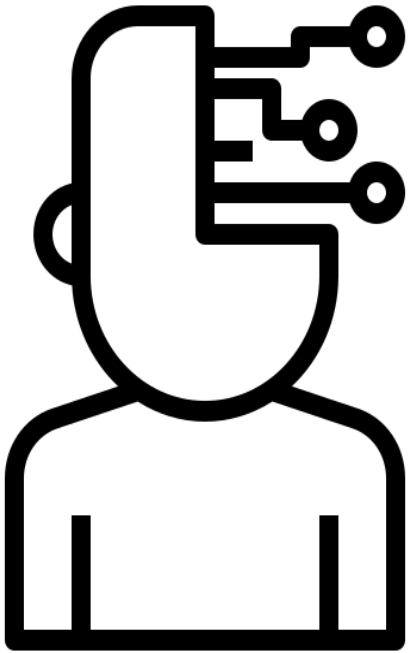
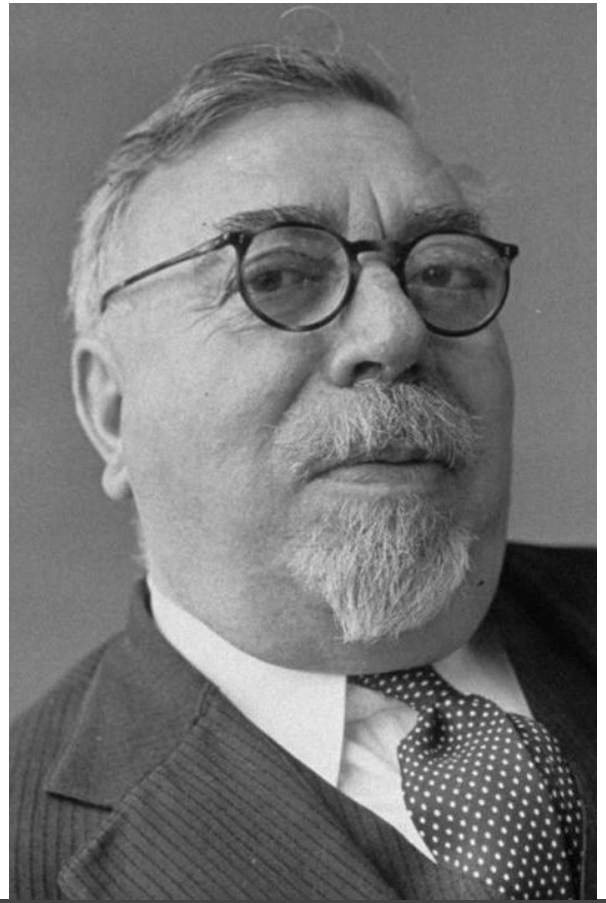
퍼셉트론으로 살펴보는 인공신경망(딥러닝)의 원리

교사 김진관



사이버네틱스(Cybernetics)

생물의 자기 제어의 원리를 기계 장치에 적용하여
통신, 제어, 정보, 처리 등의 기술을 종합적으로
연구하는 학문 분야



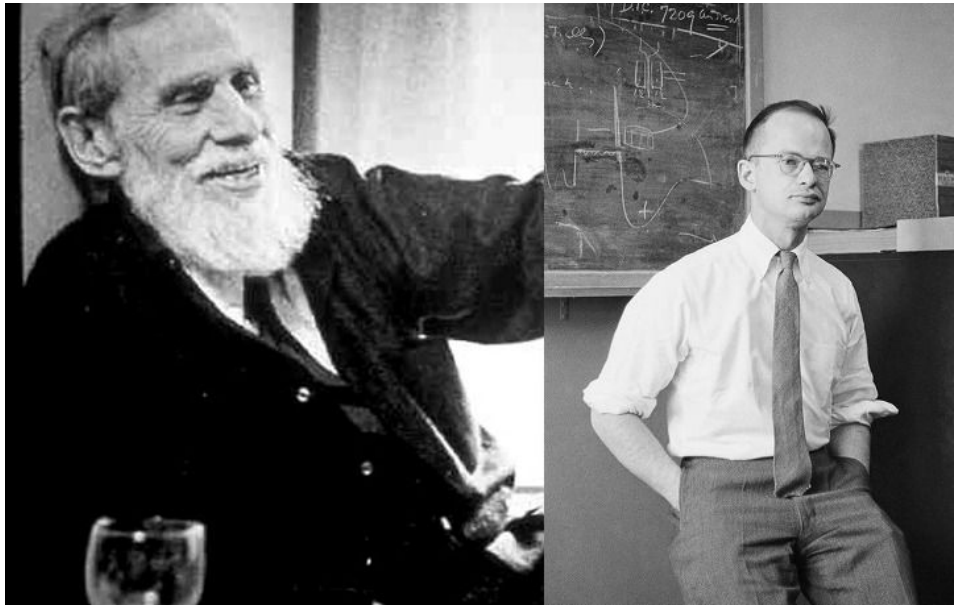
모든 동물은 피드백으로부터 배운다.

인간과 기계 사이에
‘피드백을 통한 통제’라는
공통분모가 존재한다.

인간처럼 기계도 피드백을 통해
스스로 학습할 수 있지 않을까?

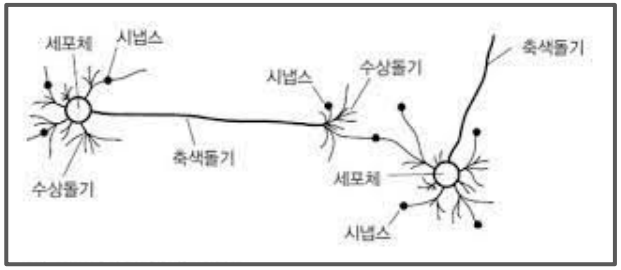
-사이버네틱스의 창시자
노버트위너-

뇌는 어떻게 작동하길래
생각하고 학습할 수 있을까?

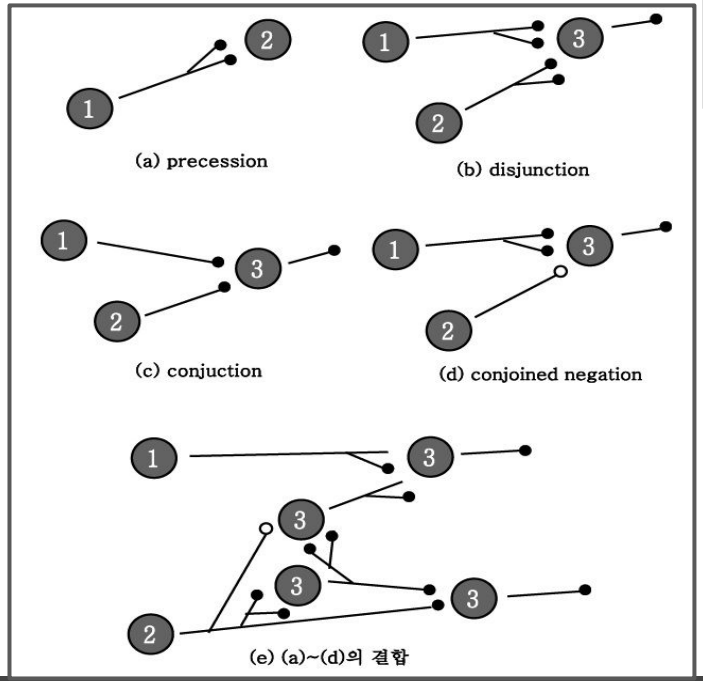


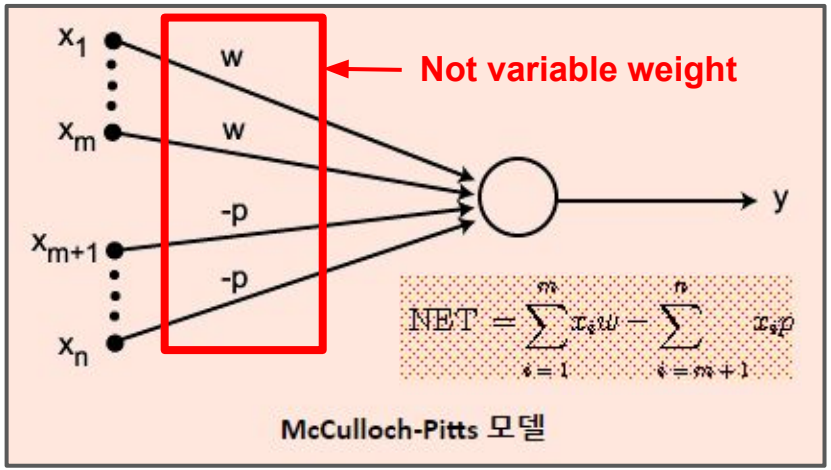
워렌 맥컬럭과 월터피츠(1943)

인간의 뇌를 최초로 수학적으로 모델링하다

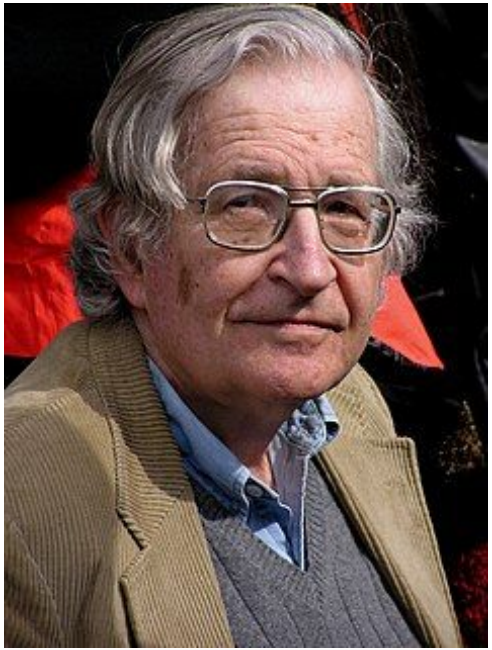


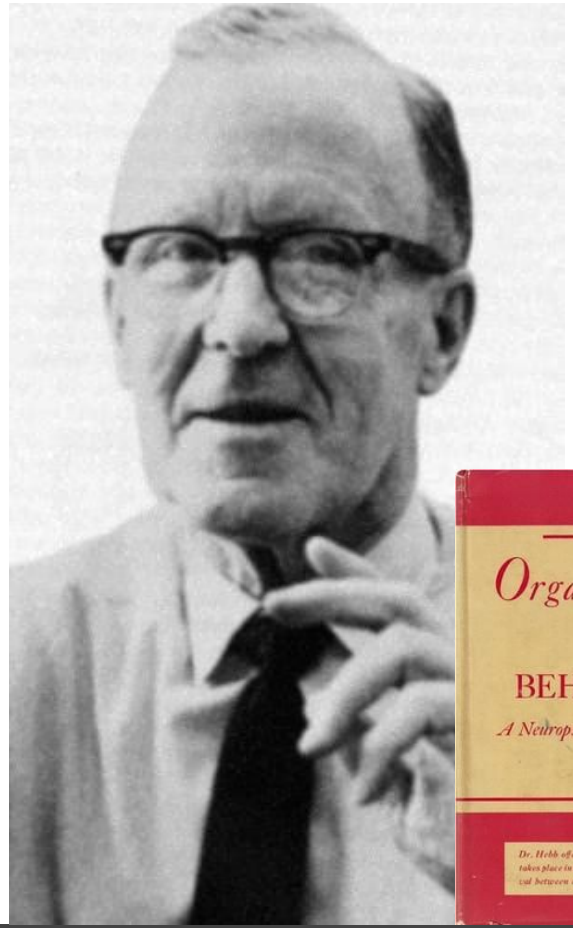
- 뉴런은 활성화되거나 활성화되지 않은 2가지 상태이다. 즉, 뉴런의 활성화는 '전부 아니면 전무' 프로세스이다.
- 어떤 뉴런을 흥분되게(excited) 하려면 2개 이상의 고정된 수의 시냅스가 일정한 시간 내에 활성화되어야 한다.
- 신경 시스템에서 유일하게 의미 있는 시간 지연(delay)은 시냅스에서의 지연(synaptic delay)이다.
- 어떠한 억제적인(inhibitory) 시냅스는 그 시각의 뉴런의 활성화를 절대적으로 방지한다.
- 신경망의 구조는 시간에 따라 변하지 않는다.

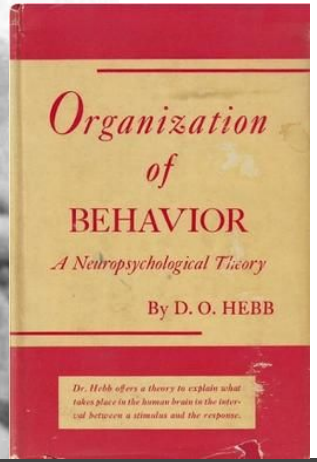




학습에 대한 논쟁

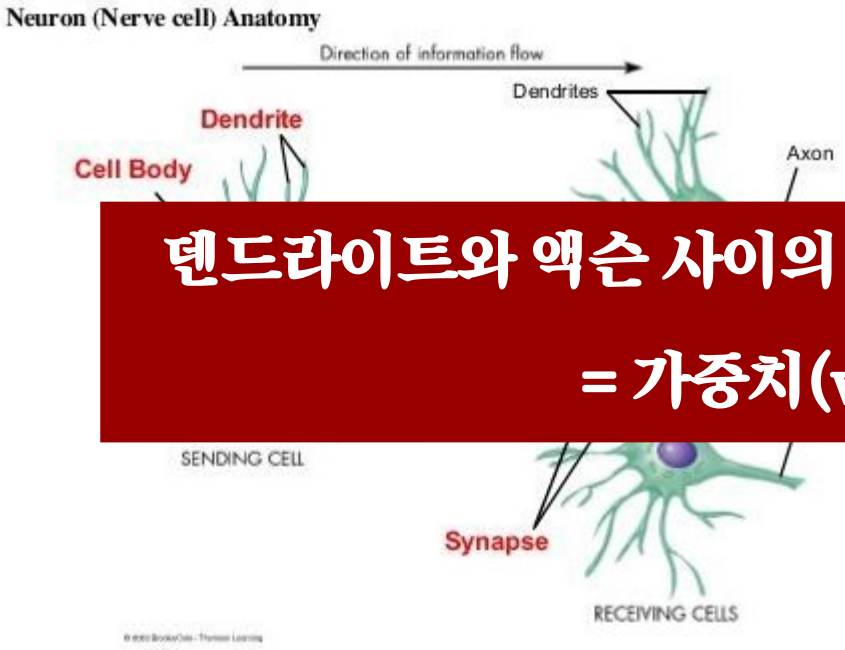






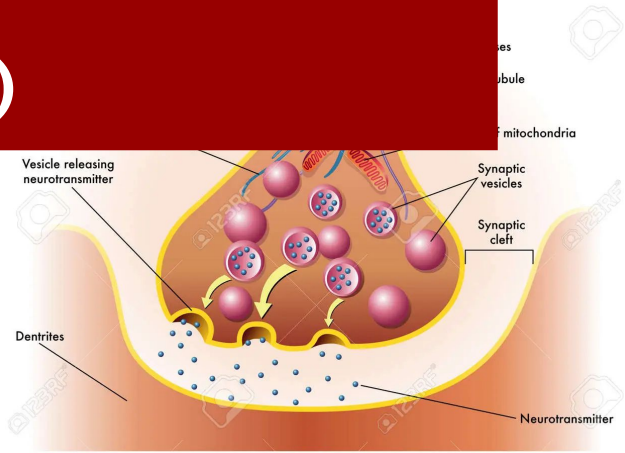
도널드 헵의 학습이론(1949)

인간 학습은
뇌세포의 연결강화를
의미한다.



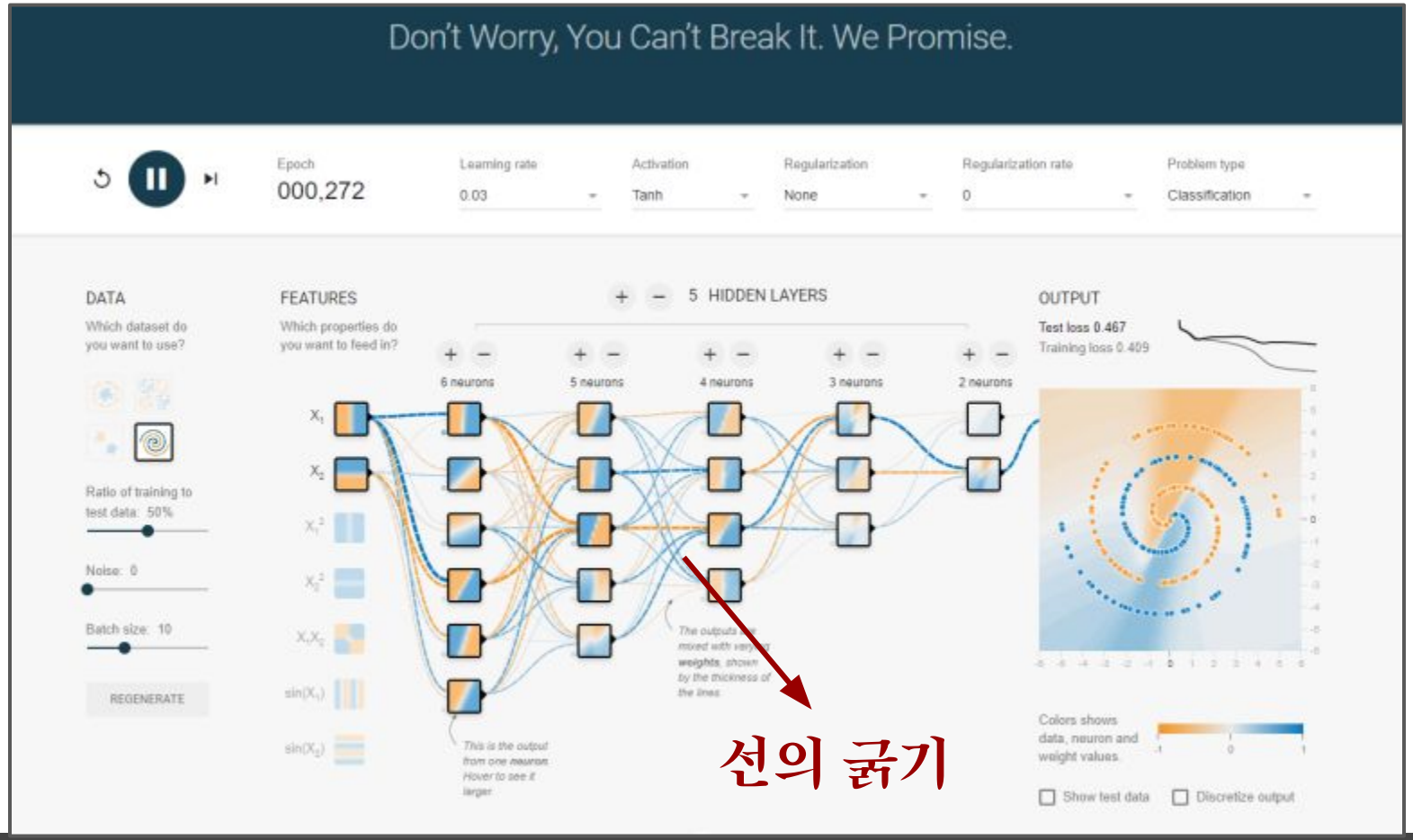
덴드라이트와 액슨 사이의 시냅스에 있는 화학물질

= 가중치(weight)

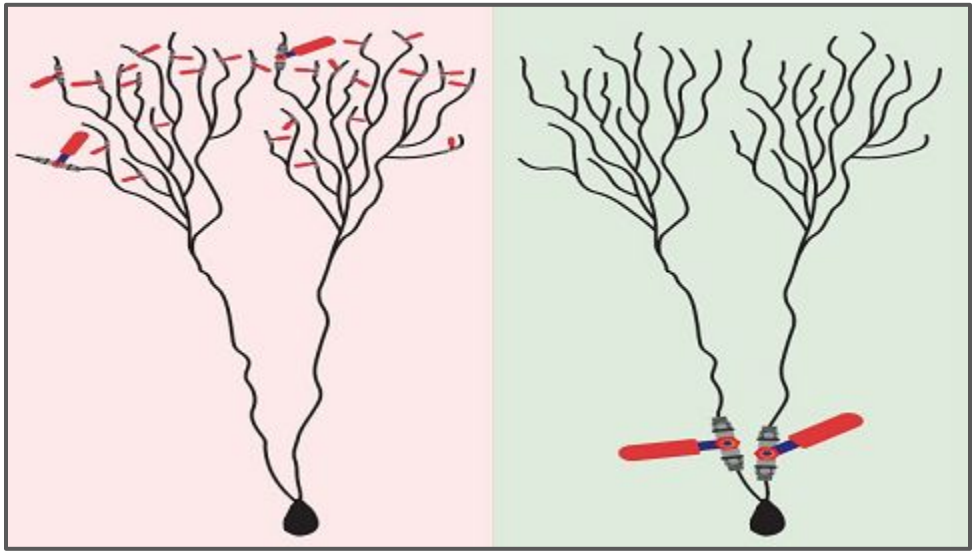


학습은 가중치의 값을
조정하는 것이다.

가중치의 값은
선의 굵기, 가스 밸브의 개폐정도로
표현될 수 있다.



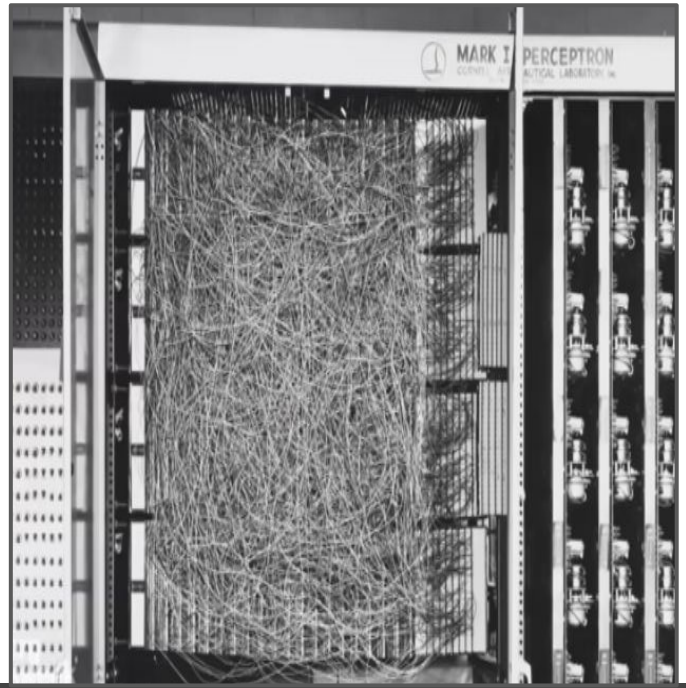
가스 밸브

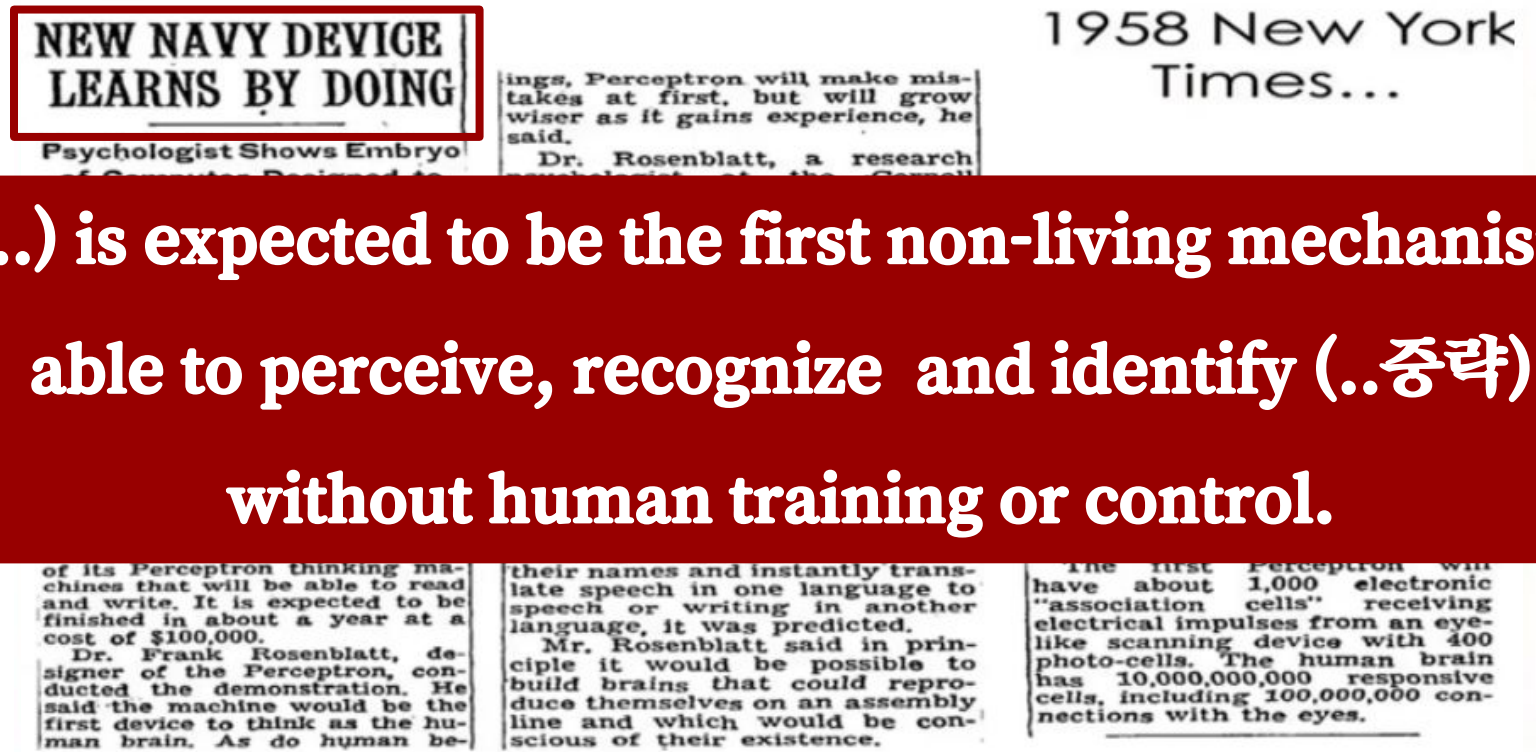


연결주의의 본격적인 시작

All the knowledge is in the connectionism.
-David Rumelhart-

로젠블랏과 최초의 신경망, 퍼셉트론





NEW NAVY DEVICE LEARNS BY DOING

Psychologist Shows Embryo

1958 New York Times...

(..) is expected to be the first non-living mechanism able to perceive, recognize and identify (..중략) without human training or control.

of its Perceptron thinking machines that will be able to read and write. It is expected to be finished in about a year at a cost of $100,000.

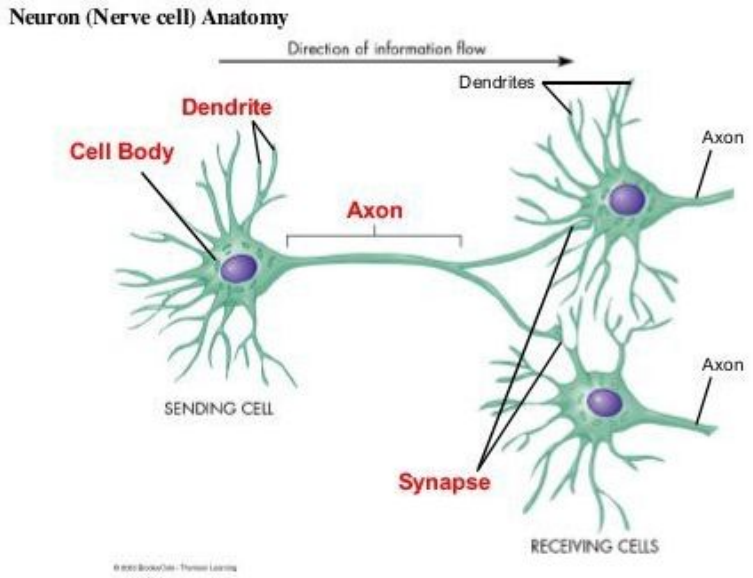
Dr. Frank Rosenblatt, designer of the Perceptron, conducted the demonstration. He said the machine would be the first device to think as the human brain. As do human beings, Perceptron will make mistakes at first, but will grow wiser as it gains experience, he said.

Dr. Rosenblatt, a research

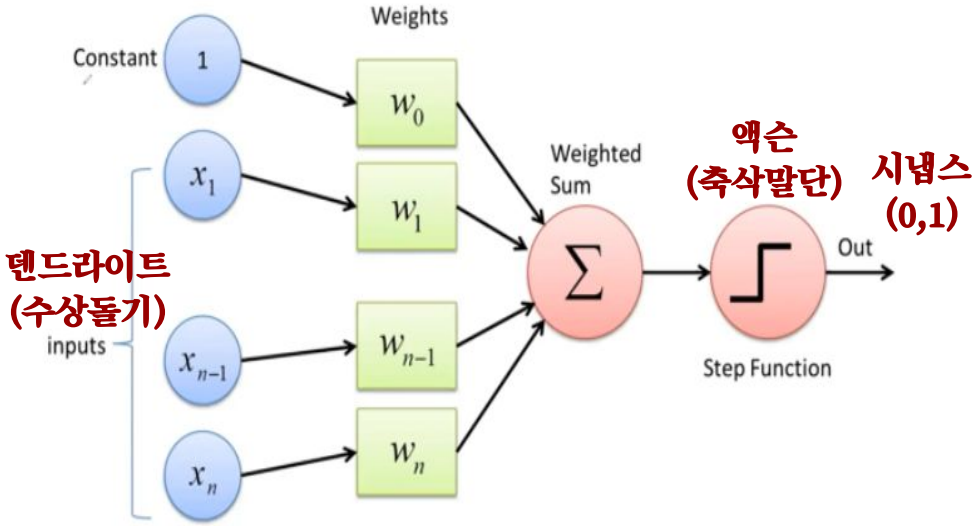
their names and instantly translate speech in one language to speech or writing in another language, it was predicted.

Mr. Rosenblatt said in principle it would be possible to build brains that could reproduce themselves on an assembly line and which would be conscious of their existence.

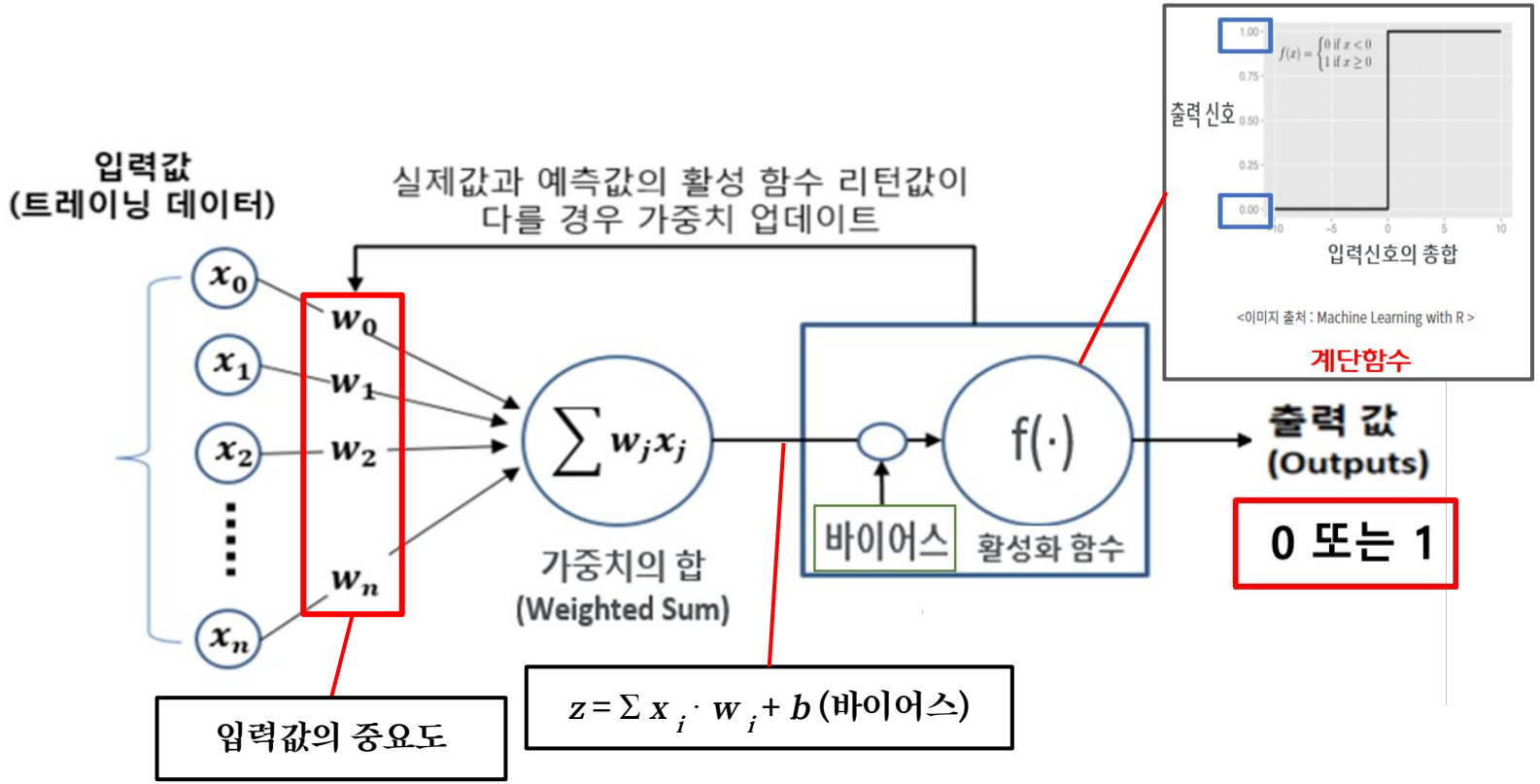
The first Perceptron will have about 1,000 electronic "association cells" receiving electrical impulses from an eye-like scanning device with 400 photo-cells. The human brain has 10,000,000,000 responsive cells, including 100,000,000 connections with the eyes.



생물학적 신경계

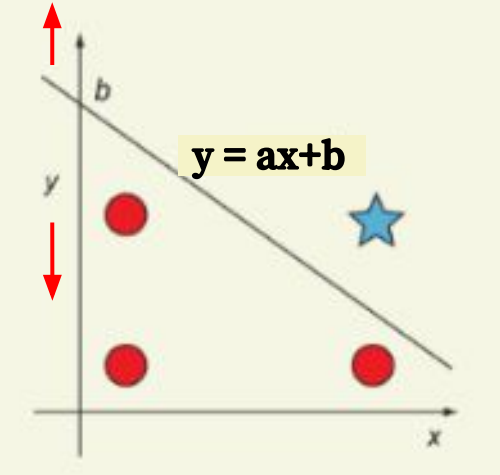


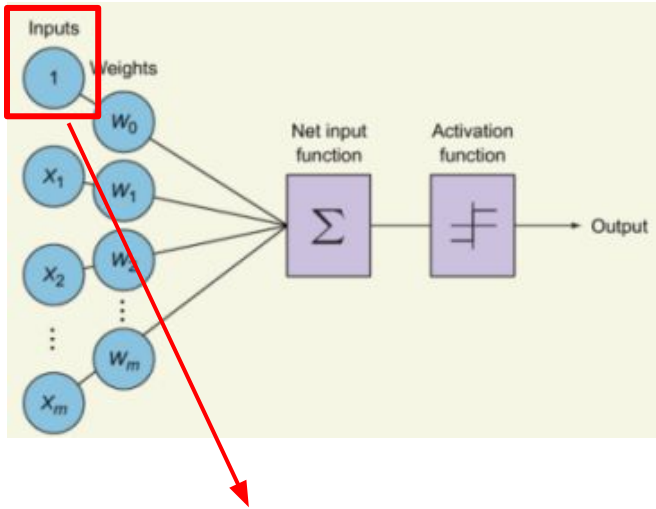
퍼셉트론 모델(1958)



가중치(W), 가중합(WS), 바이어스(편향), 활성화 함수

퍼셉트론에서 편향이 필요한 이유





편향 = y절편 = b

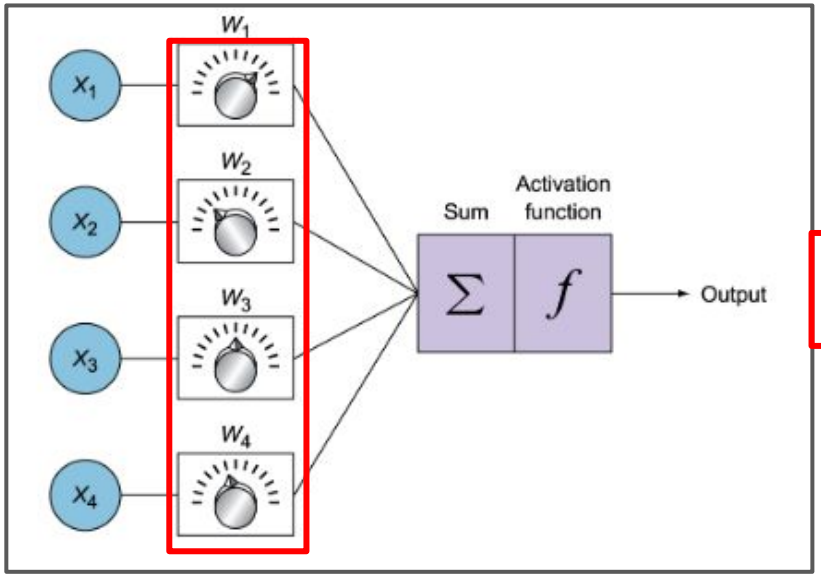
1인 입력 추가하면 편향 추가된 효과

편향은 가중치가 하나 추가된 것처럼 취급됨

퍼셉트론에서의 학습은?

1. 순방향 계산

$z = \Sigma x_i \cdot w_i + b$ (바이어스)

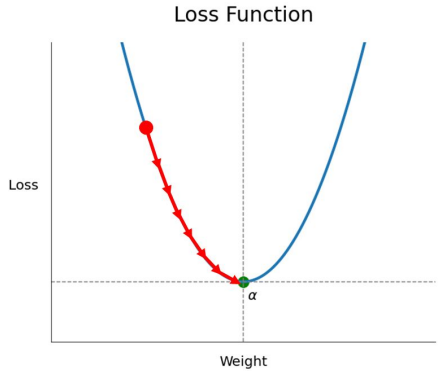


손실함수

2. 오차(예측 값 - 실제 값) 계산

3. 오차에 따라 가중치(=볼륨 손잡이) 조절

4. 1~3의 과정을 반복

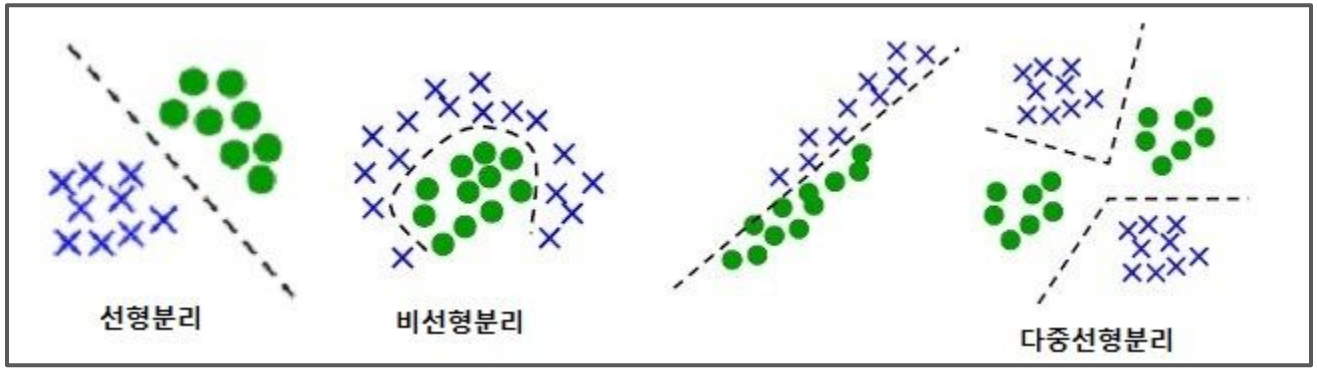


퍼셉트론에서의 학습은
적정한 가중치의 값을 찾는데 있다.

적정한 가중치란, 손실 함수의 값을 최소가 되게 하는 가중치이다.

기호주의가 외면 받자 연결주의에 반격을 개시합니다.

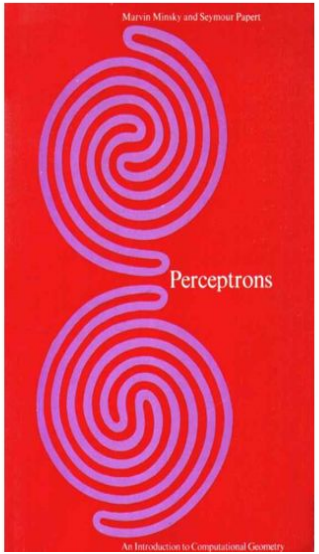
하나의 뉴런만으로
복잡한 문제를 해결할 수 있을까?



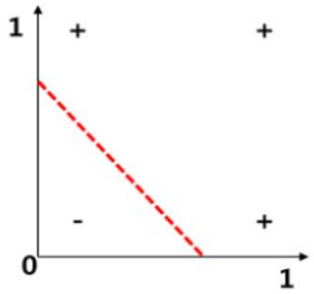
선형 데이터 셋: 하나의 직선으로 데이터 분류 가능

비선형 데이터 셋: 제대로 분류하려면 2개 이상의 직선이 필요

- XOR(Exclusive OR): 배타적 논리합(입력 값이 서로 다르면 1을 출력, 같으면 0을 출력)
- 단층 퍼셉트론으로는 XOR문제 해결 불가능

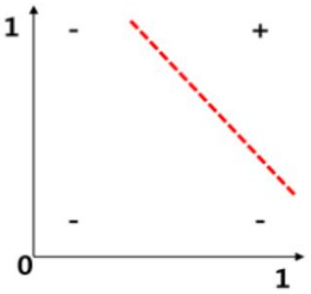


OR



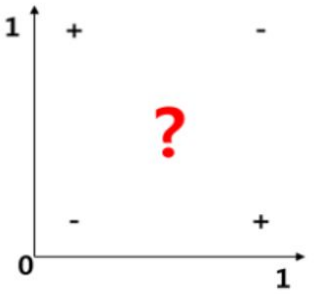
x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	1

AND

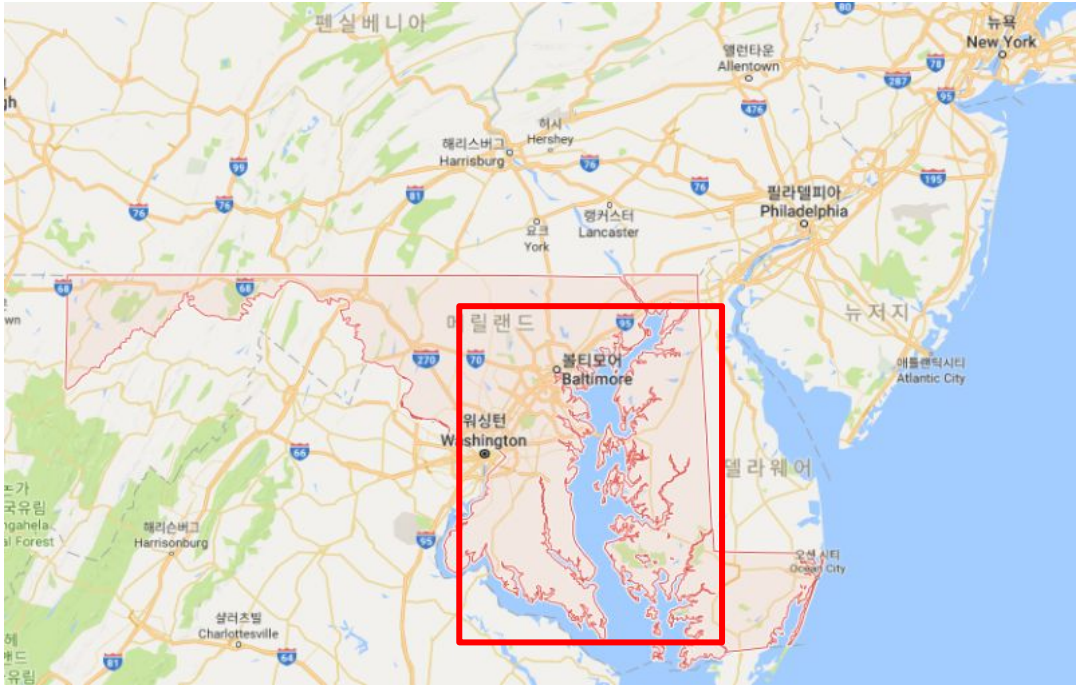


x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

XOR



x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

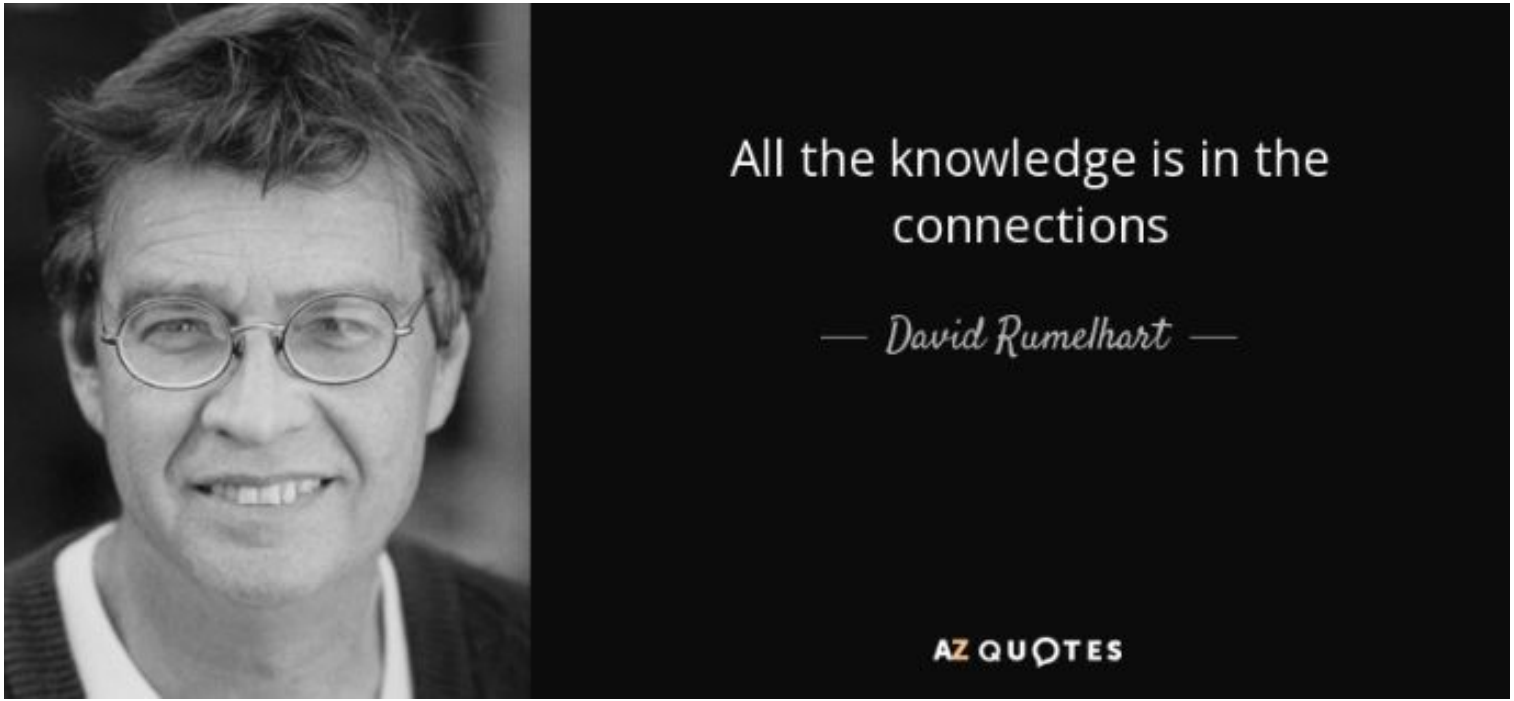


1971년,
체서피크만의 요트 위에서
비운의 죽음을 맞이한
프랭크 로젠블랏

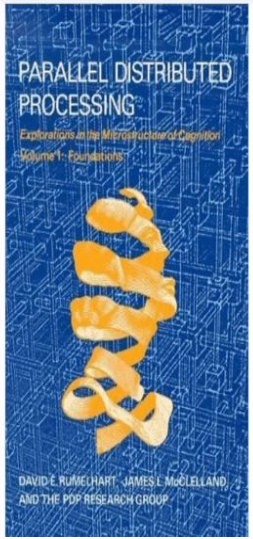
녹록치 않은 세상의 문제를 해결하기
위해서는 여러개의 퍼셉트론이 필요합니다.

여러 개의 퍼셉트론을 갖춘 신경망을 만들 수만 있다면?!

연결주의, 아직 죽지 않았어



- 다층 퍼셉트론으로 XOR문제 해결

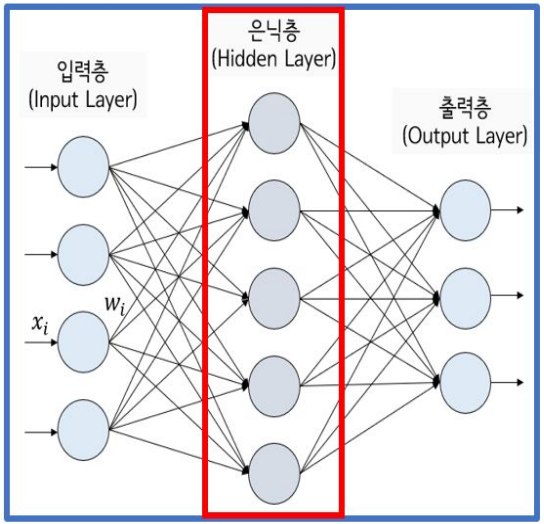


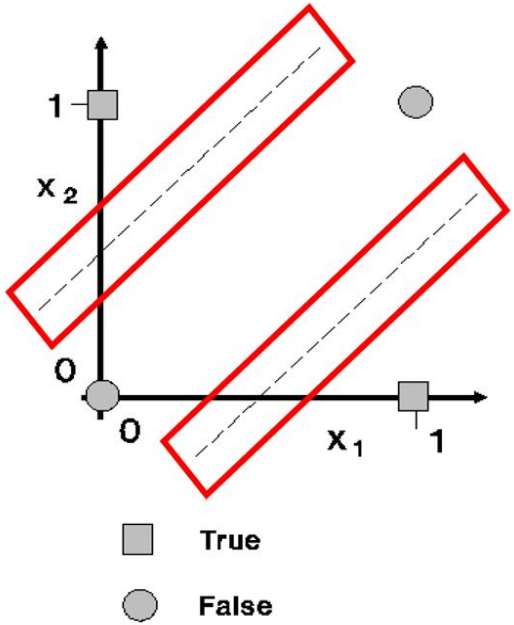
Parallel Distributed Processing, Volume 1

Explorations in the Microstructure of Cognition: Foundations

By David E. Rumelhart, James L. McClelland and PDP Research Group

A Bradford Book

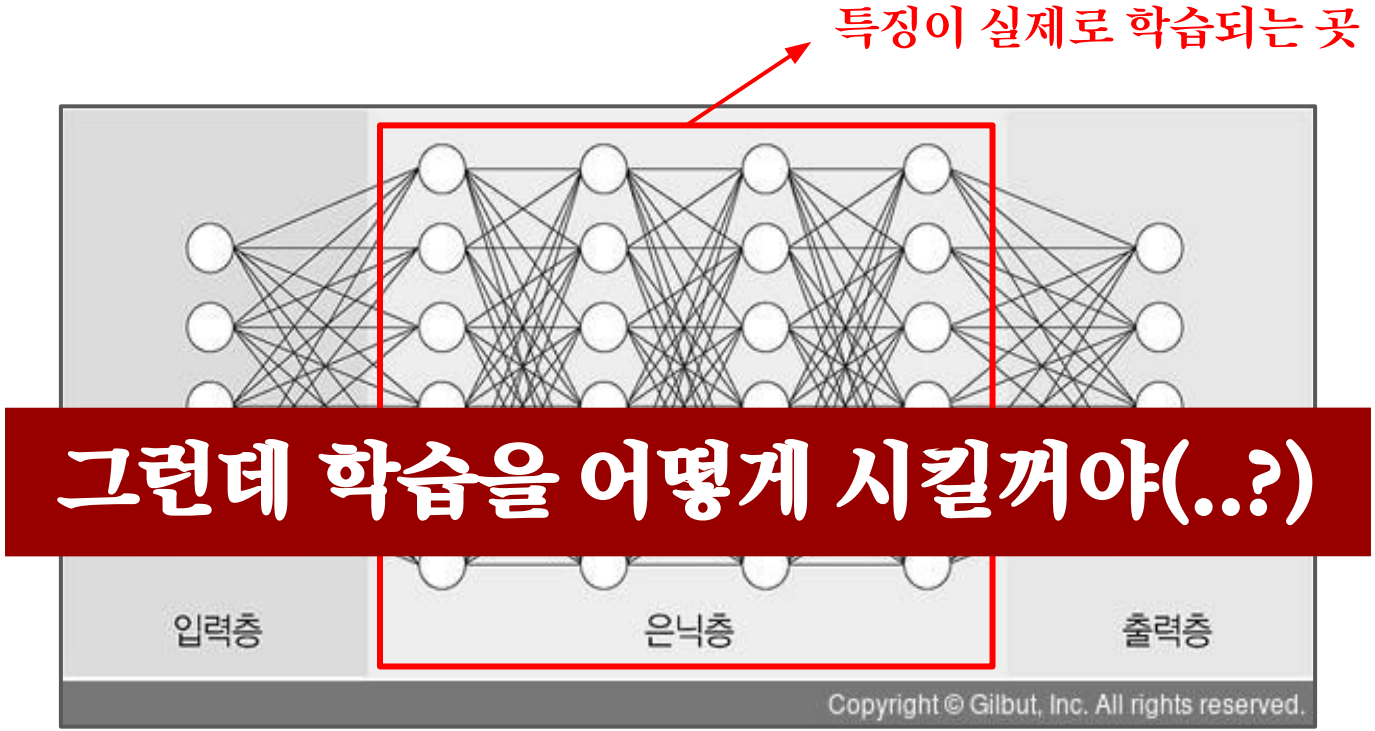




인공지능은 병렬, 분산된 처리를 해야 한다.

은닉층 추가

XOR 문제는 너무 쉽게 해결



그런데 학습을 어떻게 시킬꺼야(..?)

은닉층이 2개 이상인 신경망 = DNN(심층 신경망)

In an interview of 2018,[33] Hinton said that "David E. Rumelhart came up with the basic idea of backpropagation, so it's his invention."

Learning representations by back-propagating errors

David E. Rumelhart*, Geoffrey E. Hinton† & Ronald J. Williams*

* Institute for Cognitive Science, C-015, University of C
San Diego, La Jolla, California 92093, USA
† Department of Computer Science, Carnegie-Mellon U
Pittsburgh, Philadel

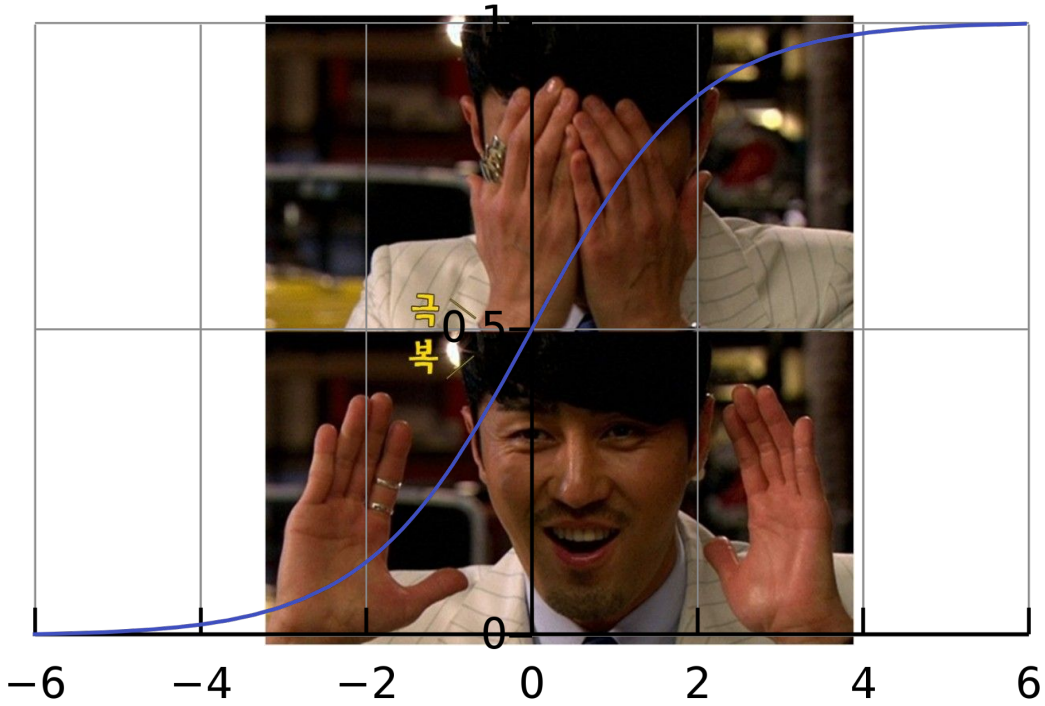
We describe a ne
networks of neuro
the weights of the
measure of the difference between the actual output
net and the desired output vector. As a result
adjustments, internal 'hidden' units which are not pa

more difficult when we introduce hidden units whose actual or desired states are not specified by the task. (In perceptrons,

층이 깊어지니까 학습이 또 잘 안되네?

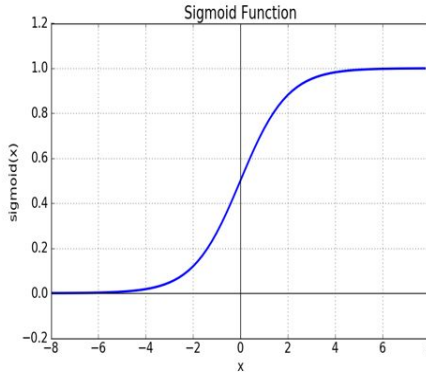
체인룰과 편미분으로 백프로퍼게이션 알고리즘(1986) 구현

ReLU 활성화 함수로 극복

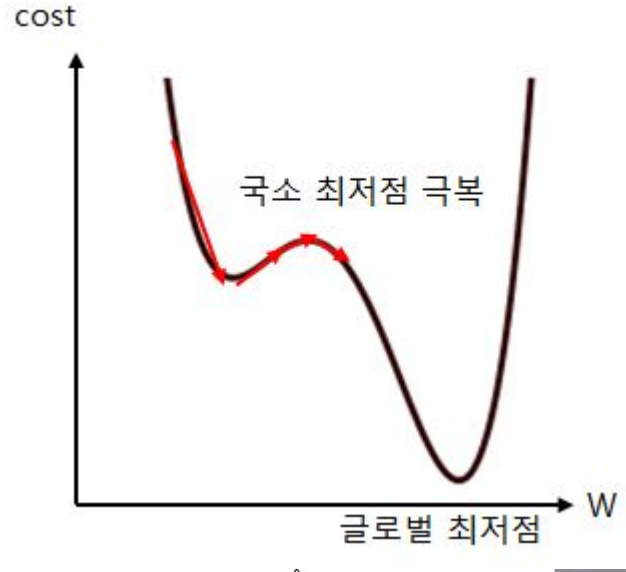


ReLU 활성화 함수로 극복

- 레이어(층)을 깊게 쌓을수
- 시그모이드 함수를 탑재한
- 다른 현실 세계의 문제에



원인: 시그모이드함수



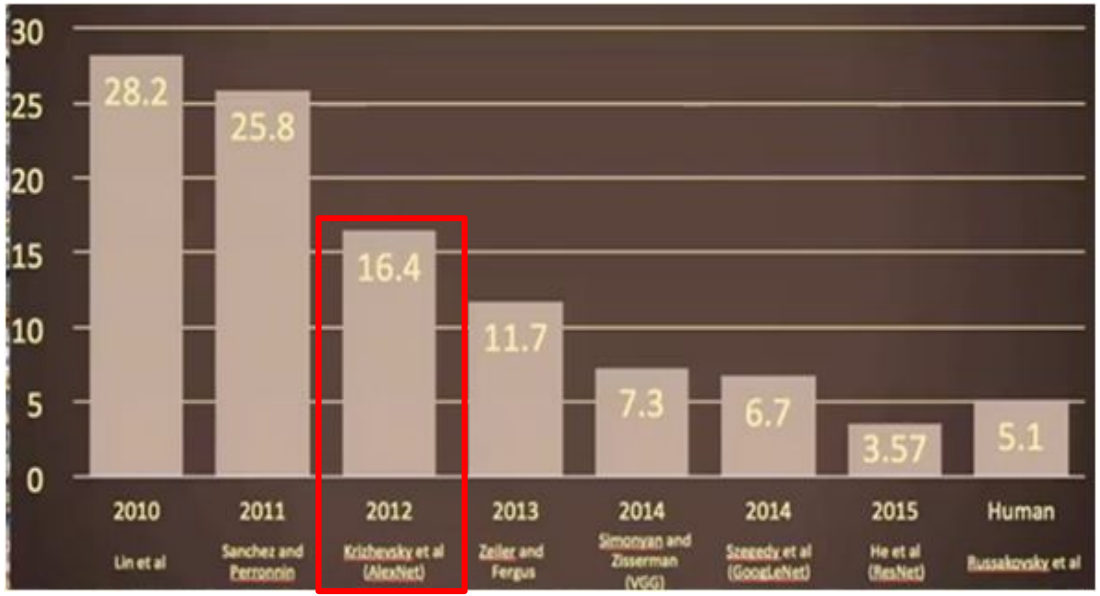
해결책: ReLU함수

shing Gradient Problem)

전파가 잘 안되네?

딥러닝을 세상에 알리다. 토론토 대학의 슈퍼비전팀
출처: https://www.wired.com/2013/03/google-hinton/

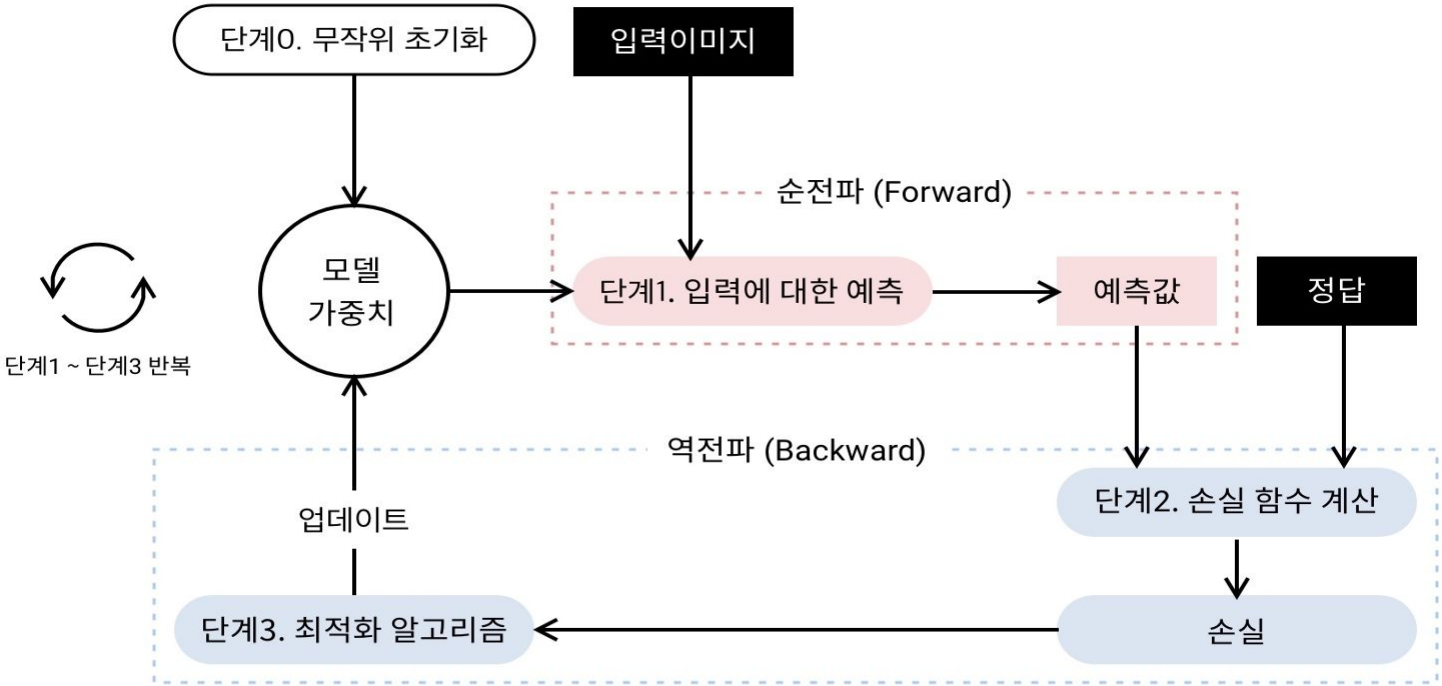
딥러닝, 월드 클래스로 변모하다.

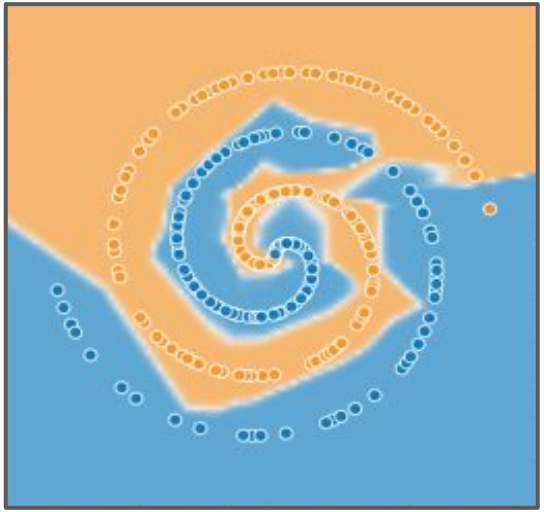


용어 정리

퍼셉트론(MLP) = 인공신경망(ANN) =
전결합층(FC) = 피드포워드 신경망 = 딥러닝

딥러닝 학습 정리



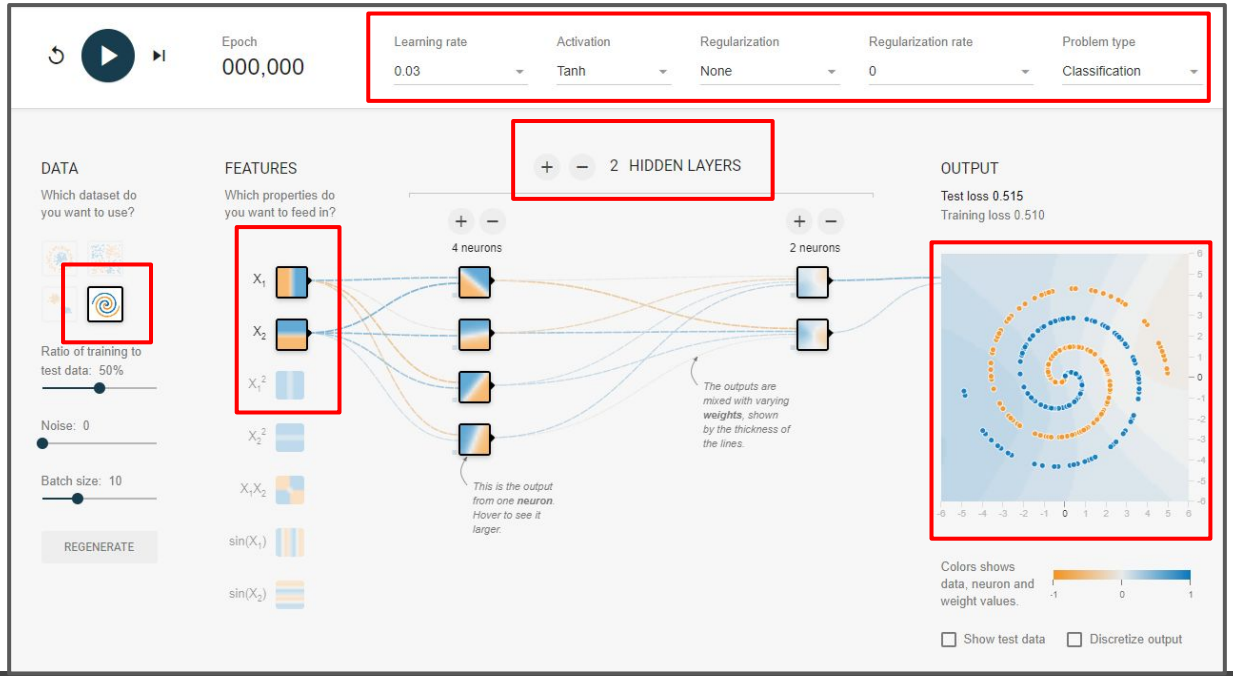


딥러닝의 진가는
비선형 데이터를 분리하는데 있다.

신경망은 구조가 중요합니다.

성능이 너무 낮다면(데이터 부합X) 층수를 늘리고,
과적합(학습데이터에만 최적화)의 느낌이 난다면
층수를 줄인다.

http://playground.tensorflow.org/



퍼셉트론으로 살펴보는 인공신경망(딥러닝)의 원리

교사 김진관