

# 딥러닝팀

## 1팀

김예찬

윤지영

채소연

한지원

홍지우

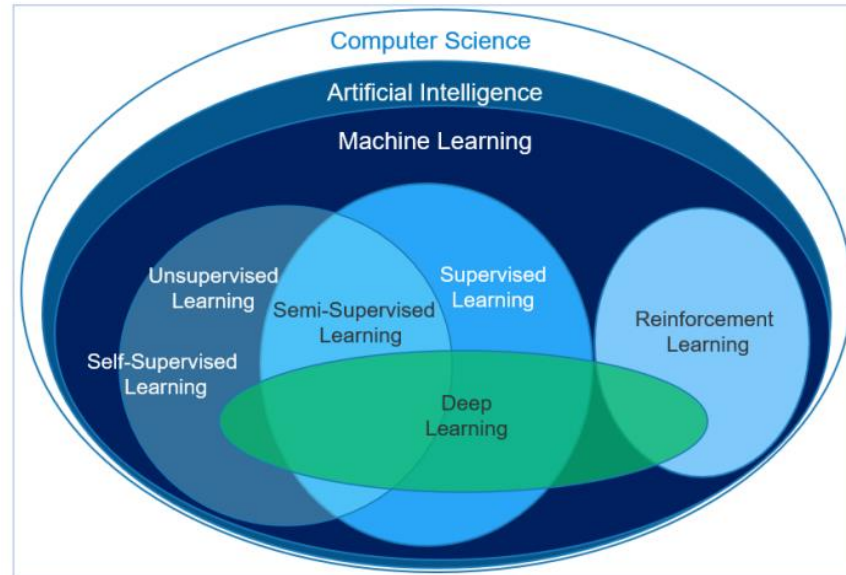
1

머신러닝

# 1 머신러닝(Machine Learning)

- 머신러닝

## 머신러닝(Machine Learning)이란?

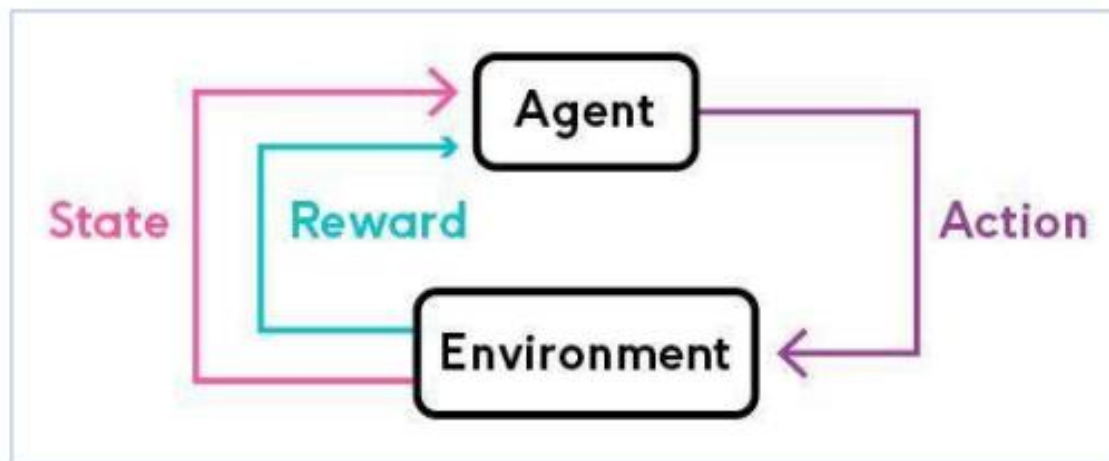


머신러닝은 **고전적 머신러닝 (지도 / 비지도 학습)**과 **강화학습**으로 나눌 수 있으며,  
딥러닝은 이 모두에 적용될 수 있음

# 1 머신러닝(Machine Learning)

- 강화학습

## 강화학습(Reinforcement Learning)



### 고전적 머신러닝 알고리즘

입력과 출력을 통해  
데이터의 **속성 파악**, 라벨을 **예측**

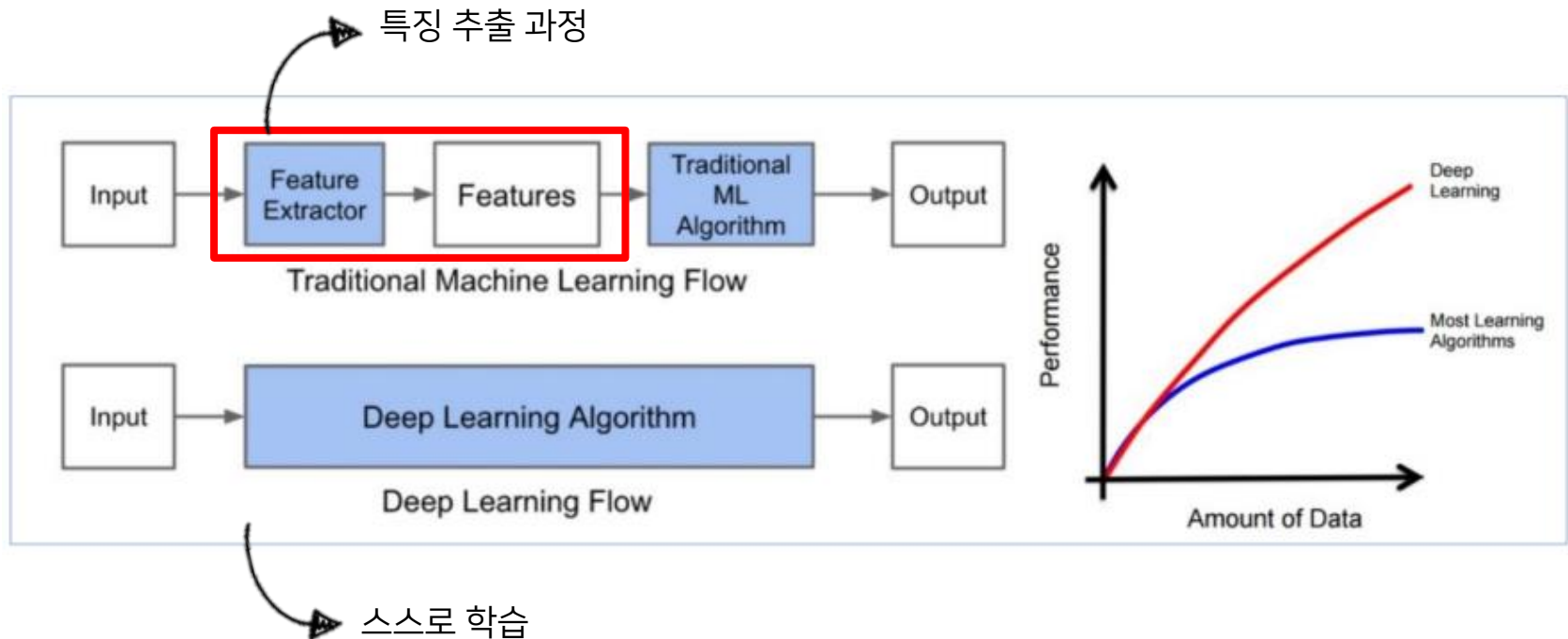
### 강화학습

(X,Y)의 형태는 물론이고  
입력과 출력이라는 개념조차 없음

# 1 머신러닝(Machine Learning)

## ● 딥러닝

### 딥러닝(Deep Learning)



- 데이터의 크기와 형태가 커질수록 더 좋은 성능 - 빅데이터 분석에 유용
- 지도학습, 비지도학습, 강화학습의 과제 모두에 적용될 수 있음

2

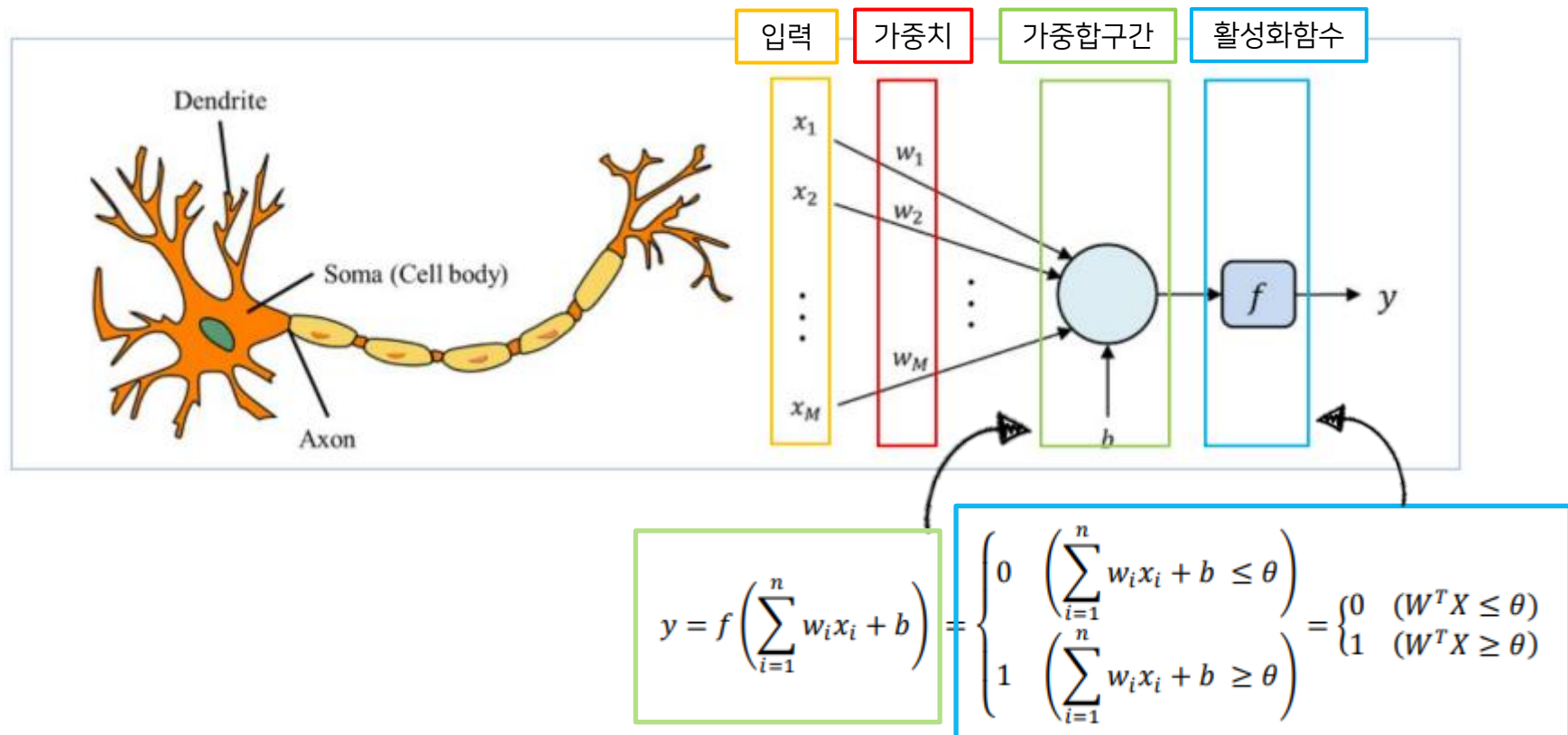
퍼 셉 트 론

## 2 퍼셉트론(Perceptron)

- 퍼셉트론

### 퍼셉트론이란?

: 다수의 입력 데이터에 대해 하나의 출력을 반환하는 형태

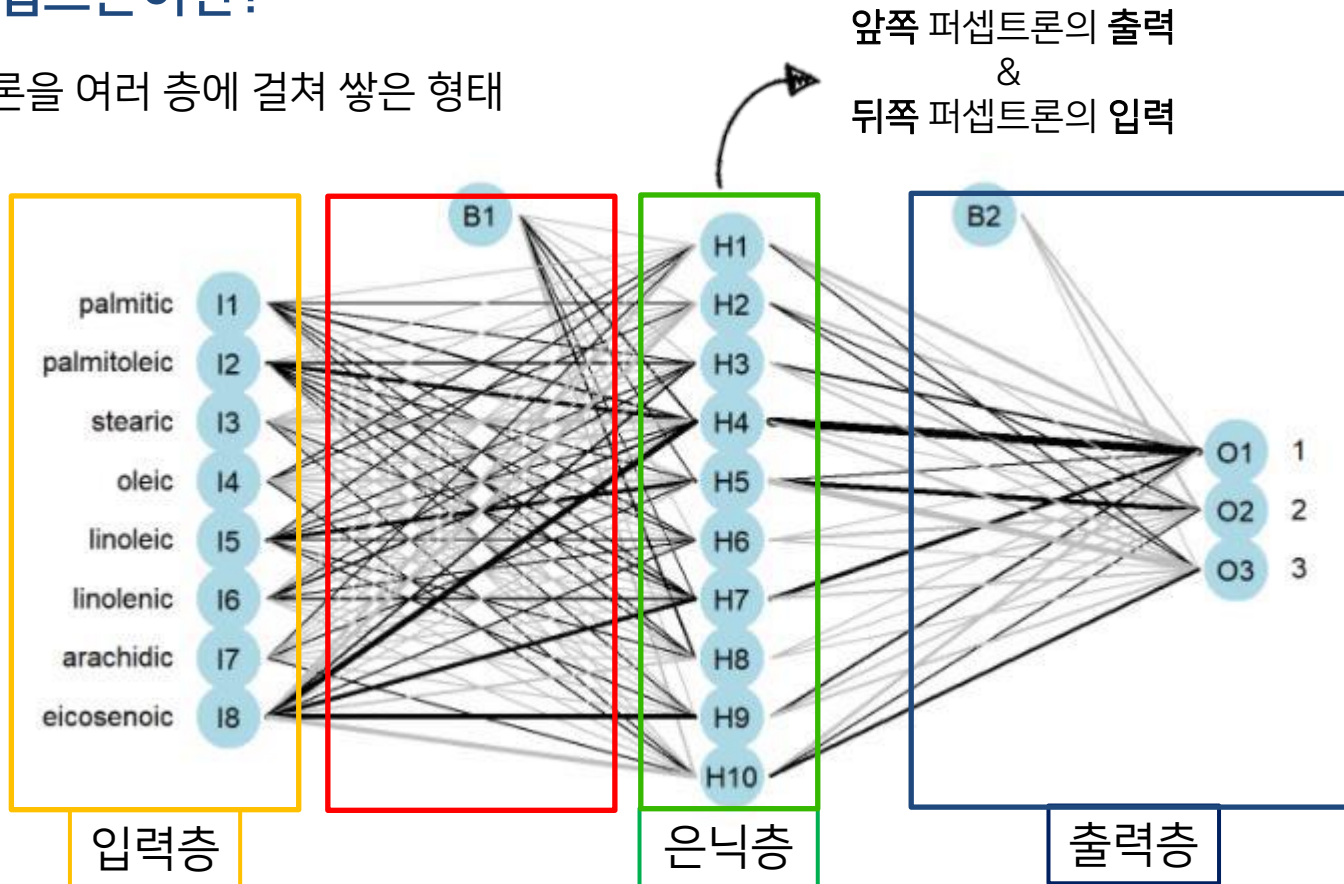


## 2 퍼셉트론(Perceptron)

- 다층 퍼셉트론

### 다층 퍼셉트론이란?

: 퍼셉트론을 여러 층에 걸쳐 쌓은 형태





# 3

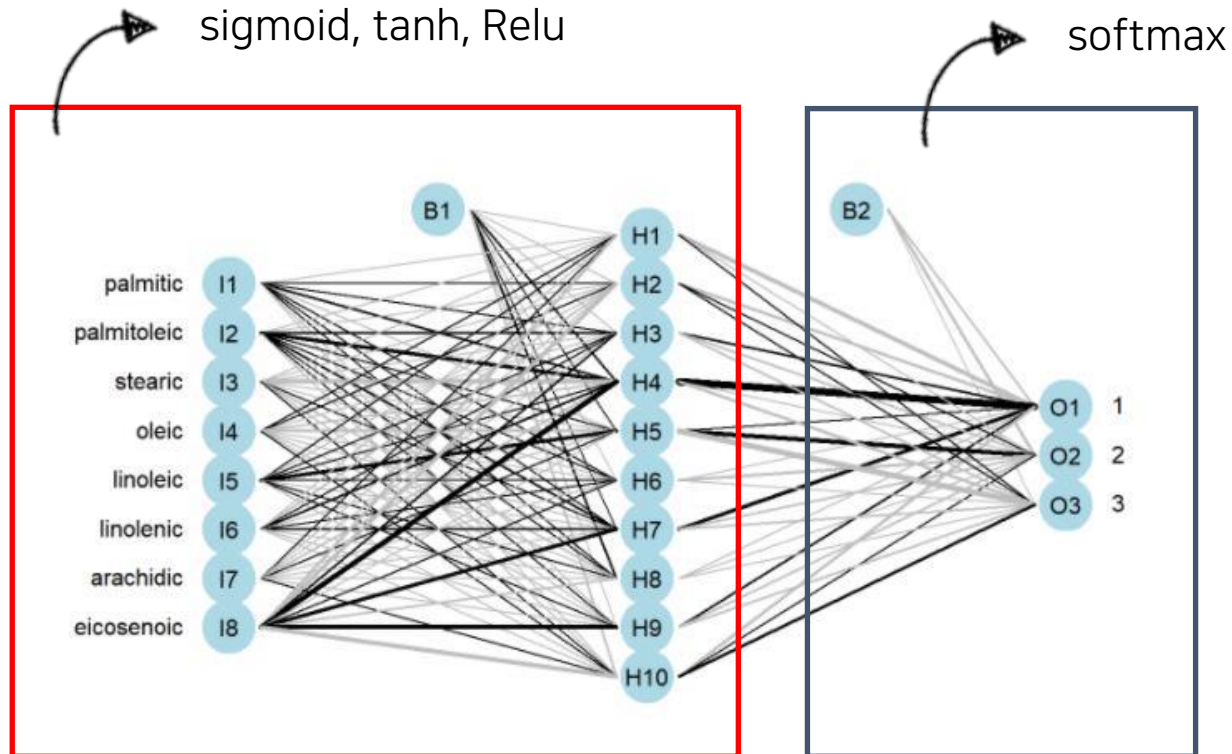
신 경 망

### 3 신경망 (Neural Network)

- 활성화 함수

#### 활성화 함수(Activation Function)란?

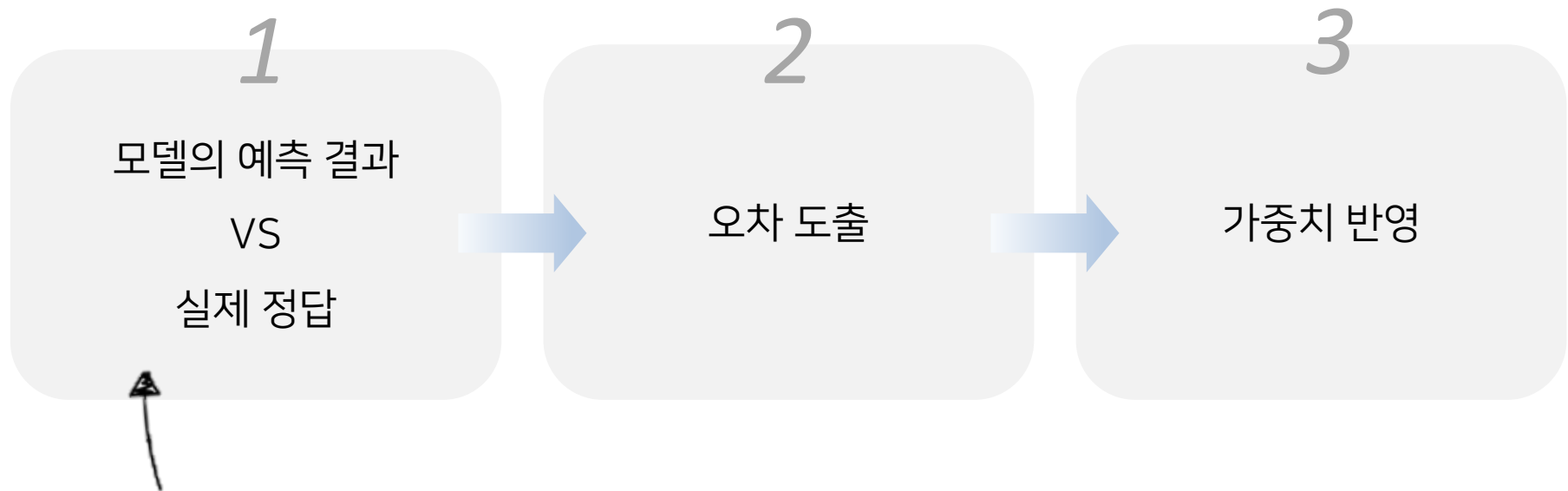
: 입력과 가중치의 선형결합에 비선형성을 부여해주는 함수



### 3 신경망 (Neural Network)

- 손실 함수

손실 함수(Loss Function)란?

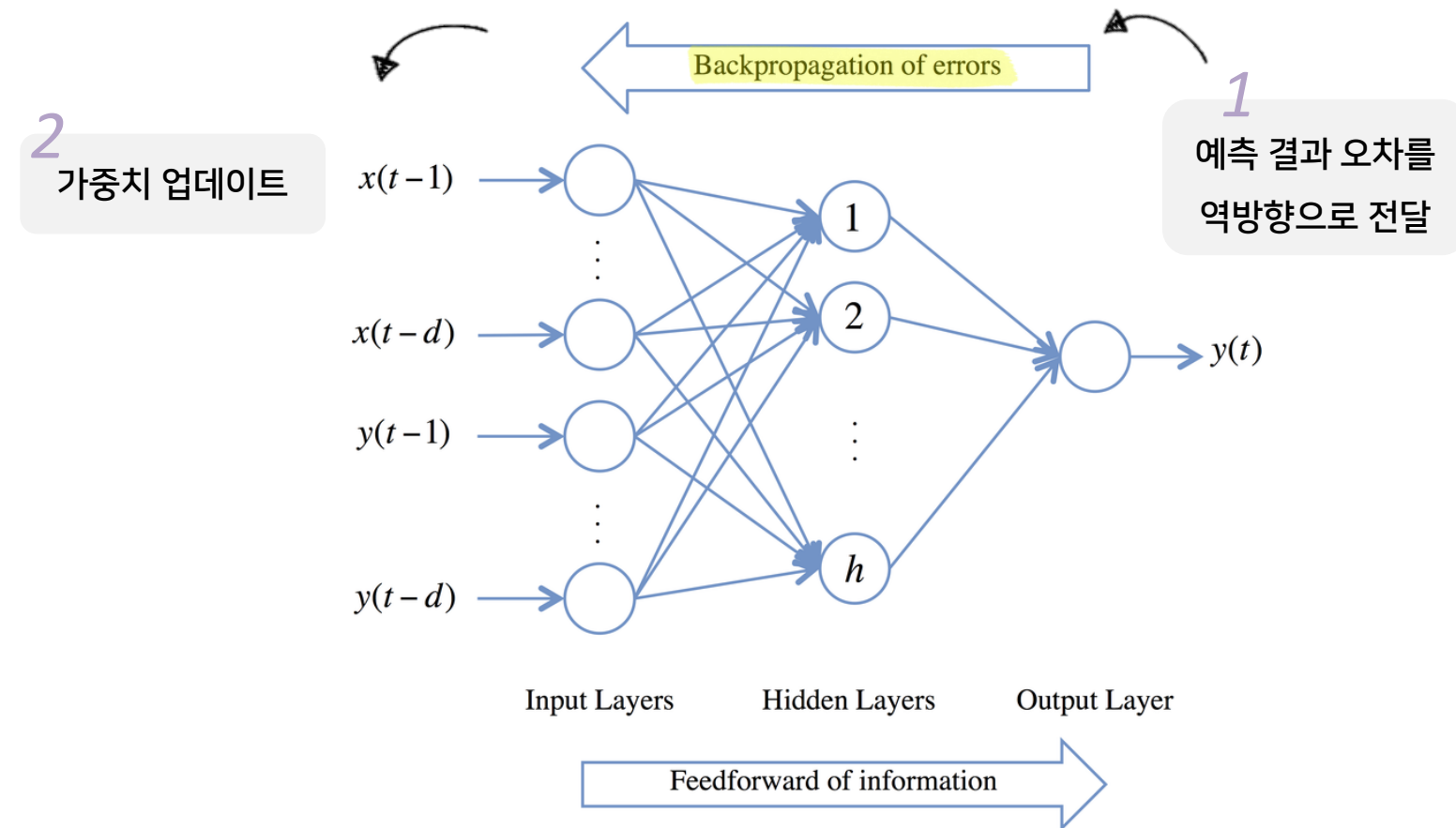


- 역전파를 위한 필수적인 단계

### 3 신경망 (Neural Network)

- 역전파

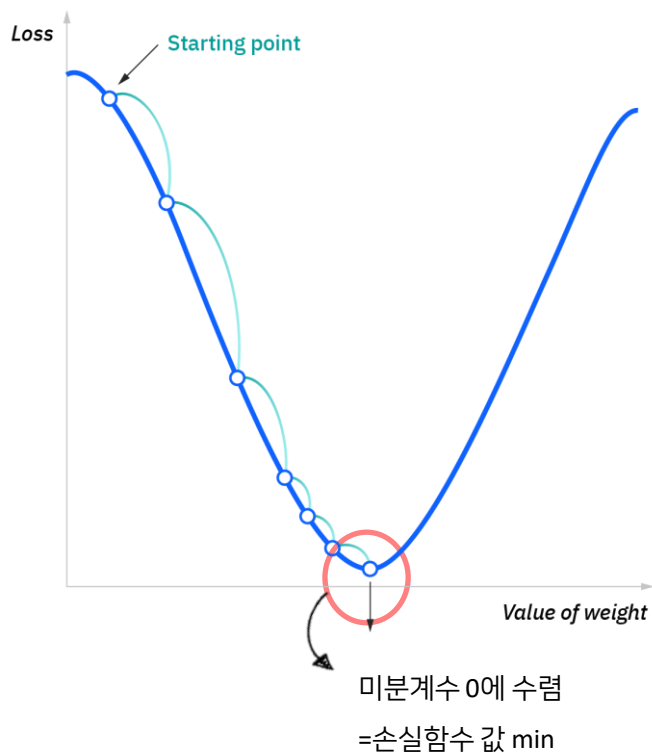
#### 진행 과정



# 3 신경망 (Neural Network)

## Optimizer

### 경사 하강법(Gradient Descent)



- 기울기를 작게 만들어 나가는 형태
- 미분계수 부호의 반대 방향으로 이동  
→ 최솟값

학습률  $\alpha$

$$x_{i+1} = x_i - \alpha \frac{df}{dx}(x_i) \quad \text{-가중치 1개}$$

$$\left. \begin{aligned} W &\leftarrow W - \eta \left( \frac{\partial E}{\partial w} \right) \\ b &\leftarrow b - \eta \left( \frac{\partial E}{\partial b} \right) \end{aligned} \right\} \quad \text{- 각각의 가중치 업데이트}$$

# 3 신경망 (Neural Network)

- Optimizer

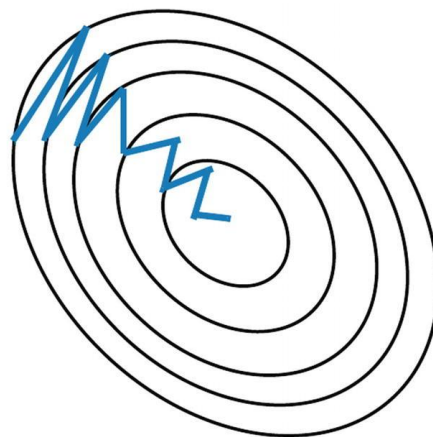
## Momentum



Stochastic Gradient  
Descent **without**  
Momentum

기존 optimizer

미분값에 따라 한 단계씩  
최적점으로 접근



Stochastic Gradient  
Descent **with**  
Momentum

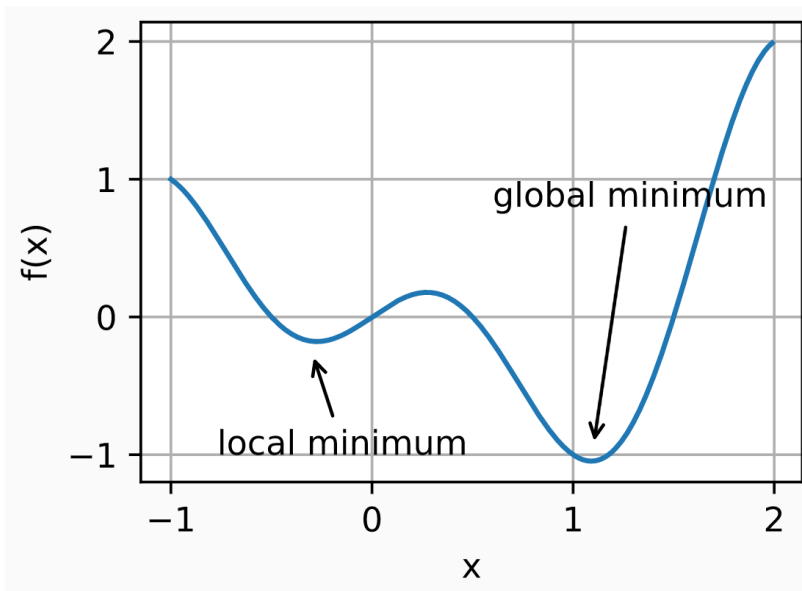
Momentum

미분값이 클 경우,  
가속도 부여(큰 보폭으로 이동)

### 3 신경망 (Neural Network)

- Optimizer의 문제점

#### Local Minima 문제



극소와 극대가 여러 곳에서 존재



학습률이 작은 경우

최소 지점(global minimum)이 아닌  
극소(local minimum)에서 고립



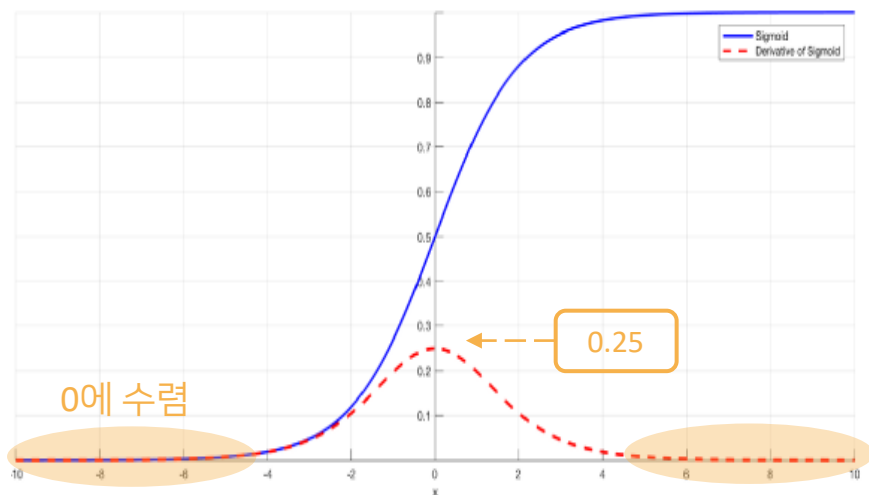
Momentum으로  
해결!

### 3 신경망 (Neural Network)

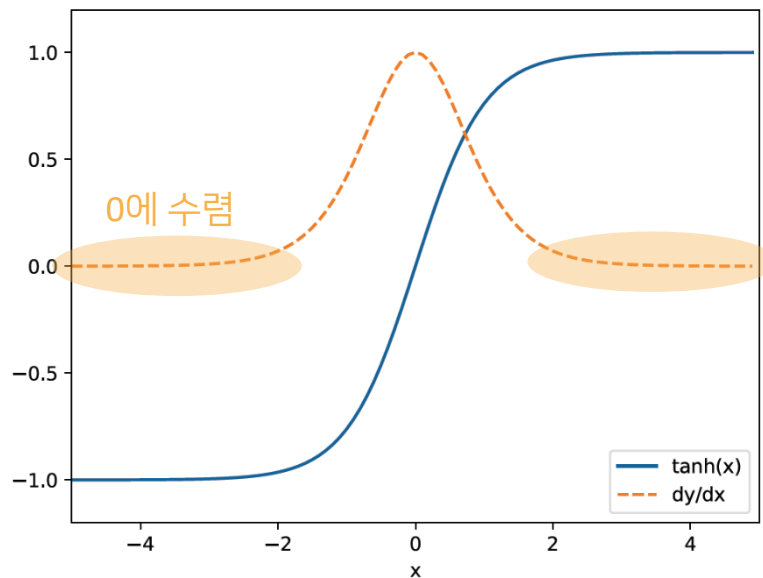
- 기울기 소실 문제 (Gradient Vanishing Problem)

미분계수가 최적점과 현위치의 차이를 적절하게 반영했는가?

시그모이드 함수



tanh 함수



함수를 통과할 때마다 기울기에  
1보다 작은 수를 곱하게 되므로 0에 수렴

정상적인 학습  $x$