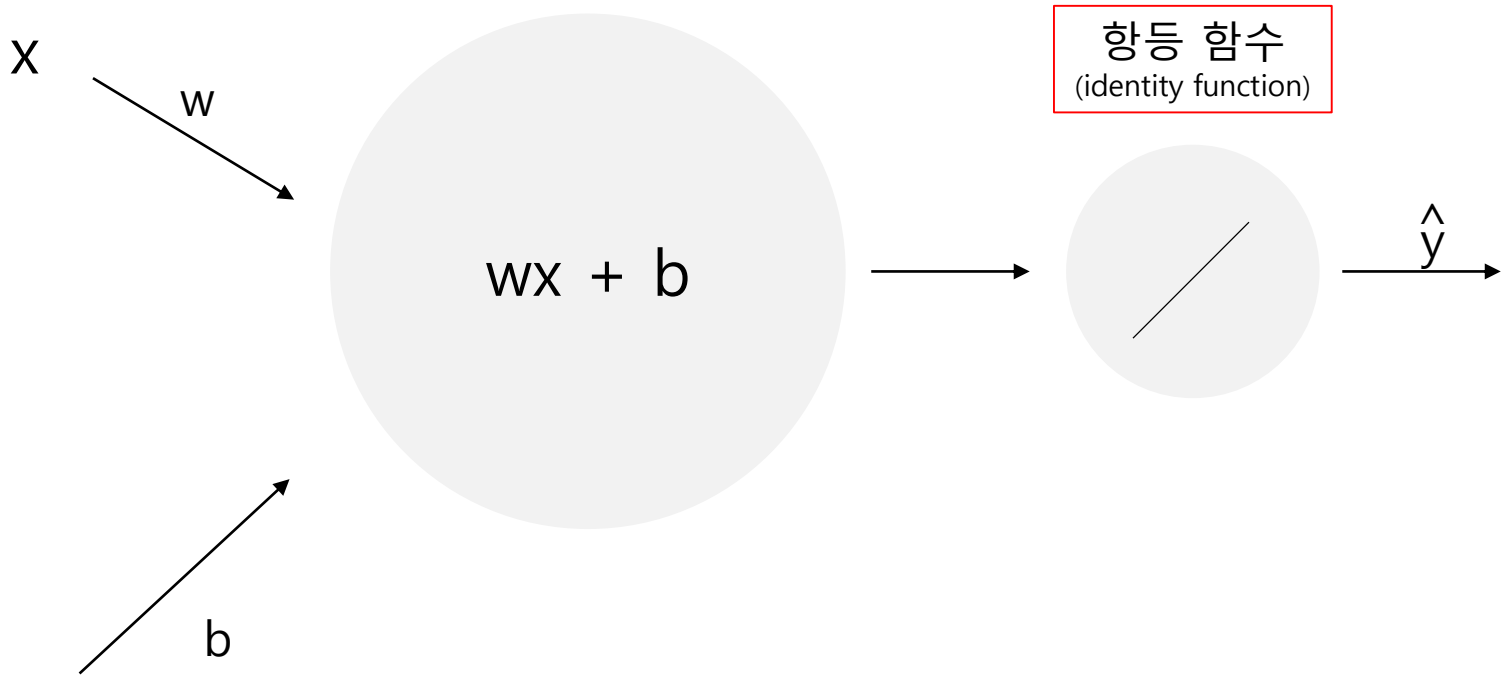
04

연습문제 – 당뇨병 예측 모델 만들기

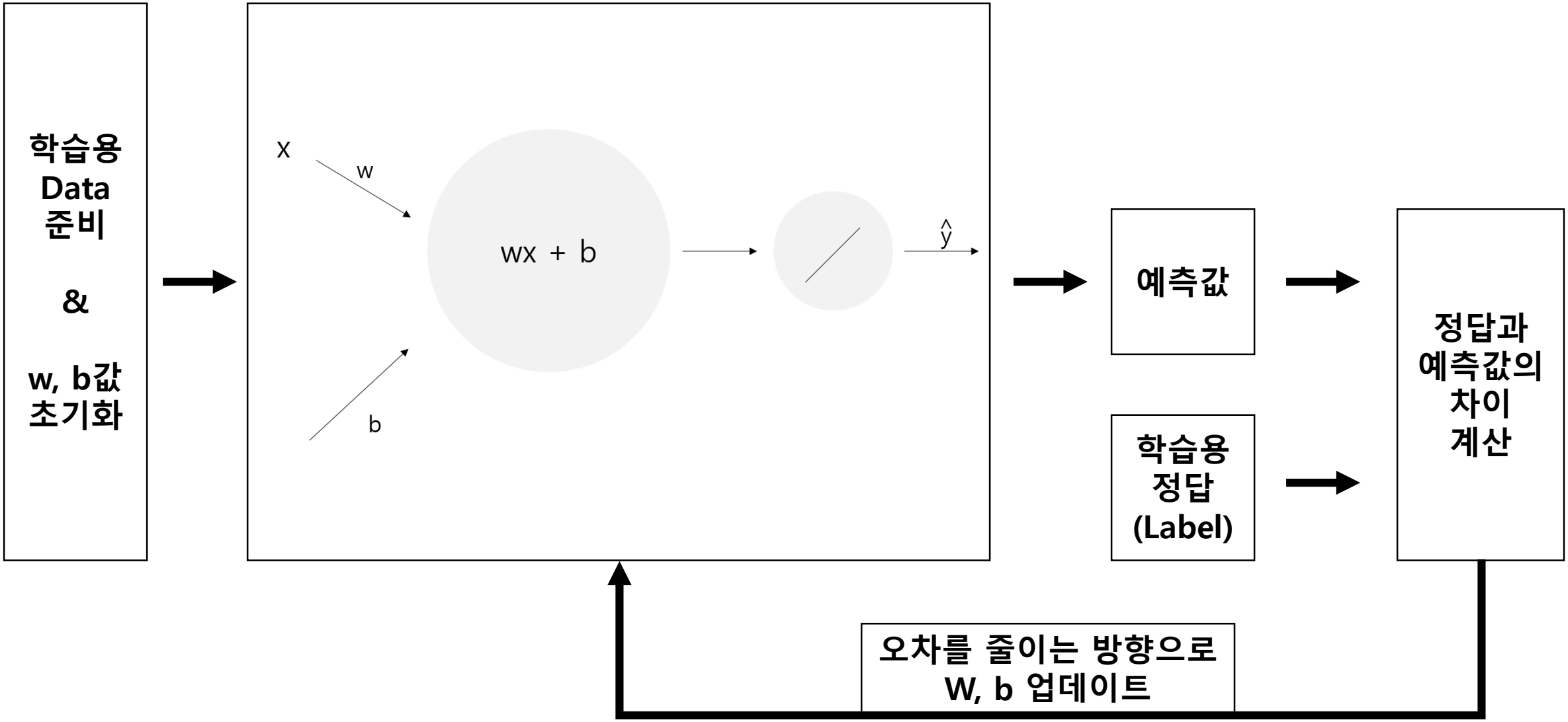
예제) 딥러닝으로 당뇨 수치 예측하기



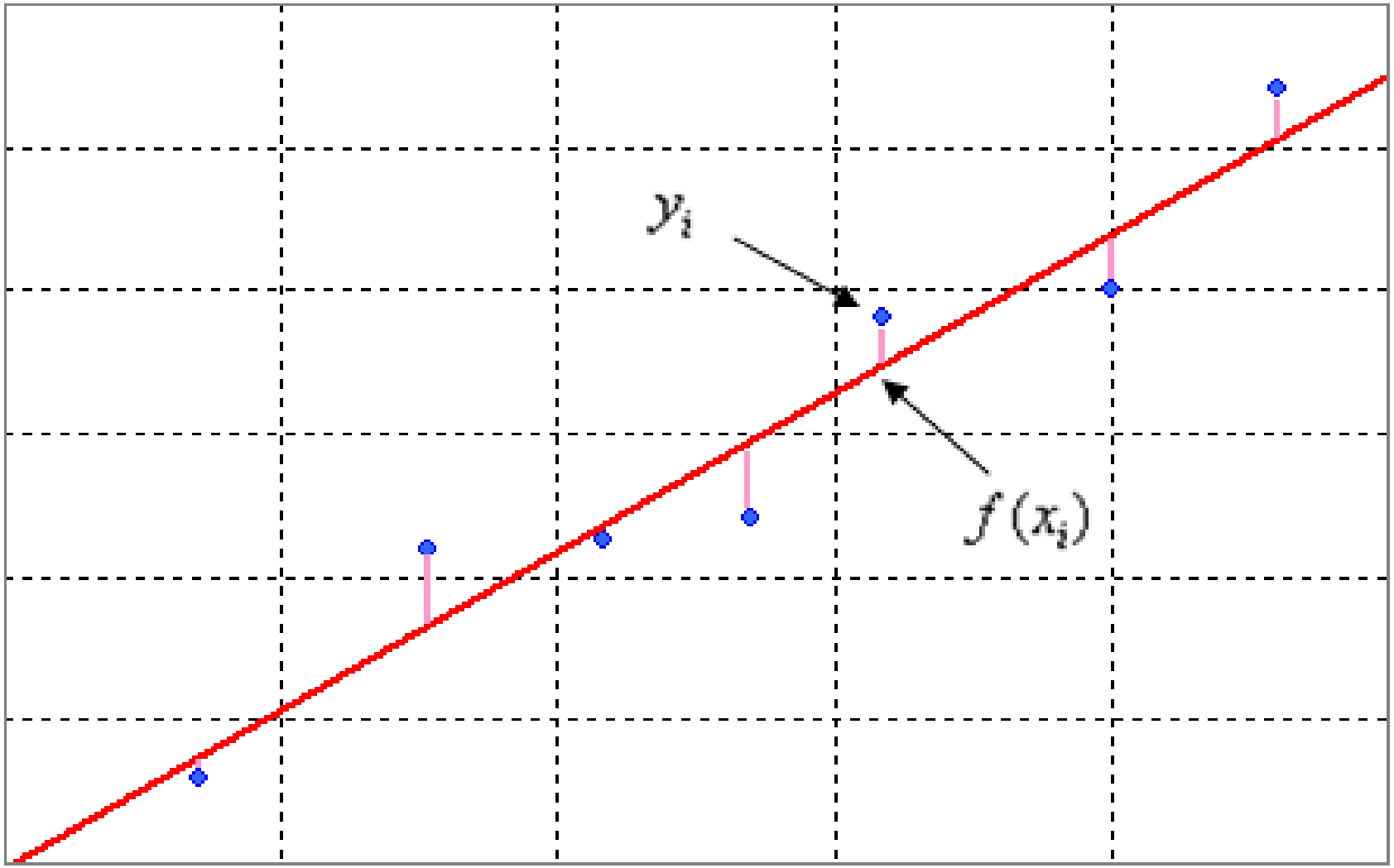
Aritificial Neuron



전체적인 학습 과정(딥러닝은 아님)



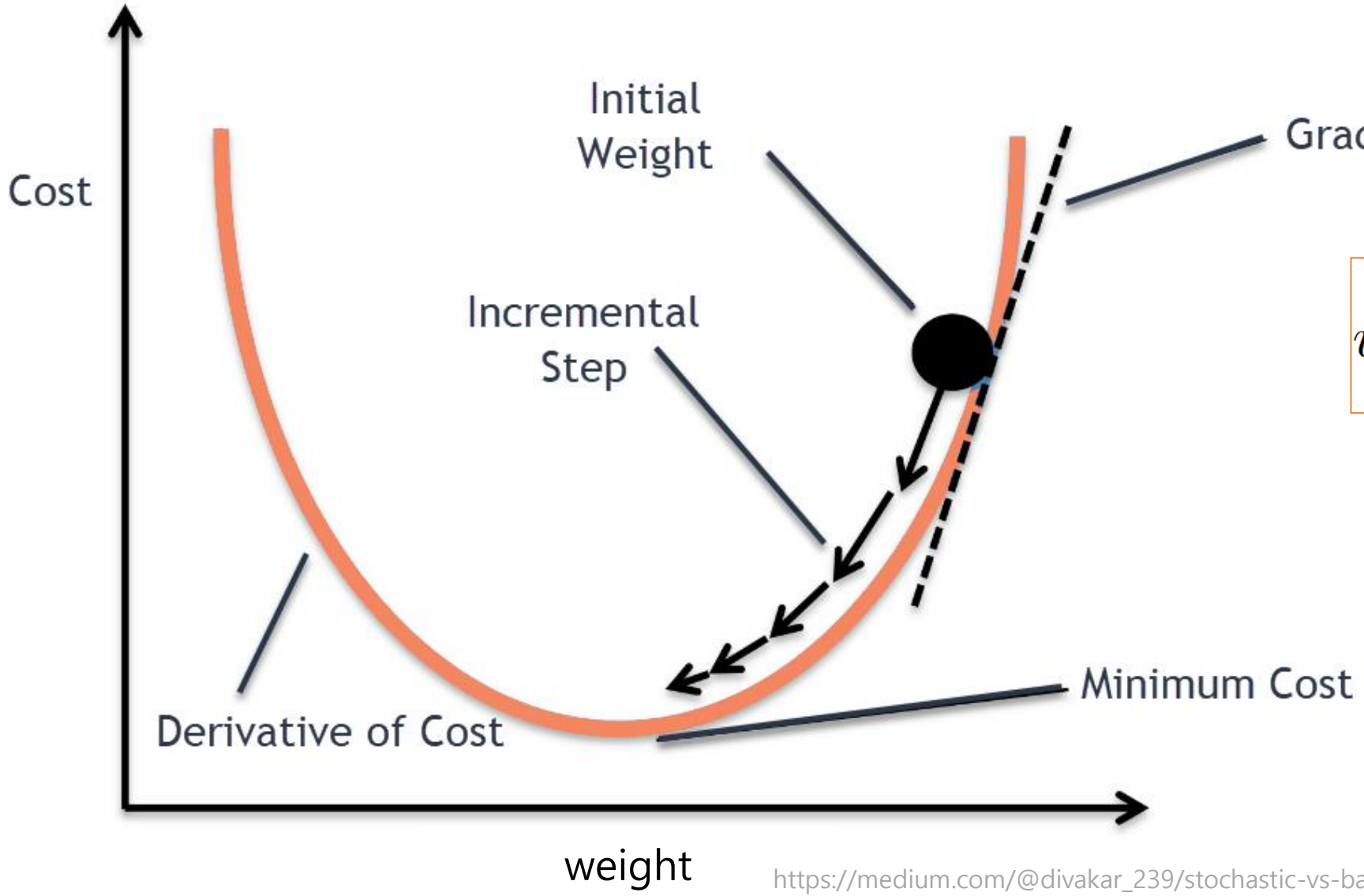
선형 모델(w, b) 찾기 – 최소 제곱법(Squared Error)



선형 모델(w, b) 찾기 – 최소 제곱법(Squared Error)

$$
\begin{aligned}
SE &= (y - \hat{y})^2 \\
&= (y - (wx + b))^2 \\
&= (y - wx - b)^2
\end{aligned}
$$

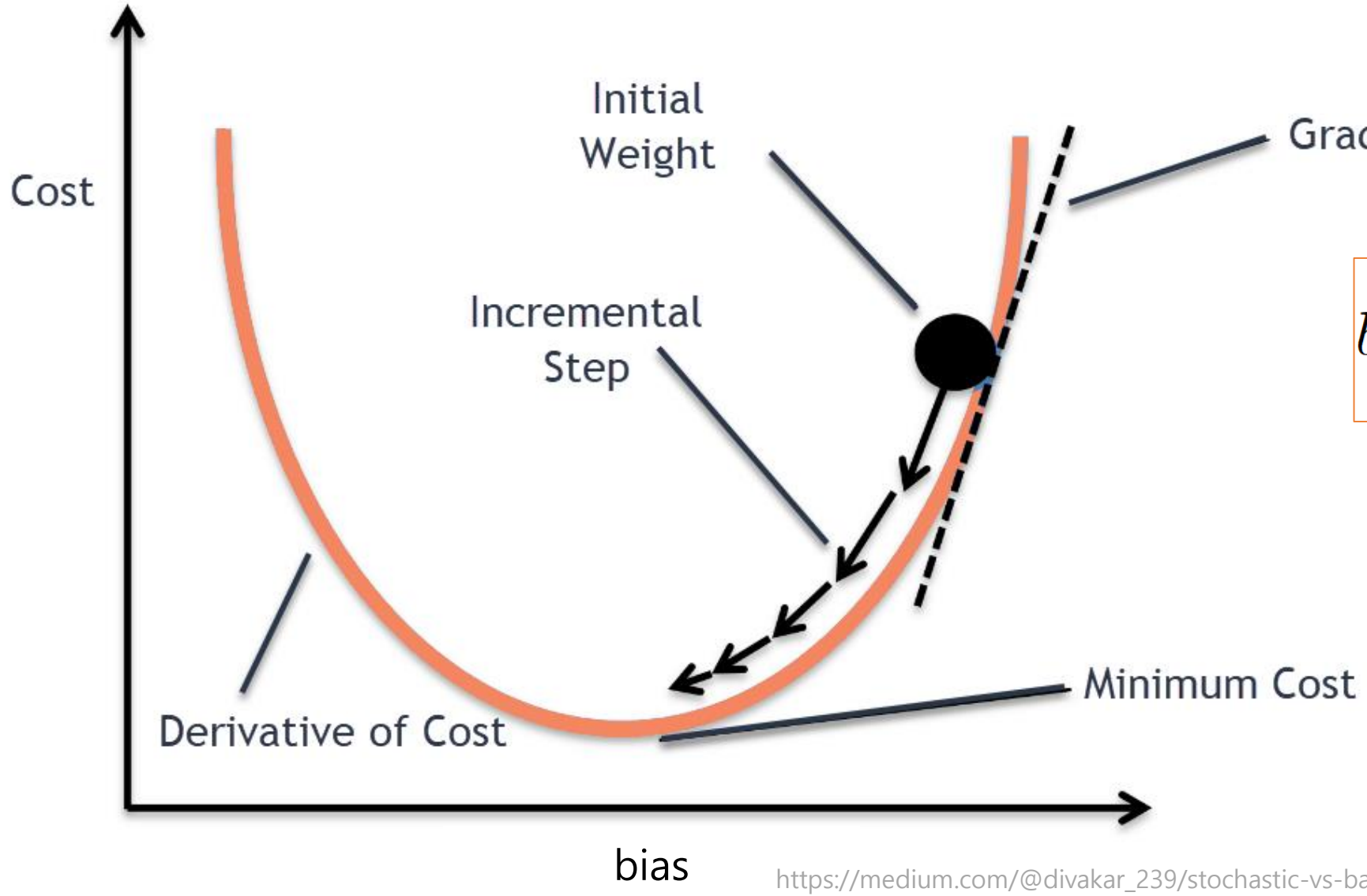
딥러닝의 주요 아이디어 – 경사 하강법



$$w = w - \frac{\partial SE}{\partial w}$$

https://medium.com/@divakar_239/stochastic-vs-batch-gradient-descent-8820568eada1

딥러닝의 주요 아이디어 – 경사 하강법



$$b = b - \frac{\partial SE}{\partial b}$$

https://medium.com/@divakar_239/stochastic-vs-batch-gradient-descent-8820568eada1

선형 회귀 모델(wx+b) 찾기 – w 편미분하기

$$SE = (y - \hat{y})^2$$

$$SE = \frac{1}{2}(y - \hat{y})^2$$

$$w = w - \frac{\partial SE}{\partial w}$$

$$w = w - \frac{\partial}{\partial w}\frac{1}{2}(y - \hat{y})^2$$

$$w = w - \frac{1}{2} \times 2(y - \hat{y})\frac{\partial}{\partial w}(y - \hat{y})$$

$$w = w - (y - \hat{y})\frac{\partial}{\partial w}(y - wx - b)$$

$$w = w - (y - \hat{y}) \times -x$$

$$w = w + (y - \hat{y}) \times x$$

선형 회귀 모델(wx+b) 찾기

$$b = b - \frac{\partial SE}{\partial b}$$

$$b = b - (y - \hat{y}) \frac{\partial}{\partial b}(y - wx - b)$$

$$b = b - (y - \hat{y}) \times -1$$

$$b = b + (y - \hat{y})$$

선형 회귀 모델(wx+b) 찾기 – w, b 업데이트 식

$$w = w + (y - \hat{y}) \times x$$

$$b = b + (y - \hat{y})$$

05

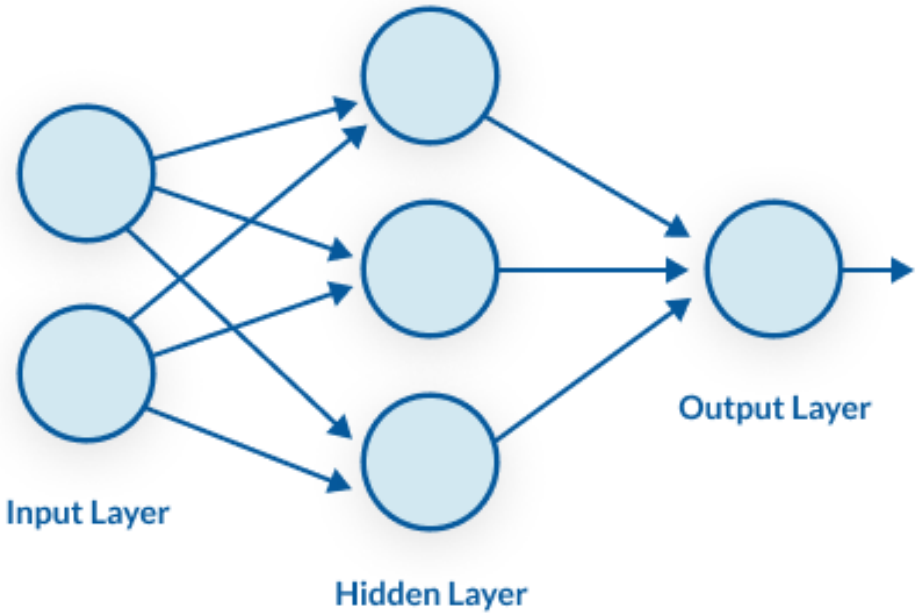
딥러닝의 학습 과정과 학습 준비

딥러닝(지도학습) 의 전체적인 학습 과정

학습용
Data
준비

&

w, b값
초기화



예측값

학습용
정답
(Label)

정답과
예측값의
차이
계산

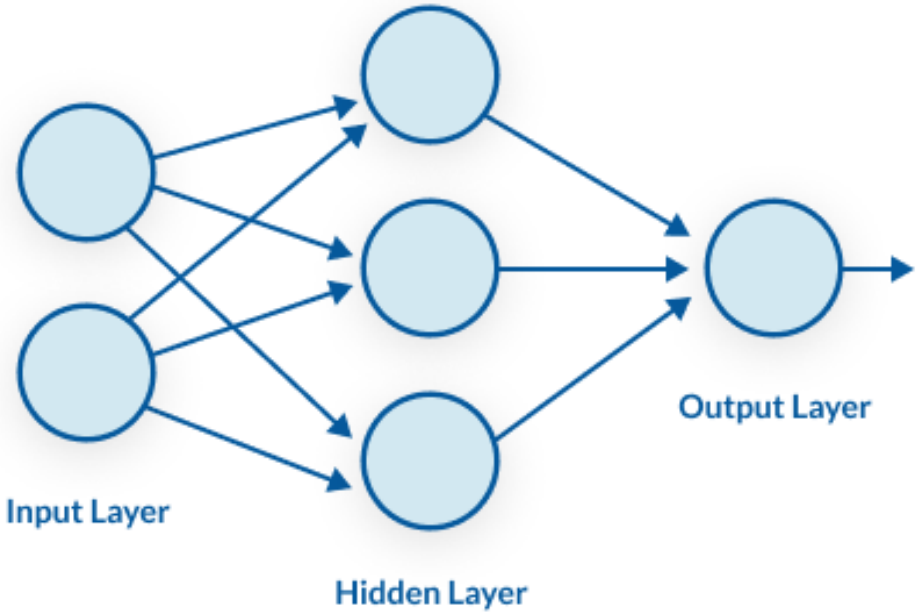
오차를 줄이는 방향으로
W, b 업데이트

학습을 위한 준비(1) – 데이터 준비(학습/테스트)

학습용 Data 준비

&

w, b값 초기화



예측값

학습용 정답 (Label)

정답과 예측값의 차이 계산

오차를 줄이는 방향으로 W, b 업데이트

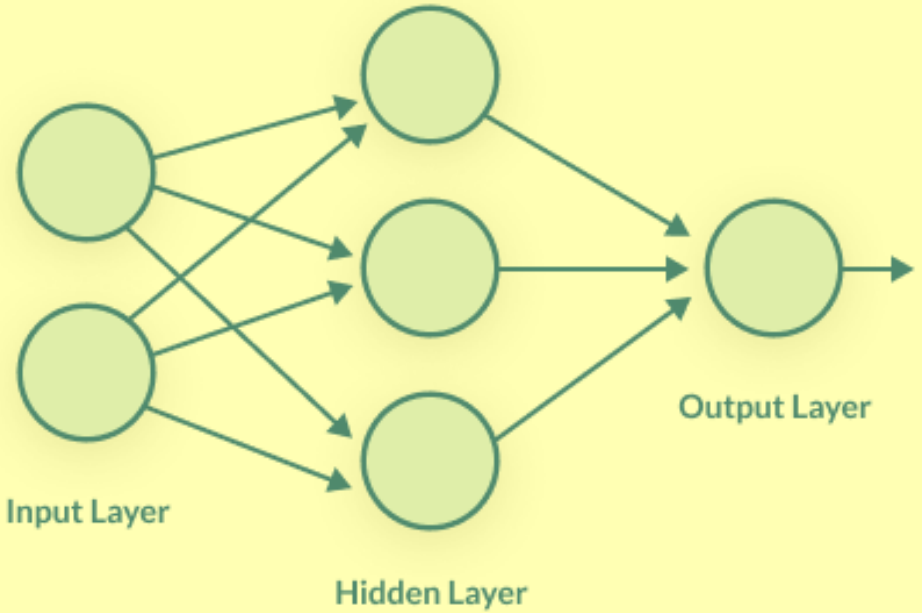
학습을 위한 준비(2) – 네트워크 구성

학습용 Data 준비

&

w, b값 초기화

Network를 어떻게 구성할 것인가?
(몇 층, 층별 노드 수, 층별 활성화함수 등)



예측값

학습용 정답 (Label)

정답과 예측값의 차이 계산

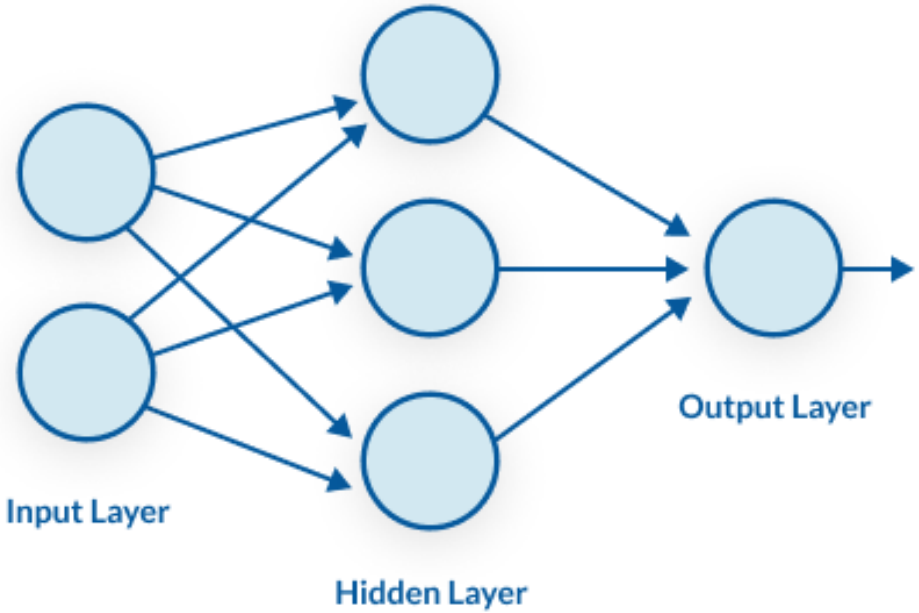
오차를 줄이는 방향으로 W, b 업데이트

학습을 위한 준비(3) – 손실함수

학습용
Data
준비

&

w, b값
초기화



예측값

학습용
정답
(Label)

정답과
예측값의
차이
계산

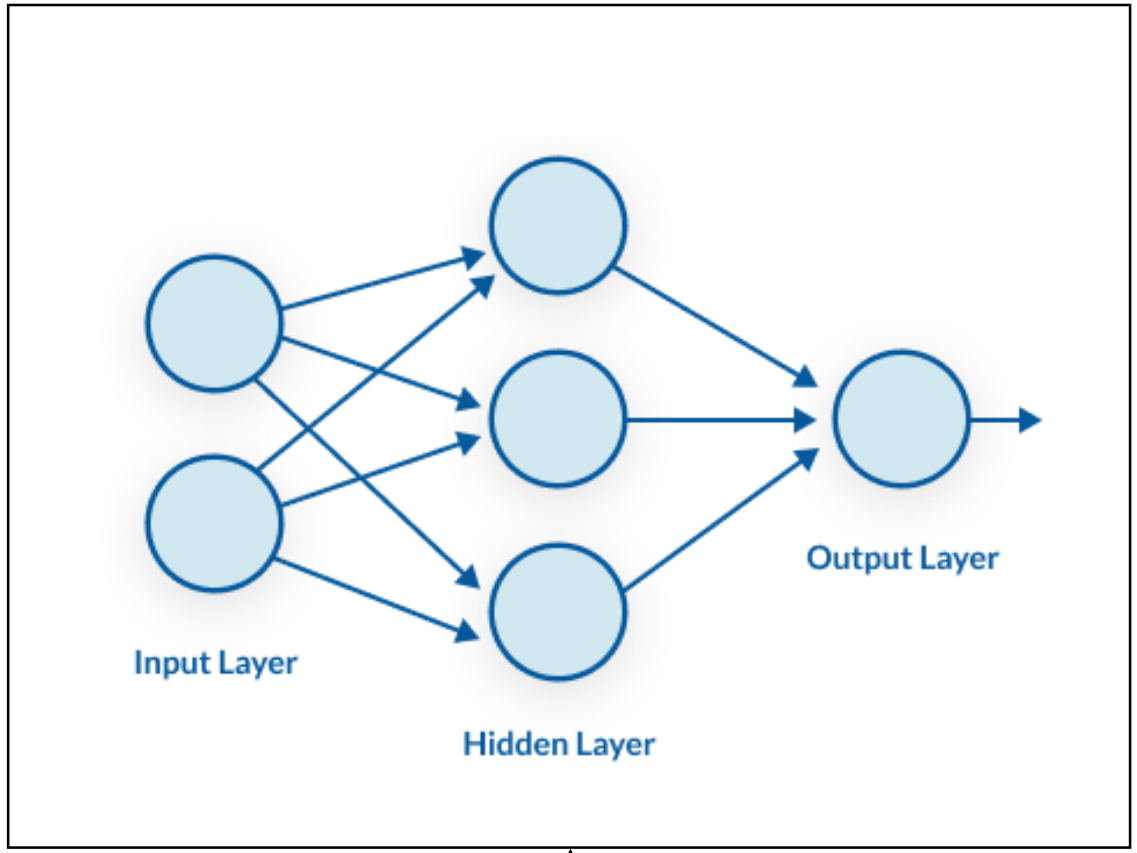
오차를 줄이는 방향으로
W, b 업데이트

학습을 위한 준비(4) – 옵티마이저

학습용
Data
준비

&

w, b값
초기화



예측값

학습용
정답
(Label)

정답과
예측값의
차이
계산

오차를 줄이는 방향으로
W, b 업데이트

06

딥러닝 들어가기 프로젝트(MNIST)

MNIST(Modified National Institute of Standards and Technology)

- 손으로 쓴 숫자들로 이루어진 대형 데이터베이스
- NIST(미국 국립표준기술연구소)의 오리지널 데이터셋의 샘플을 재혼합하여 만들어짐

- 28x28 픽셀의 그레이스케일 이미지
- 60,000개의 트레이닝 이미지+10,000개의 테스트 이미지

- 출처 : 위키피디아

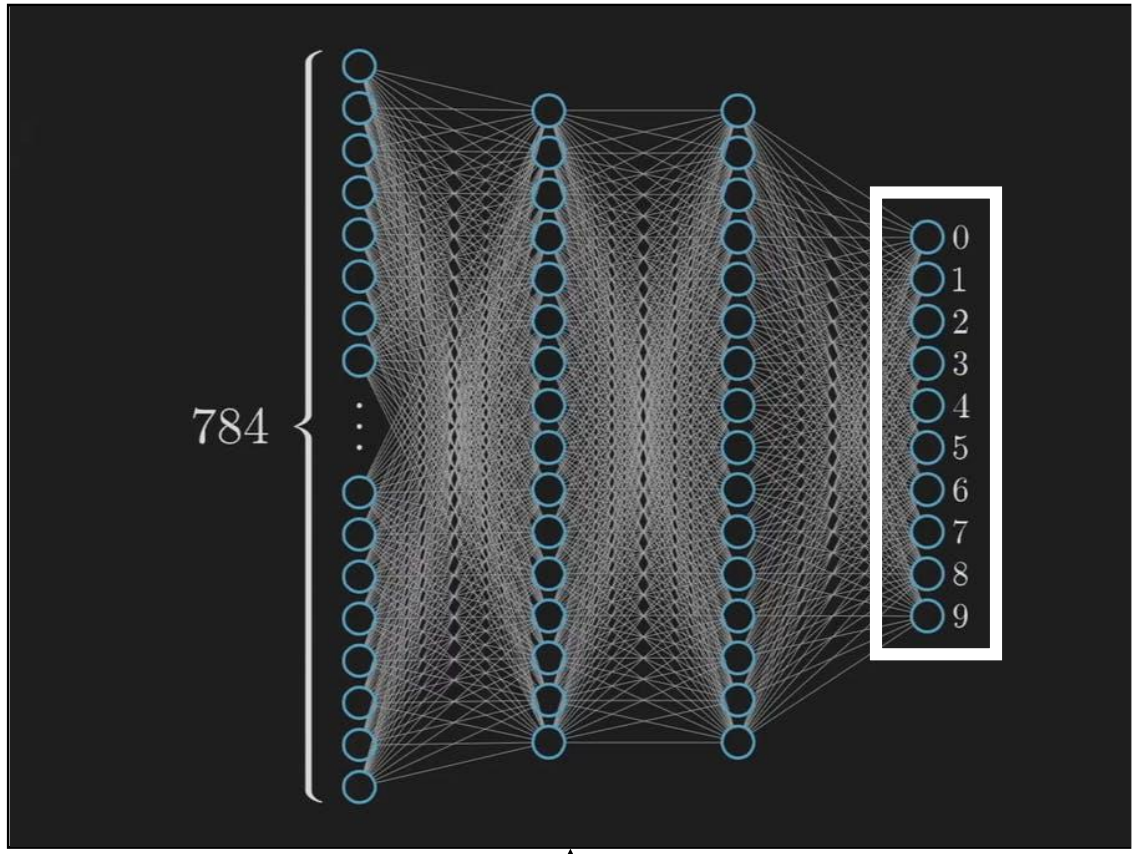
MNIST 테스트 데이터셋의 샘플 이미지.

딥러닝의 전체적인 학습 과정(MNIST 이미지 분류)

학습용
Data
준비
(28x28)

&

w, b값
초기화



예측값

학습용
정답
(Label)

손실
함수
(cross
entropy)

옵티마이저
(sgd, adam)