AI and Deep Learning

뉴런과 학습

Jeju National University
Yung-Cheol Byun

인간의 고차원 기능은
단순한 뉴런의
수많은 연결로 가능

하지만,
연결만 되었다고 가능?

학습(Learning)

학습이란 무엇인가?

Agenda

- Artificial Intelligence
- Brain and Neurons
- Learning
- Regression
- Deep Neural Networks
- CNN
- RNN
- Unsupervised Learning
- Reinforcement Learning
- AI Applications

Supervised Learning

엄청난 수의 뉴런들,

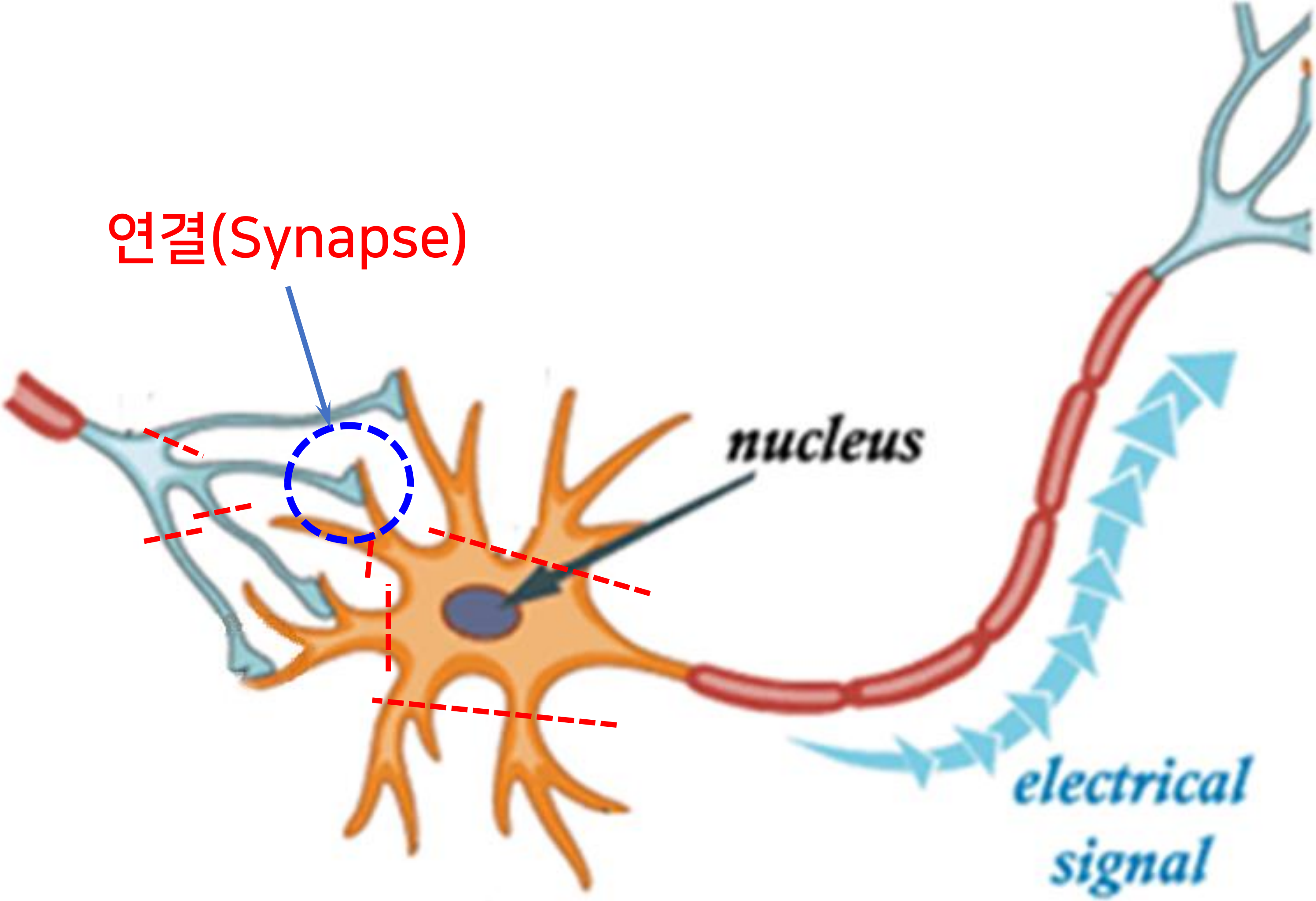
그 뉴런들의 수많은 연결

신기하게도,

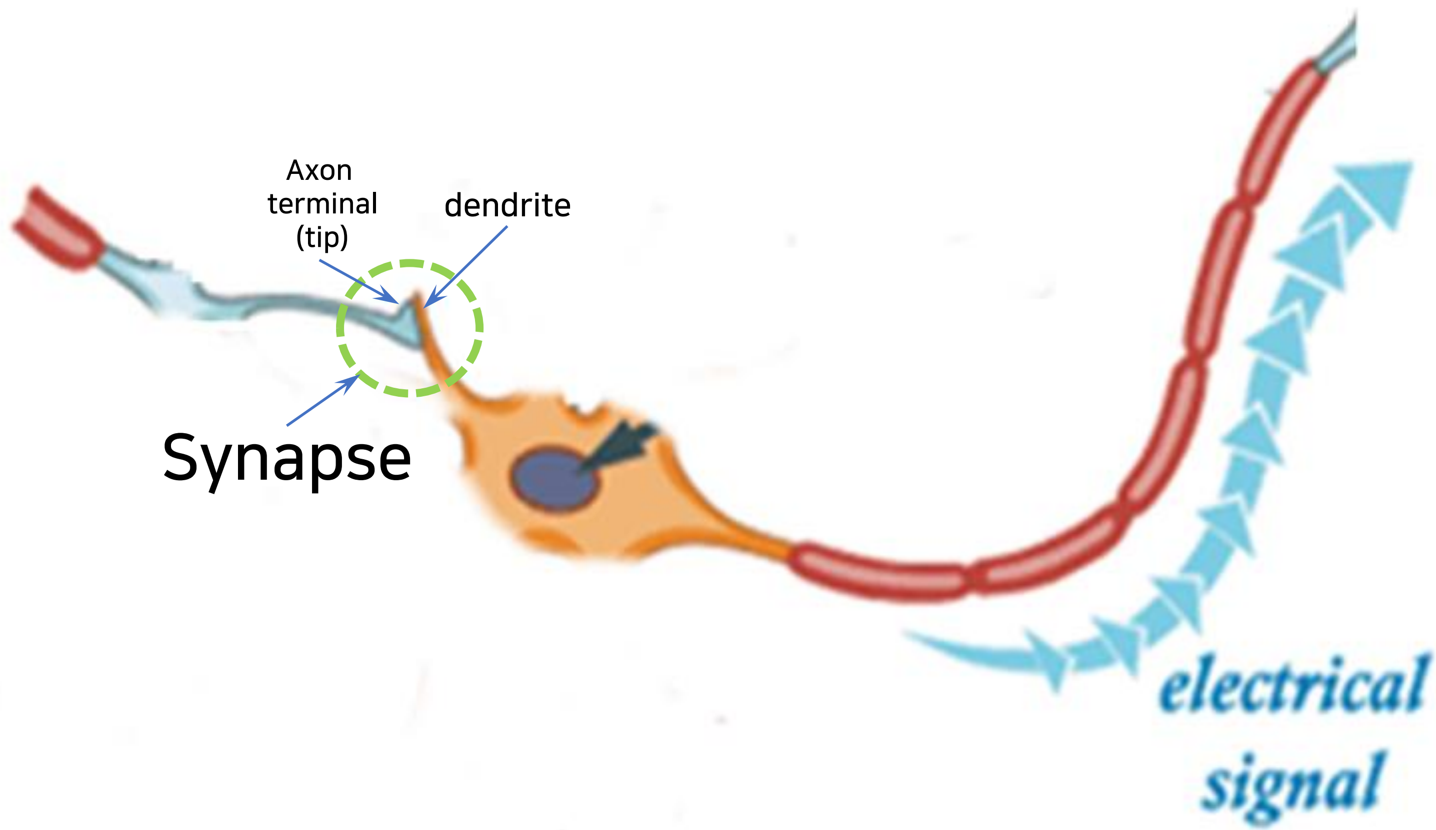
아기는 무엇인가를 경험할 때마다
뉴런 사이의 연결이
'자동으로' 조정된다.

이것이 학습

두 뉴런의 연결

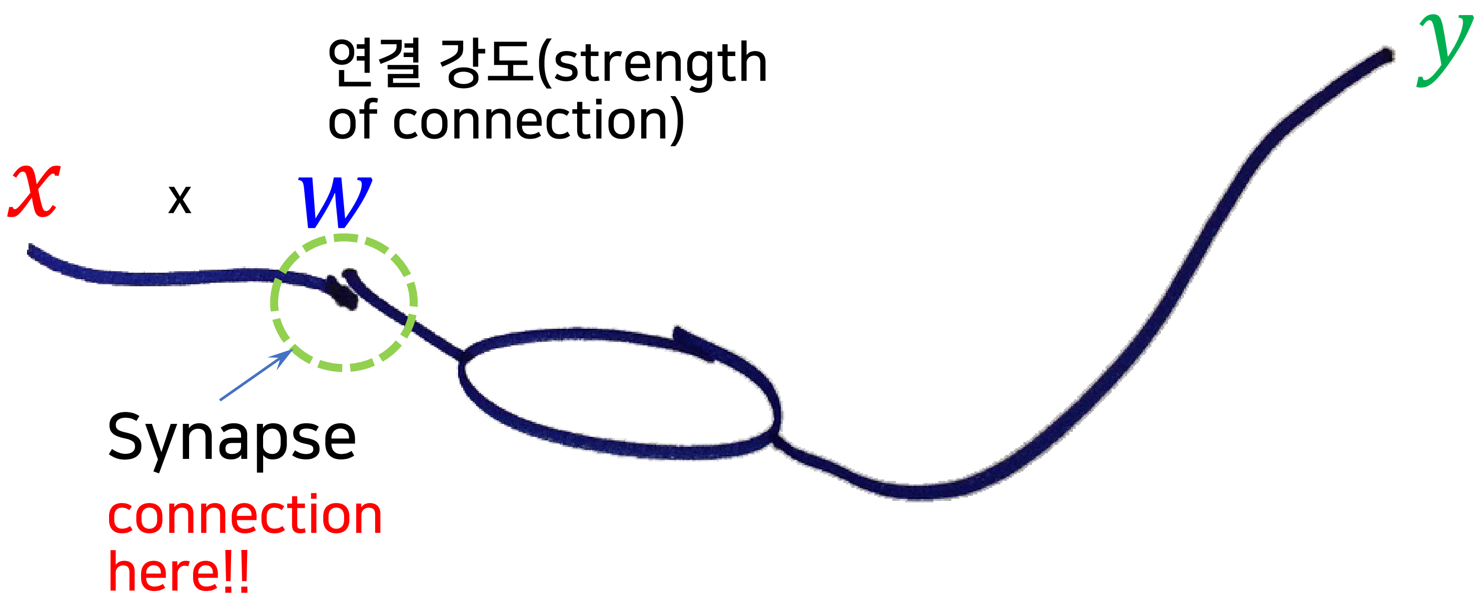


1개 입력을 갖는 뉴런



뉴런의 동작

w: 0, 7, −5 등 임의의 값



$$y = wx$$

뉴런의 동작은 매우 단순
입력(x) * 가중치(w)

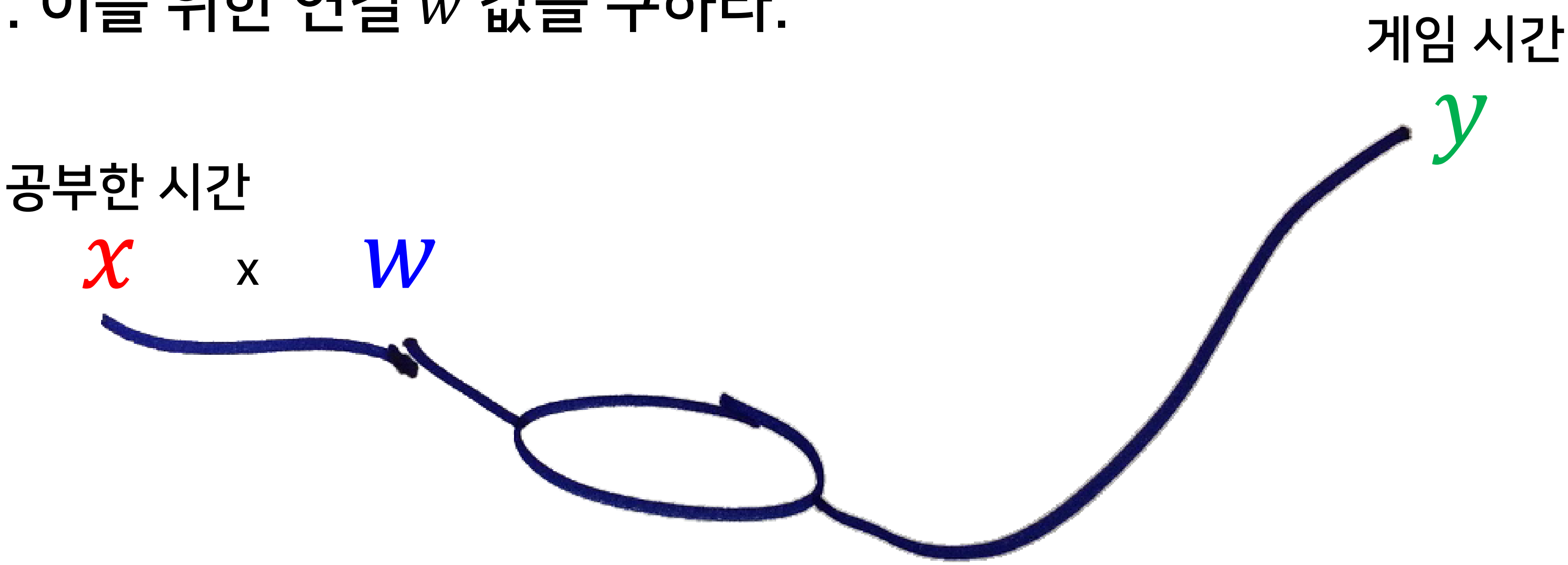
$$y = wx$$

뉴런의 응용

. 1시간(x) 공부하면 1시간(y) 게임하게 해 줄게

. How much you get if study 4 hours? (prediction)

. 이를 위한 연결 w 값을 구하라.



$$y = 1x$$

$$w = 1$$

학습이란?

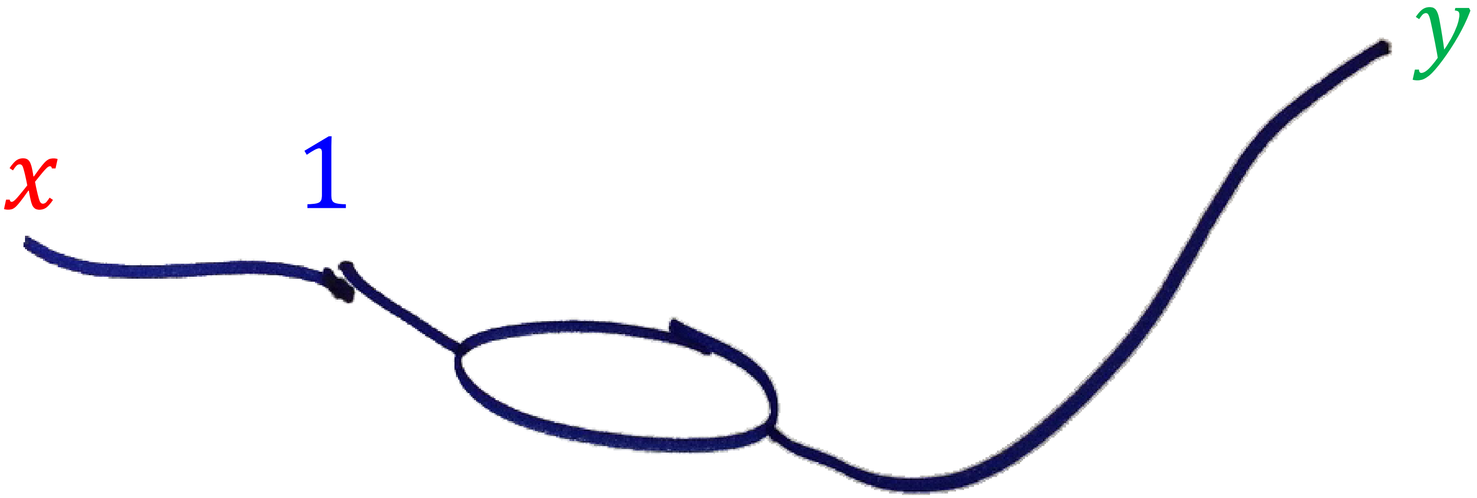
연결 값 w를 조정하는 것

{강하게, 혹은 약하게}

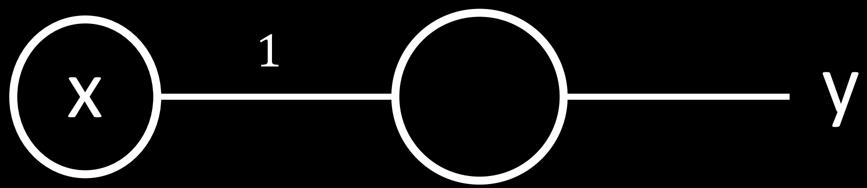
(Q) Draw a neuron

Representing the following equation:

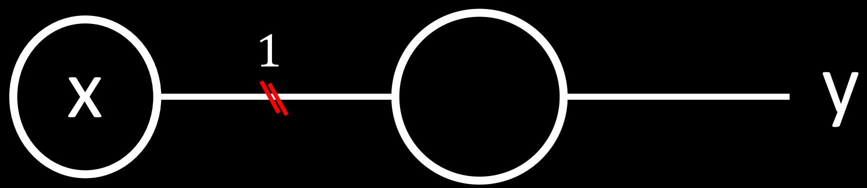
$$y = 1x$$



연결(시냅스)은 어디에 있을까?

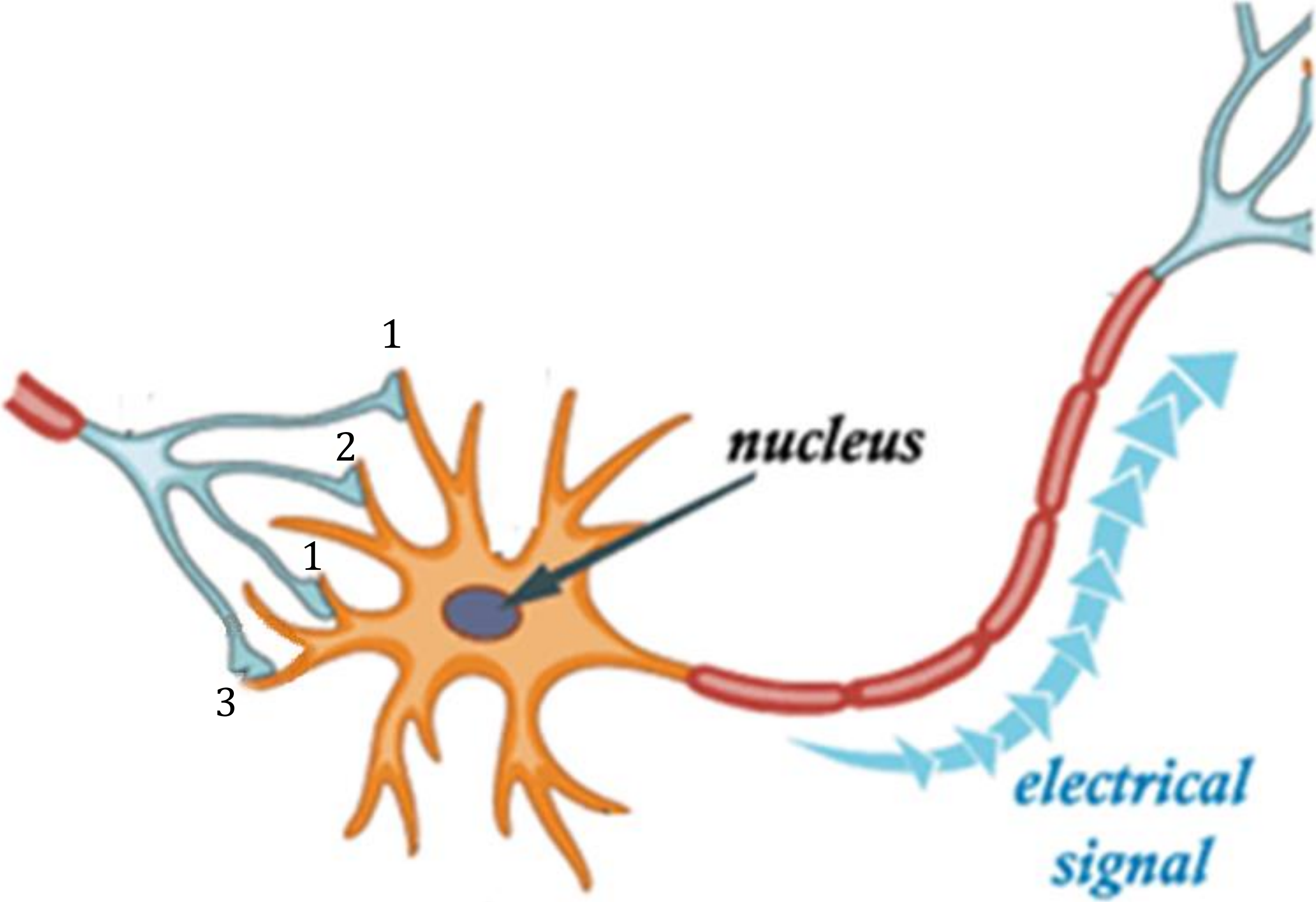


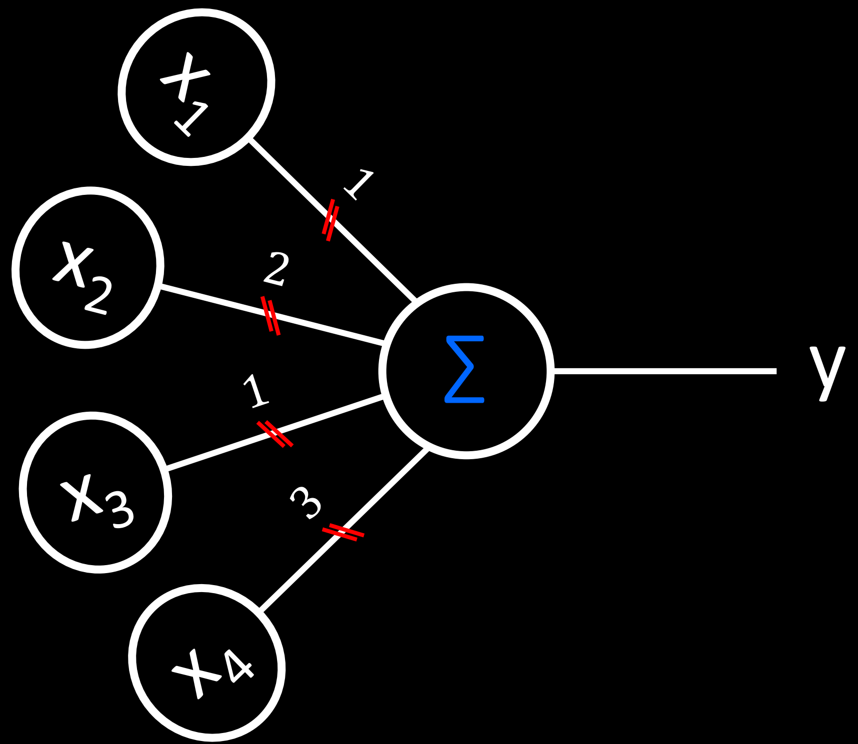
연결(시냅스)은 어디에 있을까?



x
1
y

여러 입력을 갖는 뉴런



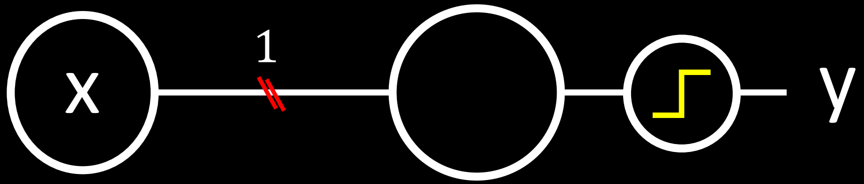


입력에 가중치를 곱하여
모두 더해서 (weighted sum) 출력
(x가 각각 1,1,1,1이면 출력 값은?)

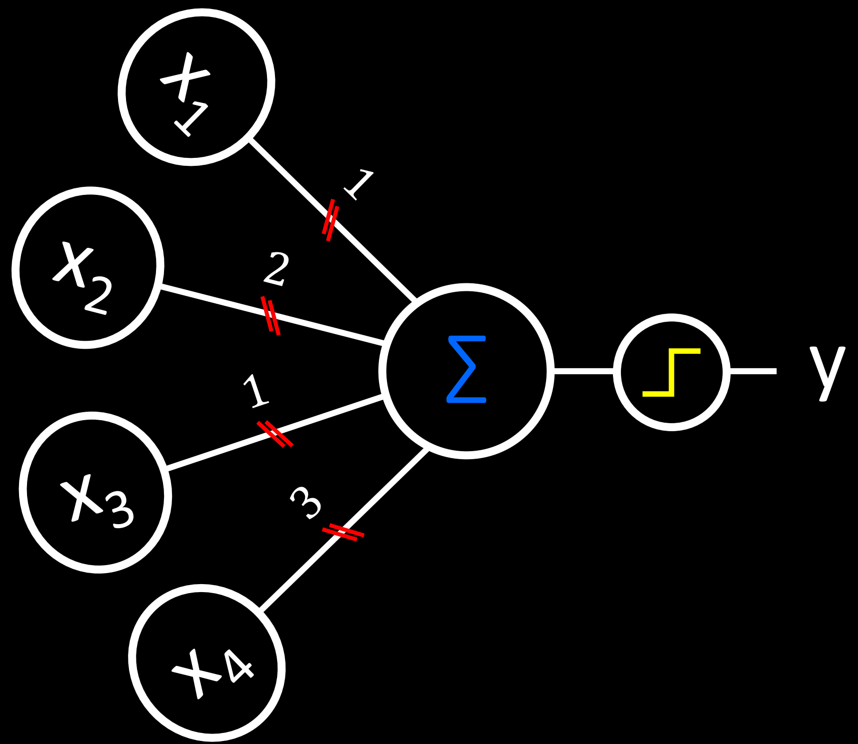
입력의 수만큼 연결이 존재
(Synapses, Weights)

사실은..

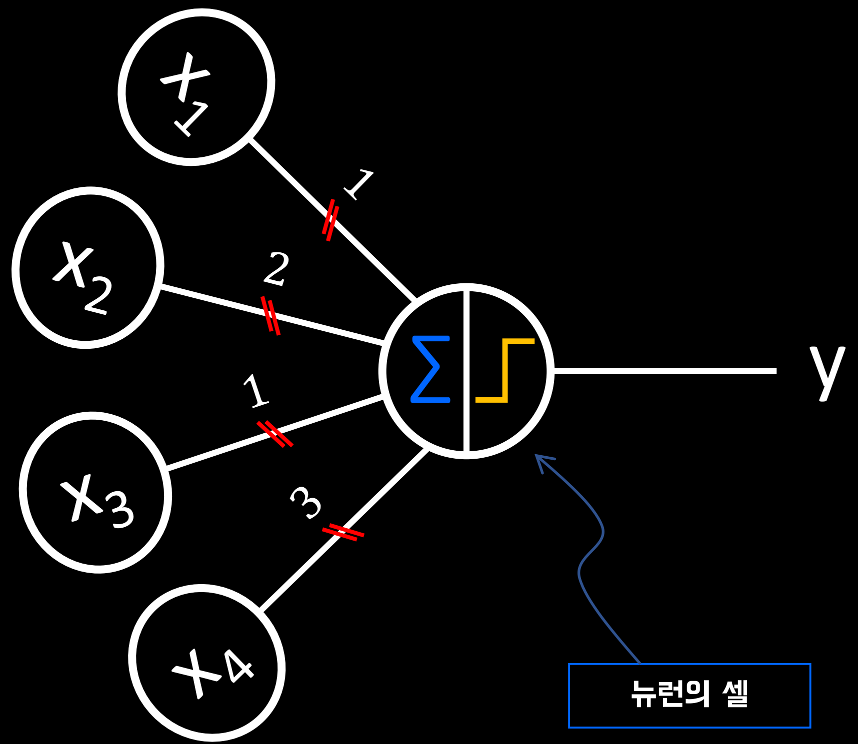
- 뉴런은 모두 더해서(weighted sum) 일정한 값 이상일 때만 시그널 ON
- 그렇지 않으면 시그널 OFF



**특정 값(T) 이상이면 ON(1),
아니면 OFF(0)**



모두 더해서 특정 값(T) 이상이면 ON(1),
아니면 OFF(0)



모두 더해서 특정 값(T) 이상이면 ON(1),
아니면 OFF(0)

다음 뉴런을 그려보자.

(1) $y = 1x$

(2) $y = x_1 + 2x_2 + x_3 + 2x_4$

(3) $y = \begin{cases} 0 & if\ x_1 + 2x_2 + x_3 + 2x_4 > T \\ 1 & otherwise \end{cases}$

학습이란 무엇이다?

어떻게 자동으로 학습할 수 있을까?

요 약

- 뉴런의 연결 부분, 시냅스
- 학습은 연결을 조정하는 것
- 파라미터(W) 튜닝
- 뉴런의 동작
- 뉴런 그리기

Learning occurs…

- while experiencing something
- the strength of connection between neurons is properly changed

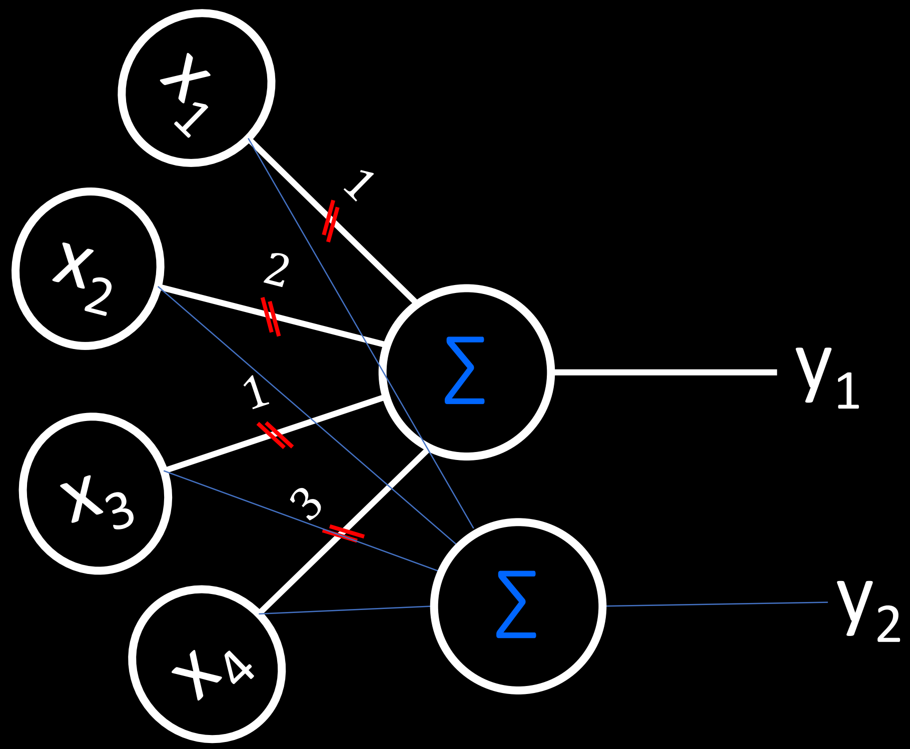
학습 = 연결 값을 조정하는 것

{강하게, 약하게}

$$y = w1x1 + w2x2 + w3x3 + w4x4$$

$$y = [w1, w2, w3, w4] \begin{bmatrix} x1 \\ x2 \\ x3 \\ x4 \end{bmatrix}$$

$$\boldsymbol{y = wx}$$



x1
x2
x3
x4
1
2
1
3
Σ
Σ
y1
y2

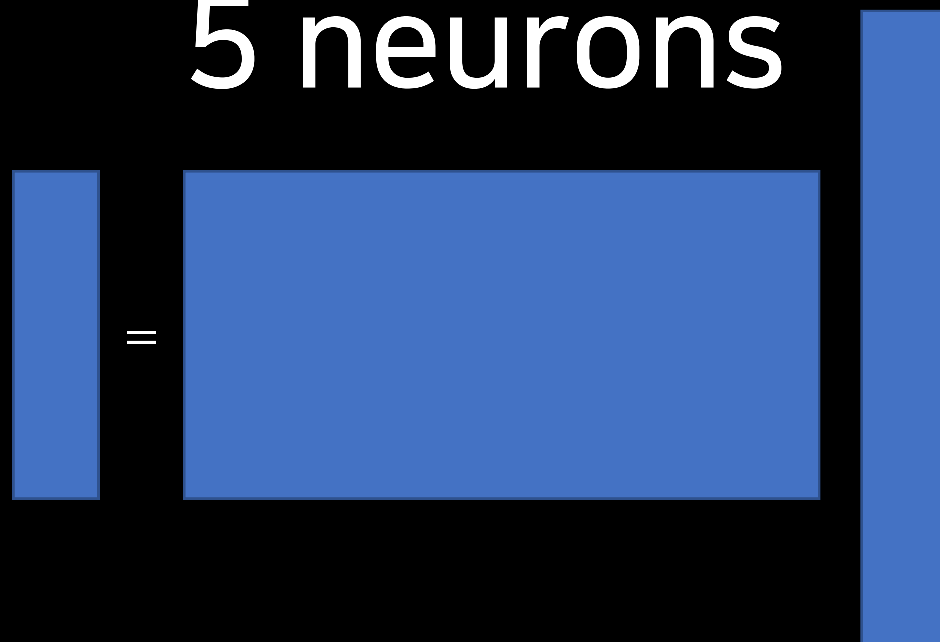
column vector

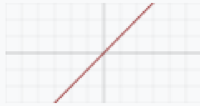
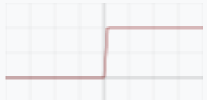
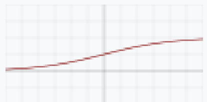
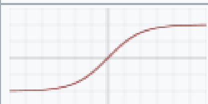
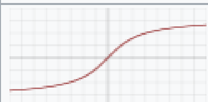
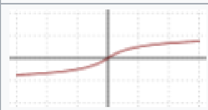
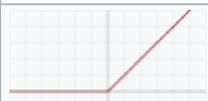
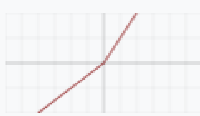
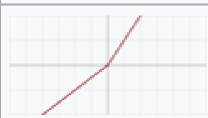
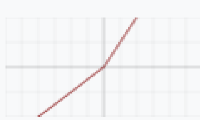
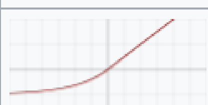
$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_1{}^1, w_2{}^1, w_3{}^1, w_4{}^1 \\ w_1{}^2, w_2{}^2, w_3{}^2, w_4{}^2 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \\ x3 \\ x4 \end{bmatrix}$$

matrix

$$\boldsymbol{y} = W\boldsymbol{x}$$

7 inputs
5 neurons



Name	Plot	Equation	Derivative (with respect to x)
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic (a.k.a. Soft step)		$f(x) = \frac{1}{1 + e^{-x}}$	$f'(x) = f(x)(1 - f(x))$
TanH		$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
ArcTan		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2 + 1}$
Softsign [7][8]		$f(x) = \frac{x}{1 + \lvert x \rvert}$	$f'(x) = \frac{1}{(1 + \lvert x \rvert)^2}$
Rectified linear unit (ReLU)[9]		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Leaky rectified linear unit (Leaky ReLU) [10]		$f(x) = \begin{cases} 0.01x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0.01 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parameteric rectified linear unit (PReLU)[11]		$f(\alpha, x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(\alpha, x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Randomized leaky rectified linear unit (RReLU)[12]		$f(\alpha, x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$ [1]	$f'(\alpha, x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Exponential linear unit (ELU)[13]		$f(\alpha, x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(\alpha, x) = \begin{cases} f(\alpha, x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$