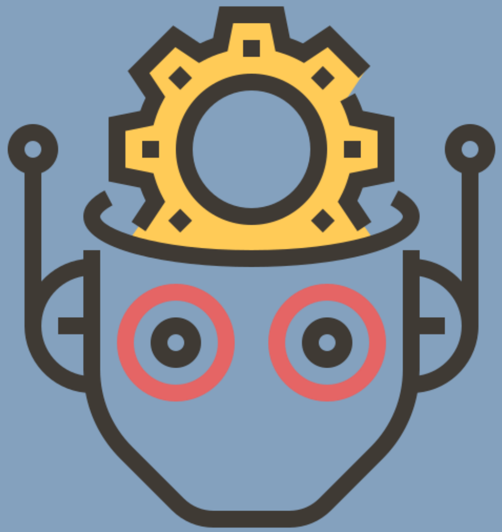
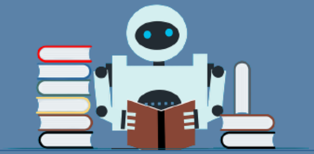


Deep Learning

스마트인재개발원
Smart Human Resources Development

정 봉 균 연구원



1	퍼셉트론의 개념을 이해할 수 있다.
2	다층 퍼셉트론의 개념을 이해할 수 있다.

딥러닝 역사

Deep
Learning

딥러닝 역사

1943년
ANN 개념 도입

1958년
Perceptron 제안

1969년
XOR 문제 지적

1986년
XOR 문제 해결
(MLP)

1986년
파라미터 개수 증가 → 가중치, 편향 학습 어려움 →
오차역전파 알고리즘

1997년
LSTM

1998년
CNN

1990대
Vanishing Gradient/과적
합/진동·발산 문제 대두

2000년 초
머신러닝의 전진
(SVM, Random forest)

2006년
과적합 문제 해결
→비지도학습-RBM→
DBN (딥러닝)제안

2007년
자기부호화기

2013년
Dropout / Maxout
(과적합 문제 해결)

ReLU 적용
(Vanishing Gradient 문제
해결)

2014년
GRU
(Vanishing Gradient 문제
해결)

2015년
DQN
(Alphago)

2023년
ChatGPT(MS Bing)

Deep Learning

딥러닝 역사

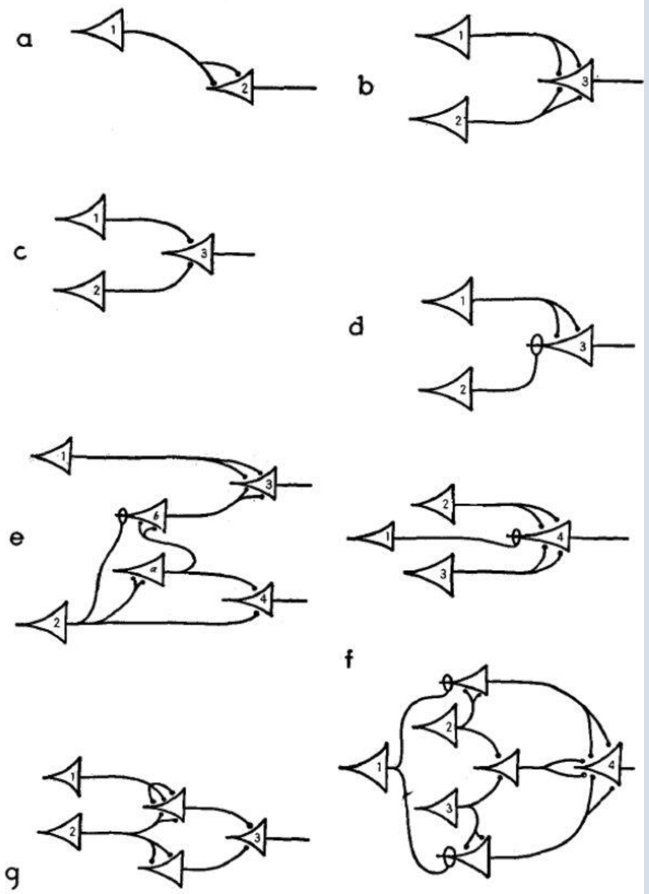
BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 5, 1943

A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY

WARREN S. MCCULLOCH AND WALTER PITTS

FROM THE UNIVERSITY OF ILLINOIS, COLLEGE OF MEDICINE,
DEPARTMENT OF PSYCHIATRY AT THE ILLINOIS NEUROPSYCHIATRIC INSTITUTE,
AND THE UNIVERSITY OF CHICAGO

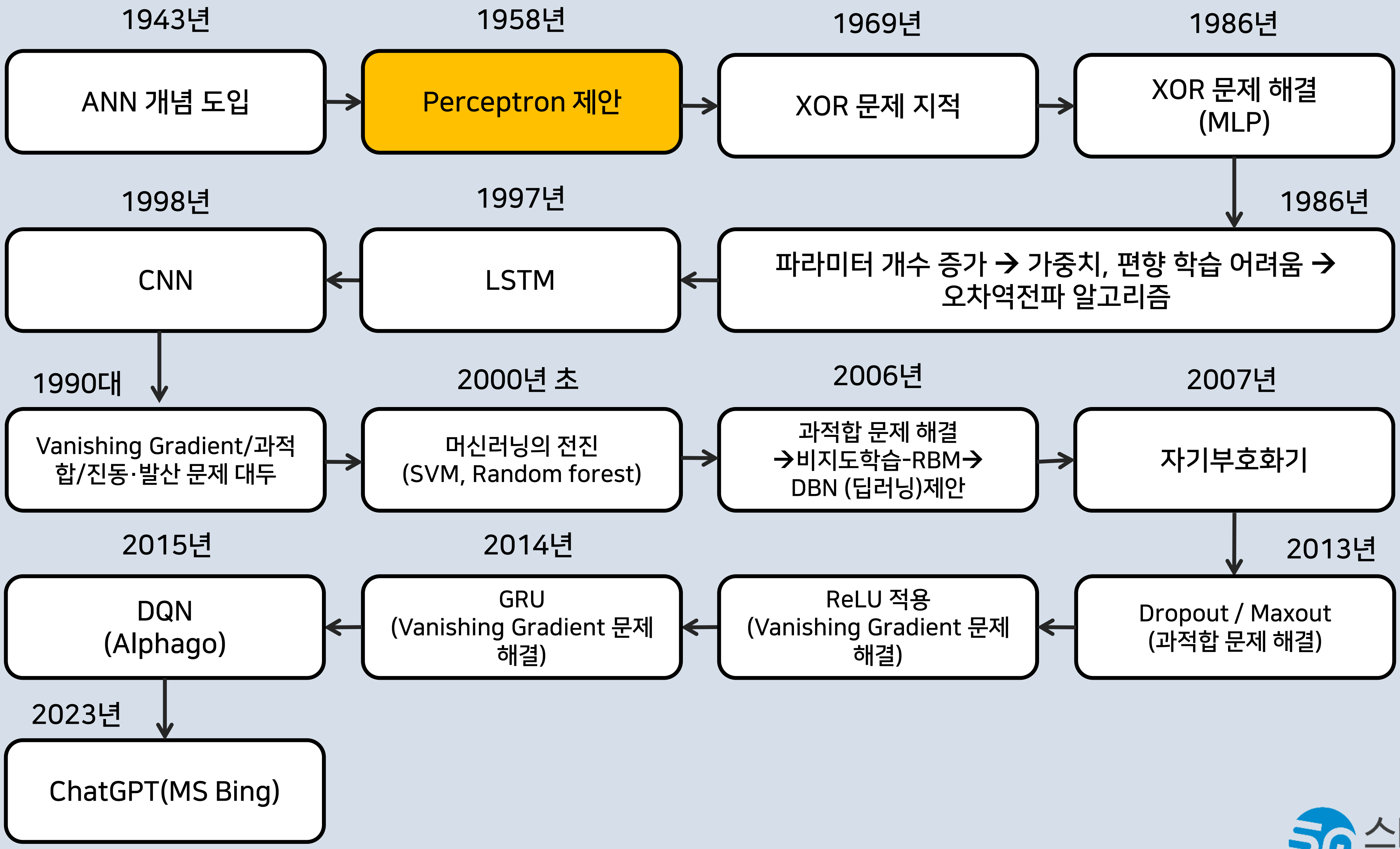
Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for nets containing circles; and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular choices among possible neurophysiological assumptions are equivalent, in the sense that for every net behaving under one assumption, there exists another net which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.



출처 : 논문_신경 활동에 있어서 관념의 논리적 미적분

스마트인재개발원
Smart Human Resources Development

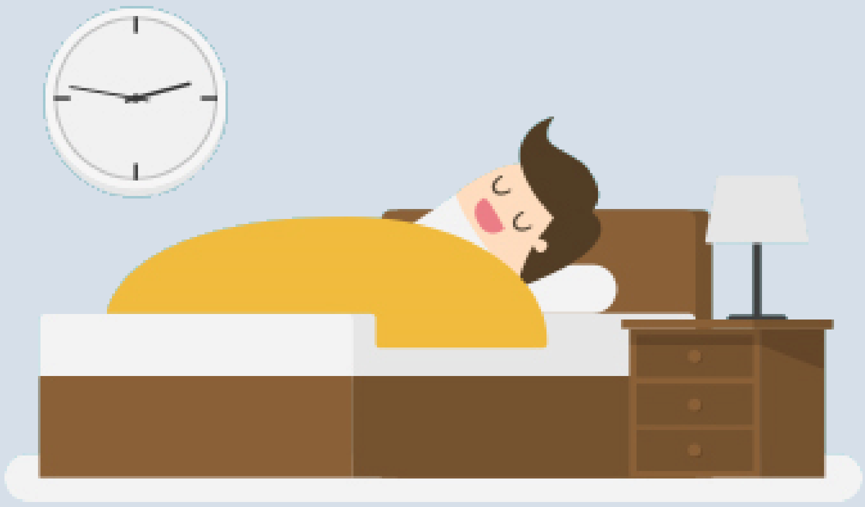
딥러닝 역사

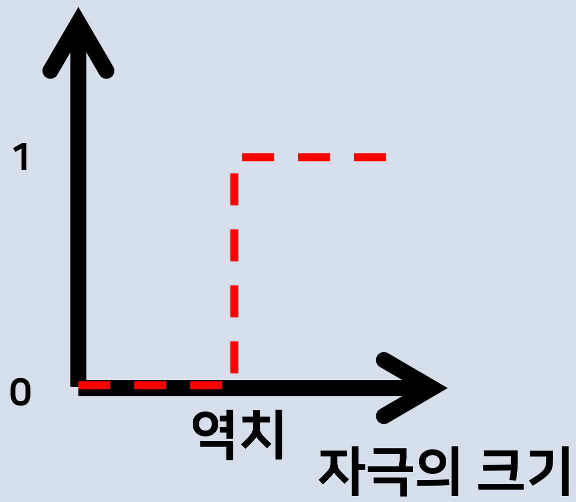


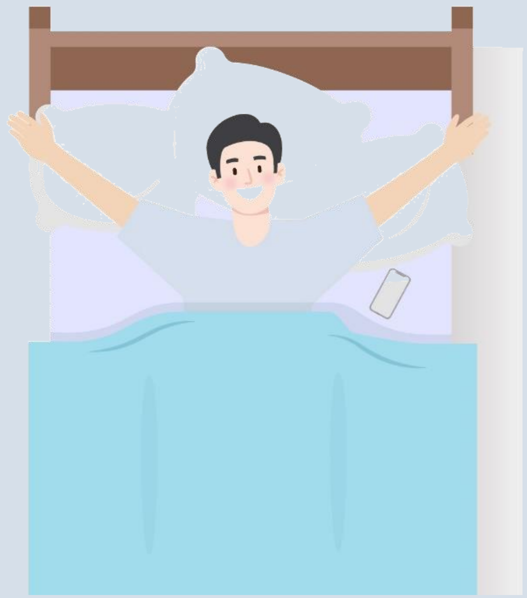
1943년
ANN 개념 도입
1958년
Perceptron 제안
1969년
XOR 문제 지적
1986년
XOR 문제 해결 (MLP)
1986년
파라미터 개수 증가 → 가중치, 편향 학습 어려움 → 오차역전파 알고리즘
1997년
LSTM
1998년
CNN
1990대
Vanishing Gradient/과적합/진동·발산 문제 대두
2000년 초
머신러닝의 전진 (SVM, Random forest)
2006년
과적합 문제 해결 →비지도학습-RBM→ DBN (딥러닝)제안
2007년
자기부호화기
2013년
Dropout / Maxout (과적합 문제 해결)
2014년
ReLU 적용 (Vanishing Gradient 문제 해결)
2014년
GRU (Vanishing Gradient 문제 해결)
2015년
DQN (Alphago)
2023년
ChatGPT(MS Bing)

퍼셉트론(Perceptron)

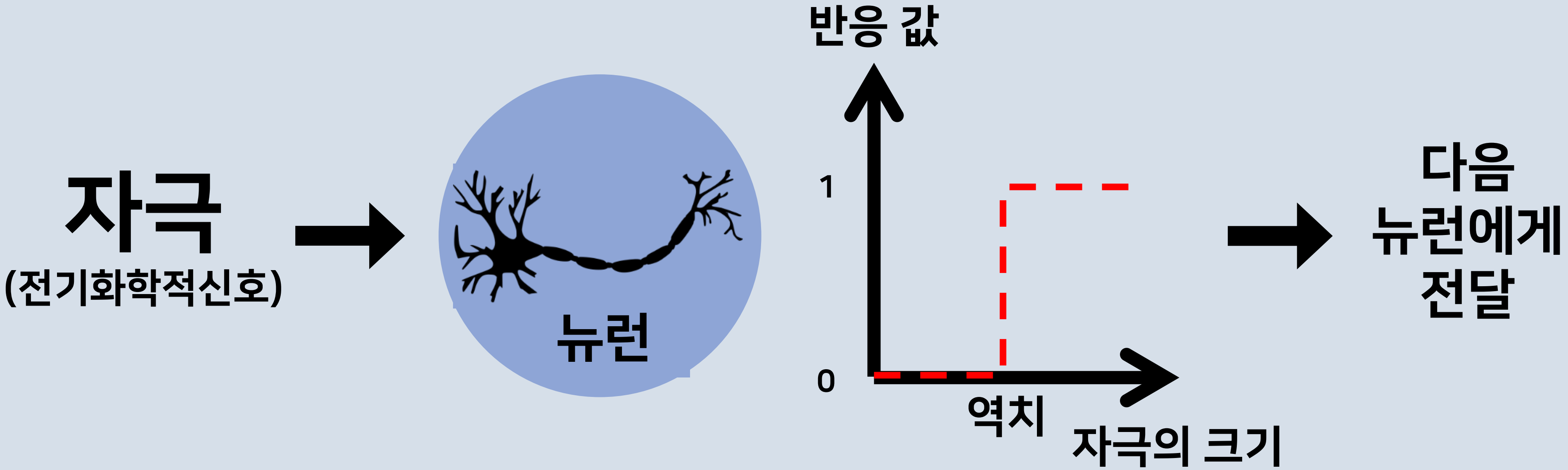
퍼셉트론(Perceptron)





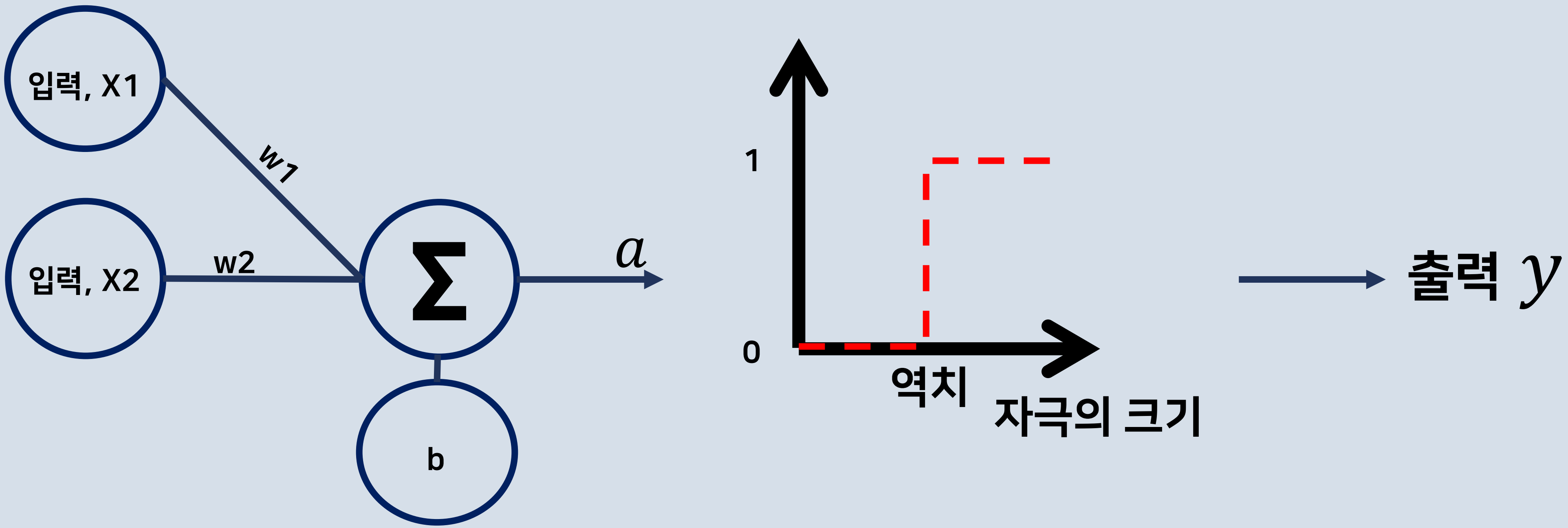


퍼셉트론(Perceptron)

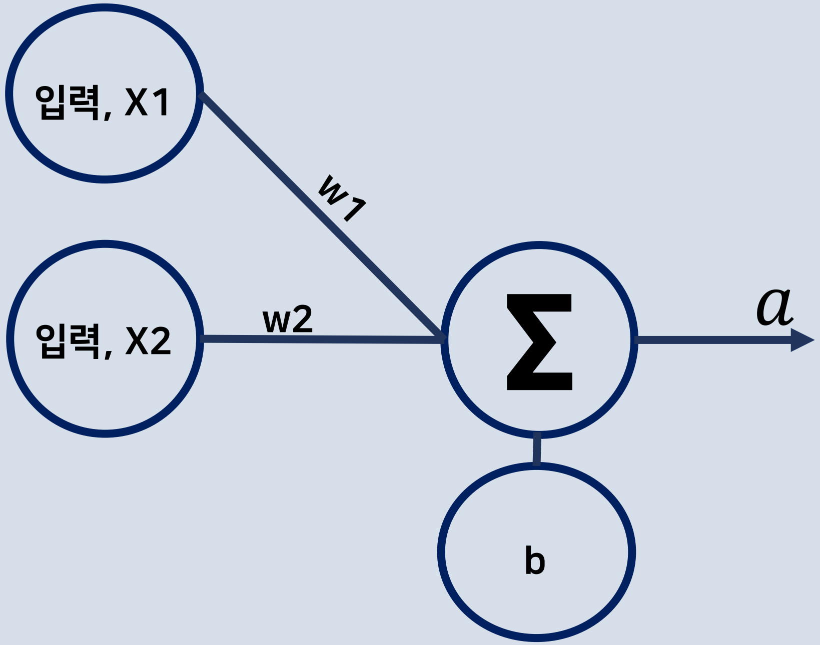


**어떠한 자극에 대해서 일정 자극 이상(역치 이상)이 되면,
다음 뉴런에게 자극을 전달**

퍼셉트론(단층)의 구조



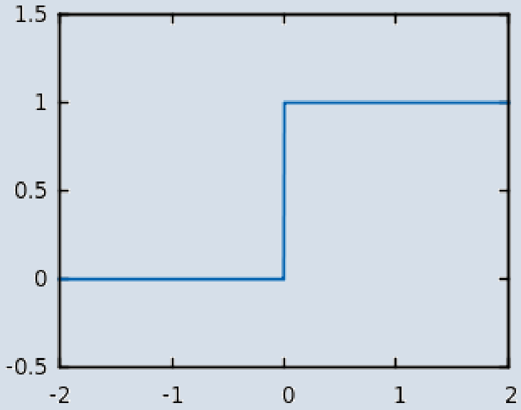
퍼셉트론(단층)의 구조



$$a = W_1X_1 + W_2X_2 + b$$
$$y = h(a)$$

$h(a)$
Activation function
(활성화함수)

출력 y

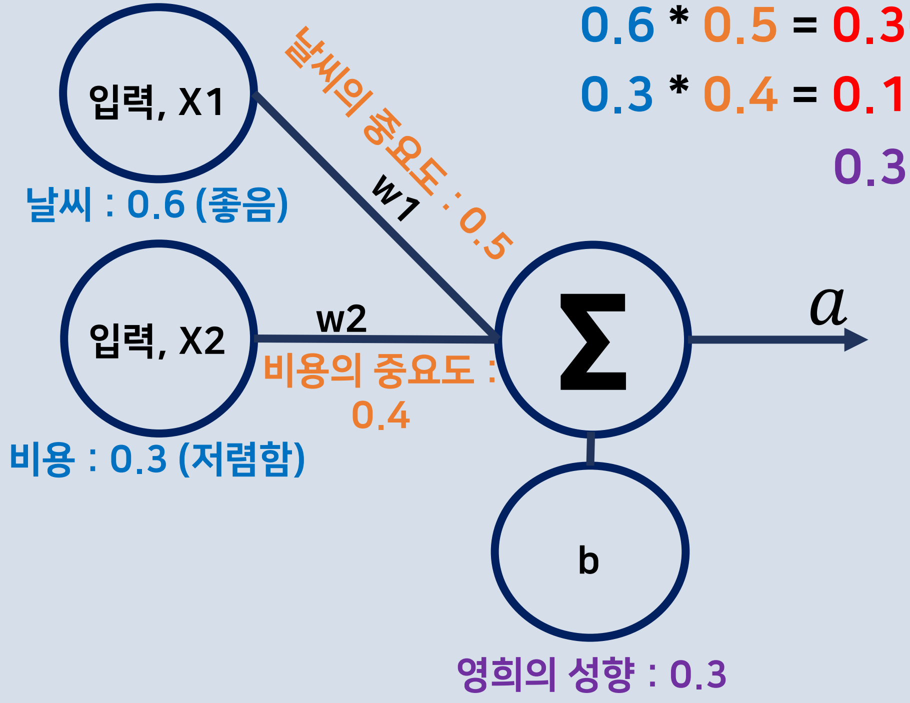


$$h(a) = \begin{cases} 0 & (a \leq 0) \\ 1 & (a > 0) \end{cases}$$

Step function(계단함수)

퍼셉트론(단층)의 작동방식

ex) 콘서트장 예매 – 영희의 경우



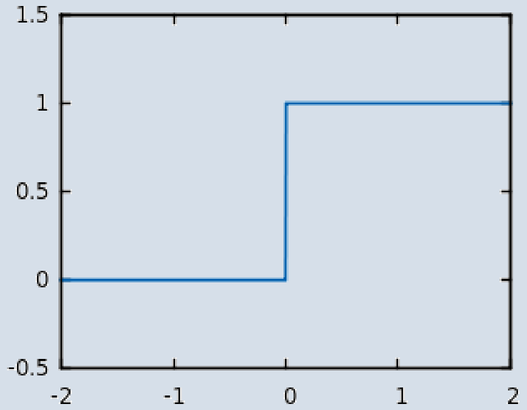
0.6 * 0.5 = 0.30
0.3 * 0.4 = 0.12
0.30

→ 0.42 + 0.30 = 0.72 → $0.72 > 0$ → 1

$h(a)$
Activation function
(활성화함수)

→ 출력 y

티켓 구매 여부



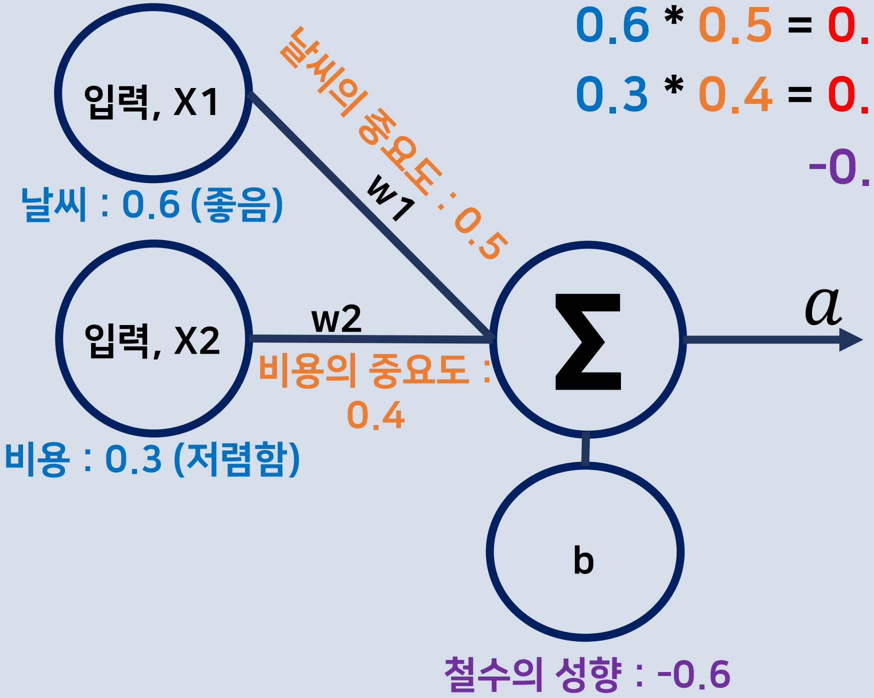
$$h(a) = \begin{cases} 0 & (a \leq 0) \\ 1 & (a > 0) \end{cases}$$

$$a = W_1X_1 + W_2X_2 + b$$
$$y = h(a)$$

Step function(계단함수)

퍼셉트론(단층)의 작동방식

ex) 콘서트장 예매 – 철수의 경우



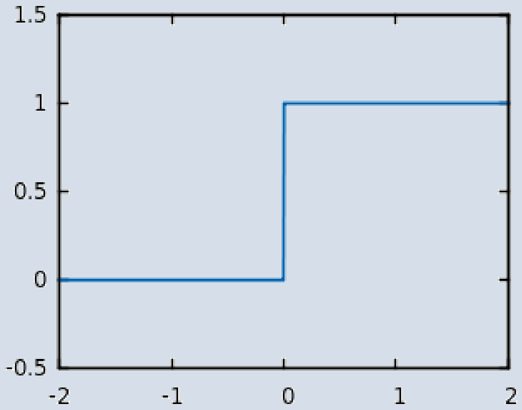
0.6 * 0.5 = 0.30
0.3 * 0.4 = 0.12
-0.60

0.42 - 0.60 = **-0.18** → $-0.18 \leq 0$ → 0

$h(a)$
Activation function
(활성화함수)

→ 출력 y

티켓 구매 여부



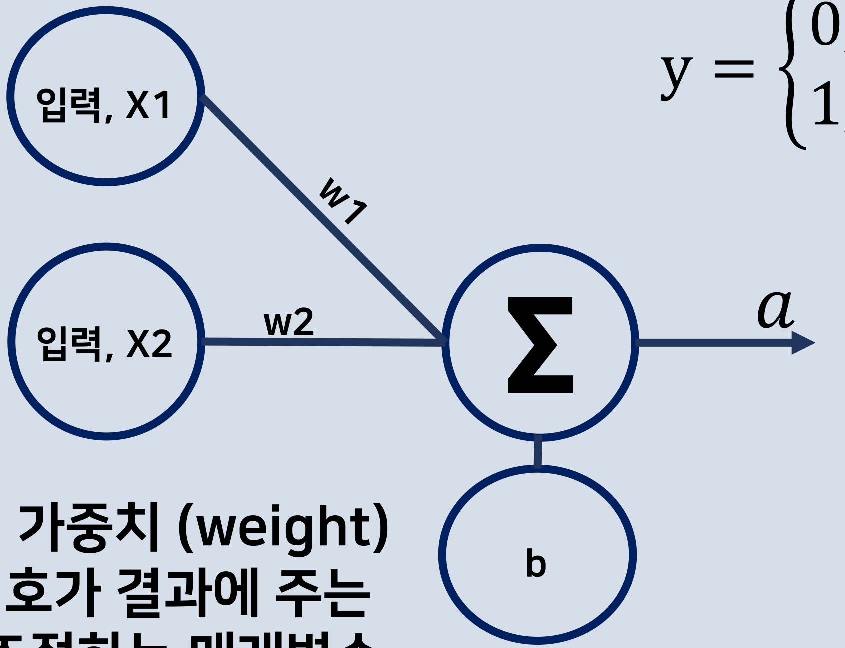
$$h(a) = \begin{cases} 0 & (a \leq 0) \\ 1 & (a > 0) \end{cases}$$

$$a = W_1X_1 + W_2X_2 + b$$
$$y = h(a)$$

Step function(계단함수)

퍼셉트론(단층)의 작동방식

$$y = \begin{cases} 0, & (W_1X_1 + W_2X_2 + b \le 0) \\ 1, & (W_1X_1 + W_2X_2 + b > 0) \end{cases}$$

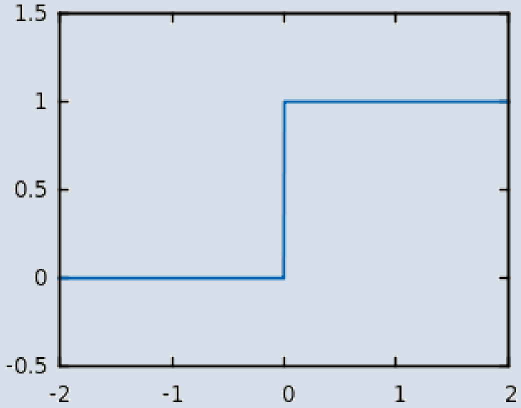


$h(a)$
Activation function
(활성화함수)

→ 출력 y

W_1, W_2: 가중치 (weight)
각 입력 신호가 결과에 주는
영향력을 조절하는 매개변수

b : 편향 (bias)
뉴런이 얼마나 쉽게
활성화하느냐를 조절하는 매개변수



Step function(계단함수)

Deep Learning

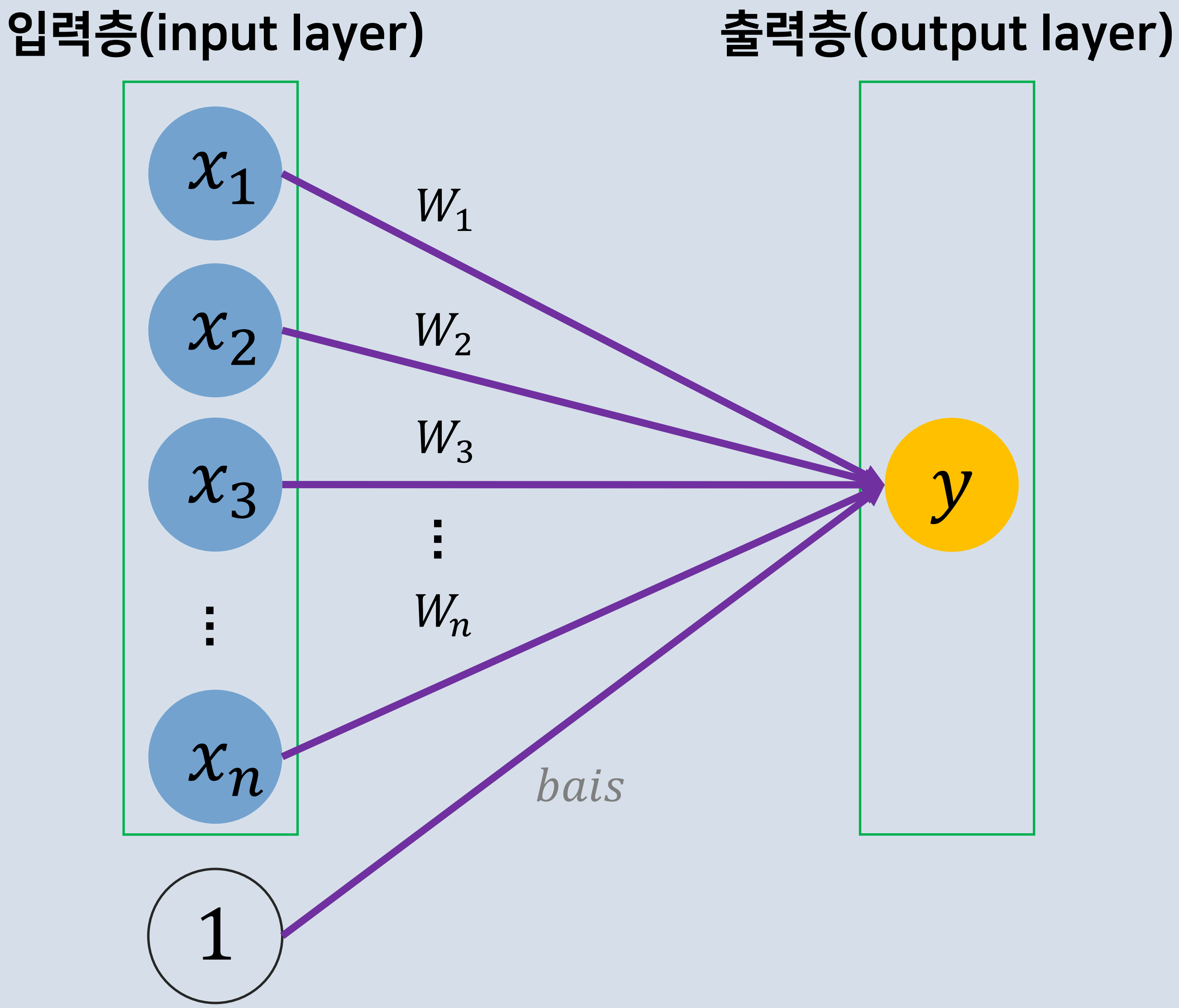
딥러닝 역사

- 1943년: ANN 개념 도입
- 1958년: Perceptron 제안
- 1969년: XOR 문제 지적
- 1986년: XOR 문제 해결 (MLP)
- 1986년: 파라미터 개수 증가 → 가중치, 편향 학습 어려움 → 오차역전파 알고리즘
- 1997년: LSTM
- 1998년: CNN
- 1990대: Vanishing Gradient/과적합/진동·발산 문제 대두
- 2000년 초: 머신러닝의 전진 (SVM, Random forest)
- 2006년: 과적합 문제 해결 →비지도학습-RBM→ DBN (딥러닝)제안
- 2007년: 자기부호화기
- 2013년: Dropout / Maxout (과적합 문제 해결)
- ReLU 적용 (Vanishing Gradient 문제 해결)
- 2014년: GRU (Vanishing Gradient 문제 해결)
- 2015년: DQN (Alphago)
- 2023년: ChatGPT(MS Bing)

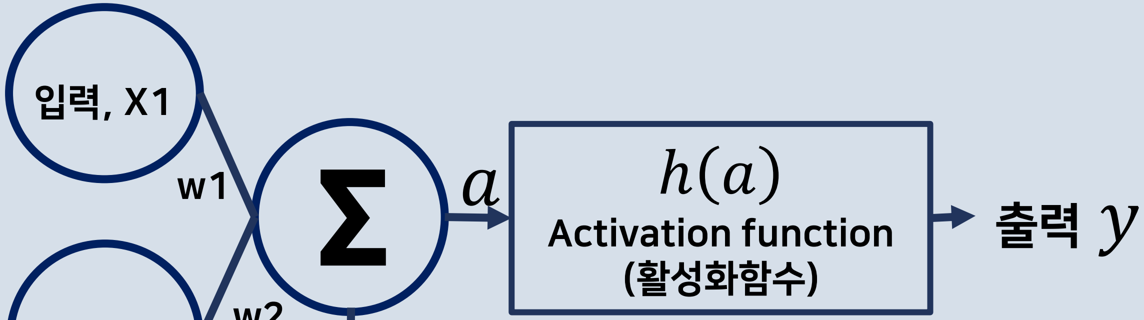
스마트인재개발원
Smart Human Resources Development

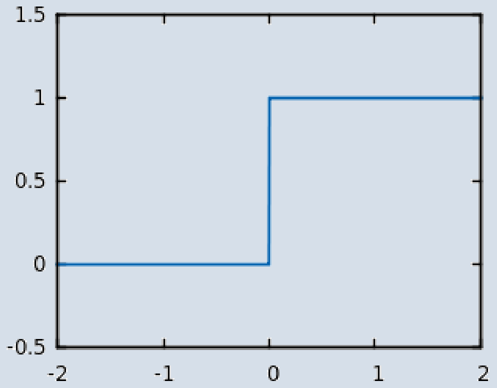
Deep Learning

단층 퍼셉트론의 한계 – XOR 문제



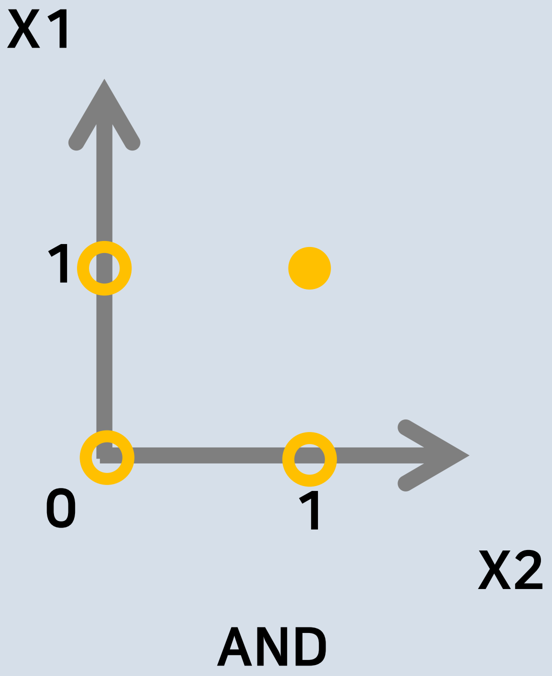
단층 퍼셉트론의 한계 – XOR 문제





AND 게이트

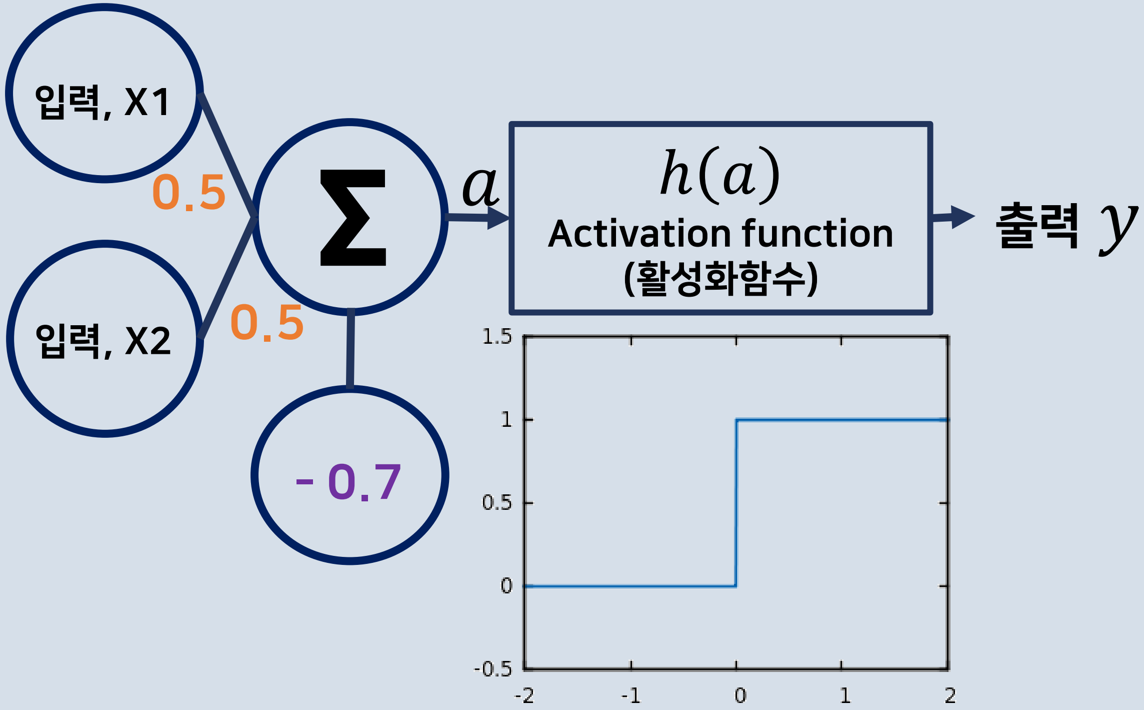
X1	X2	AND
0	0	0
0	1	0
1	0	0
1	1	1



$$y = \begin{cases} 0, & (W_1X_1 + W_2X_2 + b \le 0) \\ 1, & (W_1X_1 + W_2X_2 + b > 0) \end{cases}$$

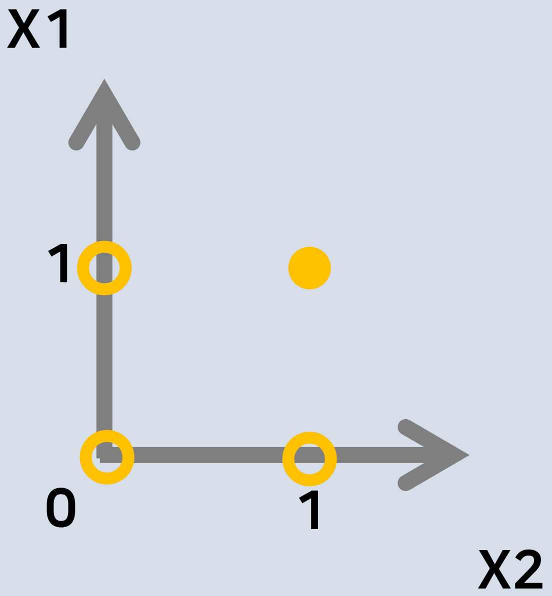
Deep Learning

단층 퍼셉트론의 한계 – XOR 문제



AND 게이트

X1	X2	AND
0	0	0
0	1	0
1	0	0
1	1	1



AND

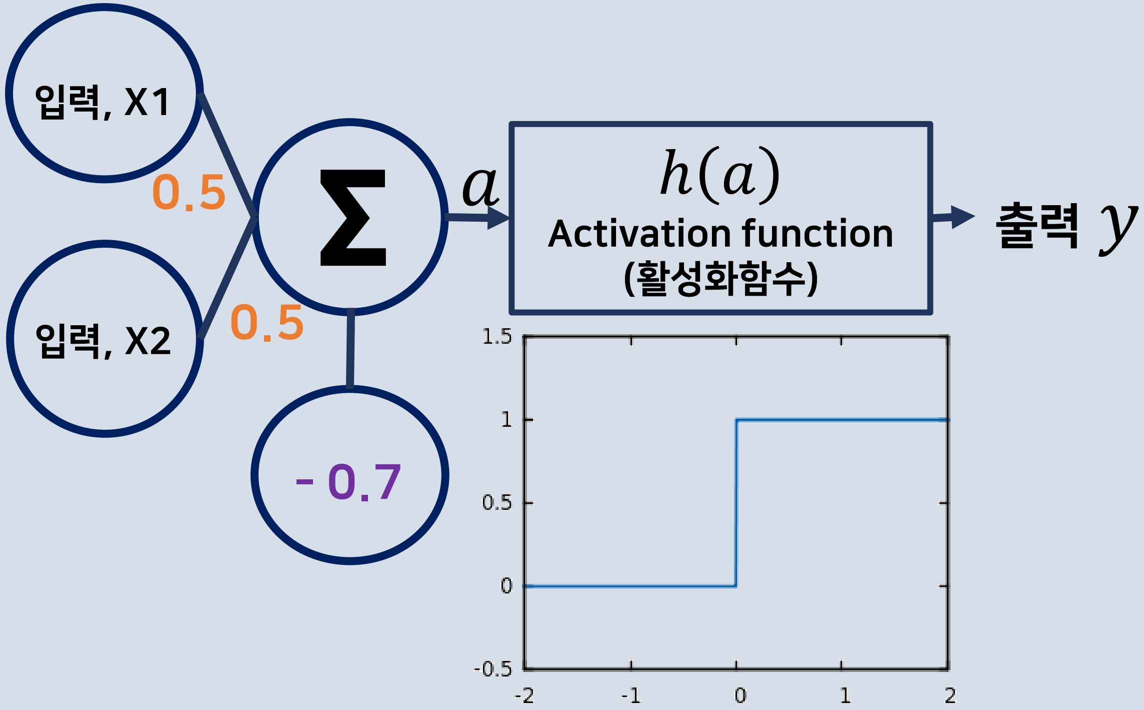
0.0 * 0.5 = 0.00
0.0 * 0.5 = 0.00
-0.70

0.00 - 0.70 = -0.70

$-0.70 \leq 0$

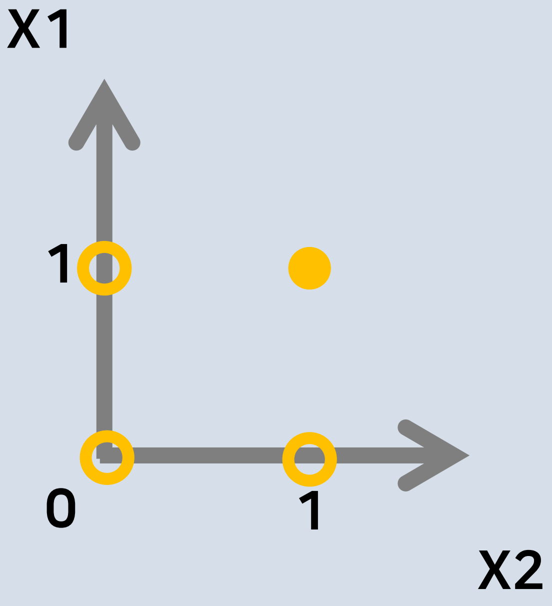
0

단층 퍼셉트론의 한계 – XOR 문제



AND 게이트

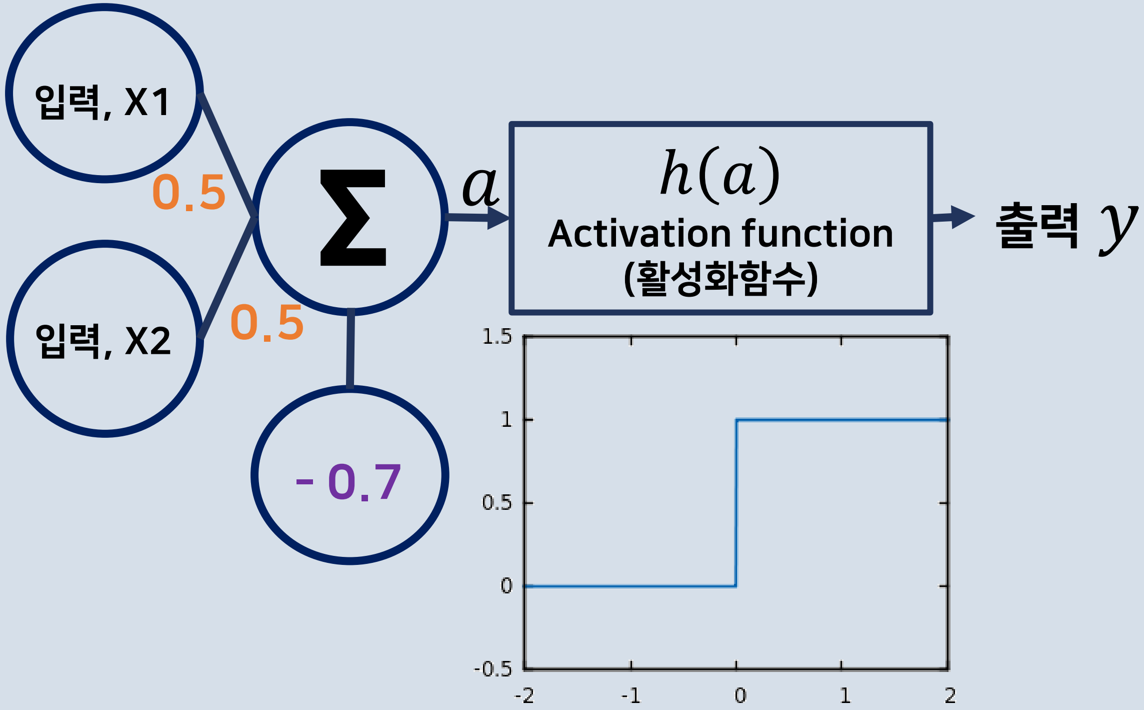
X1	X2	AND
0	0	0
0	1	0
1	0	0
1	1	1



AND

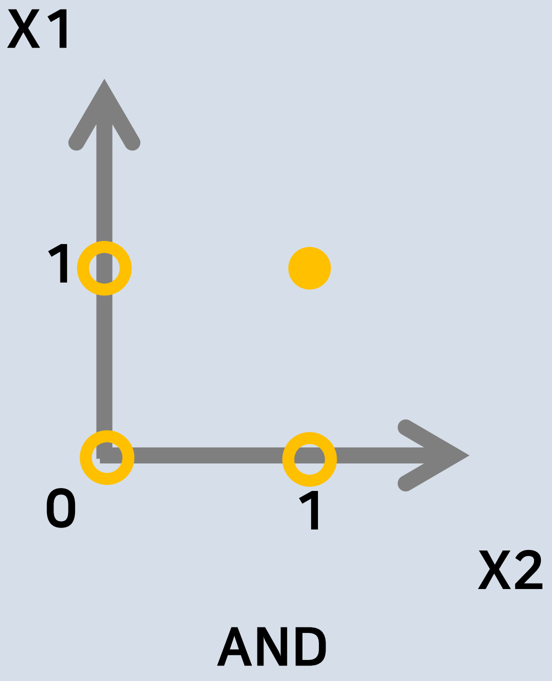
0.0 * 0.5 = 0.00
1.0 * 0.5 = 0.50 → 0.50 - 0.70 = -0.20 → $-0.20 \leq 0$ → 0
-0.70

단층 퍼셉트론의 한계 – XOR 문제



AND 게이트

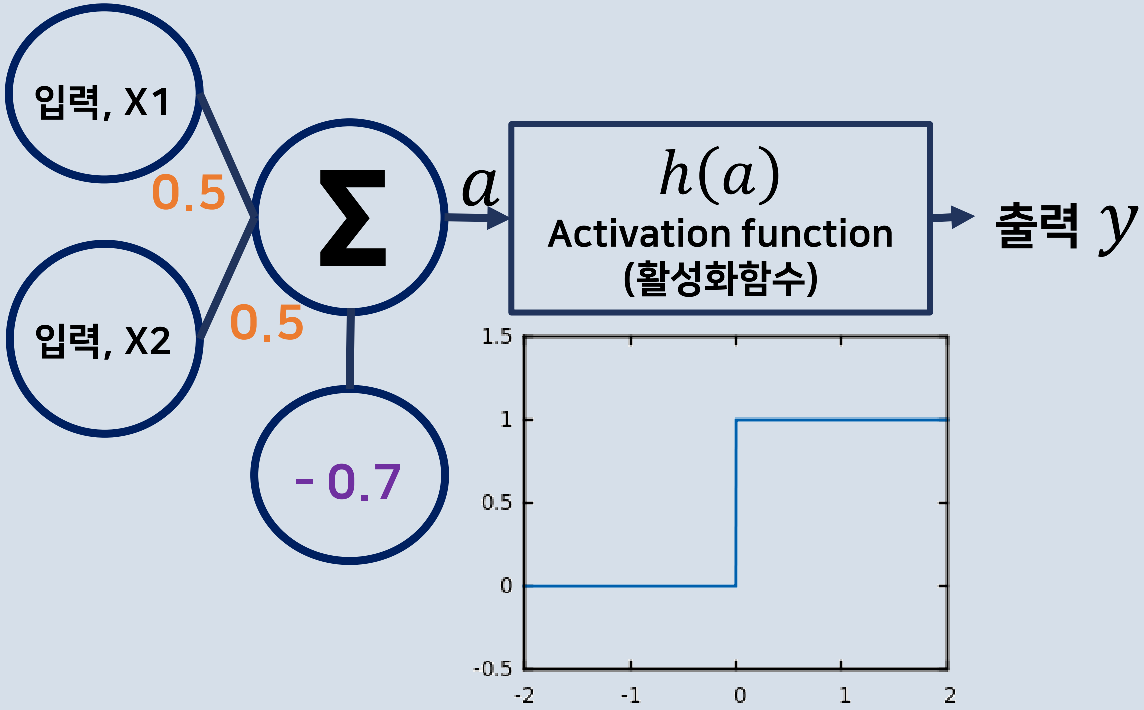
X1	X2	AND
0	0	0
0	1	0
1	0	0
1	1	1



1.0 * 0.5 = 0.50
0.0 * 0.5 = 0.00
-0.70

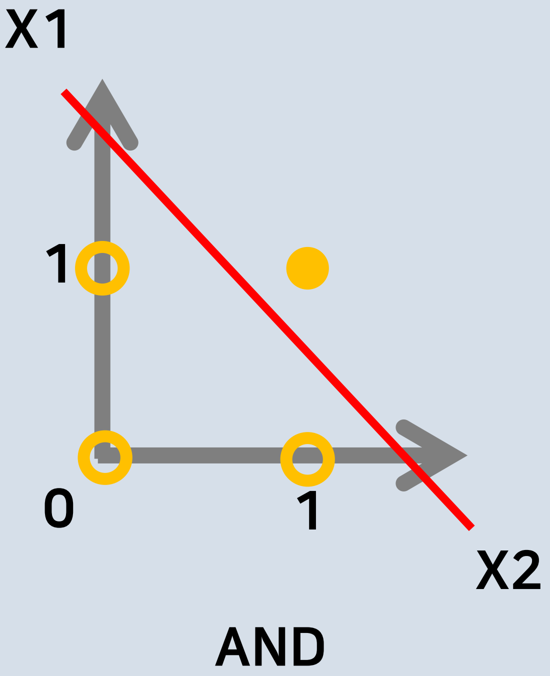
0.50 - 0.70 = -0.20 → $-0.20 \leq 0$ → 0

단층 퍼셉트론의 한계 – XOR 문제



AND 게이트

X1	X2	AND
0	0	0
0	1	0
1	0	0
1	1	1

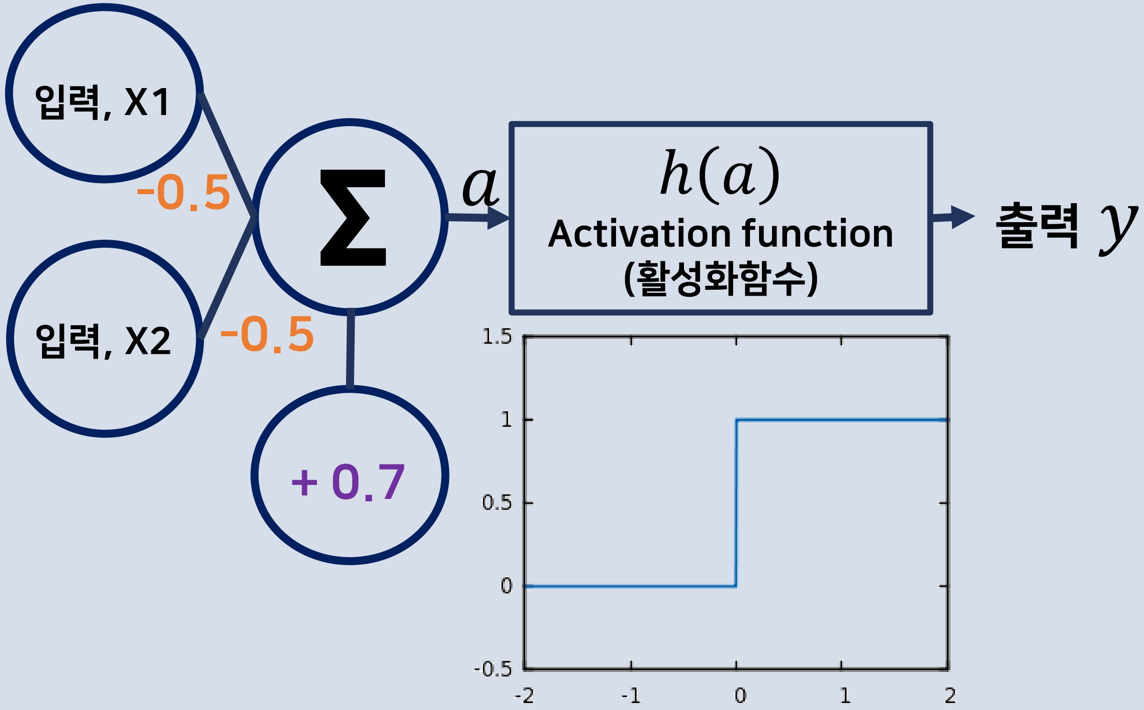


1.0 * 0.5 = 0.50
1.0 * 0.5 = 0.50
-0.70

→ 1.00 - 0.70 = 0.30 → $0.30 > 0$ → 1

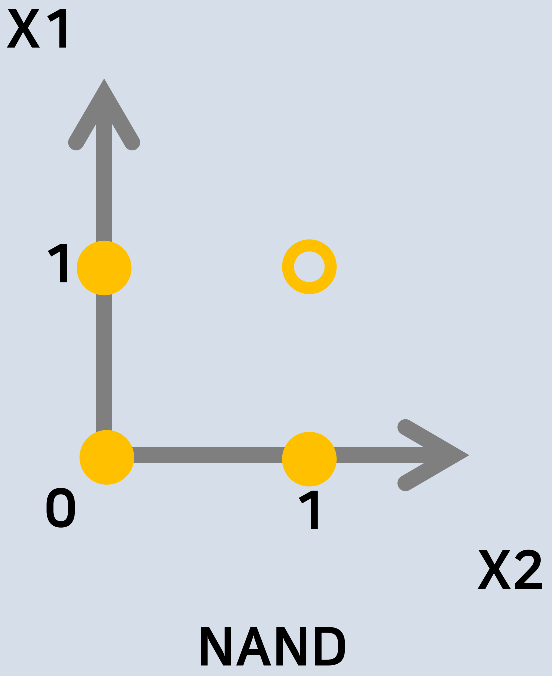
Deep Learning

단층 퍼셉트론의 한계 – XOR 문제



NAND 게이트

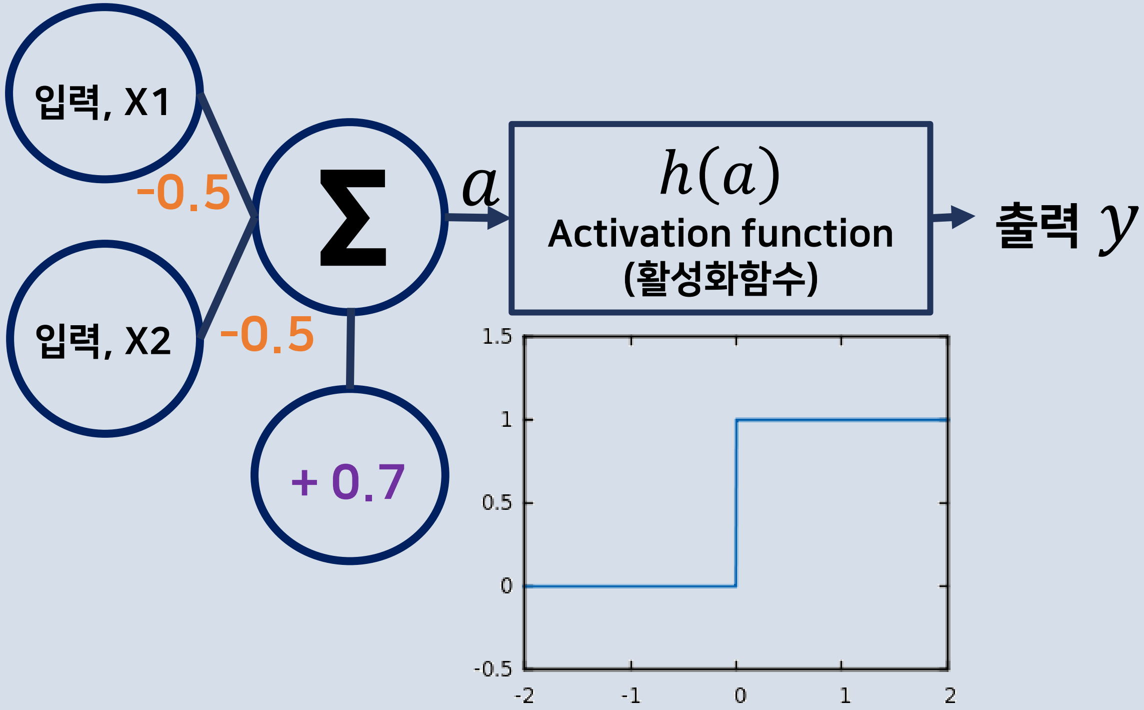
X1	X2	NAND
0	0	1
0	1	1
1	0	1
1	1	0



0.0 * -0.5 = 0.00
0.0 * -0.5 = 0.00
+ 0.70

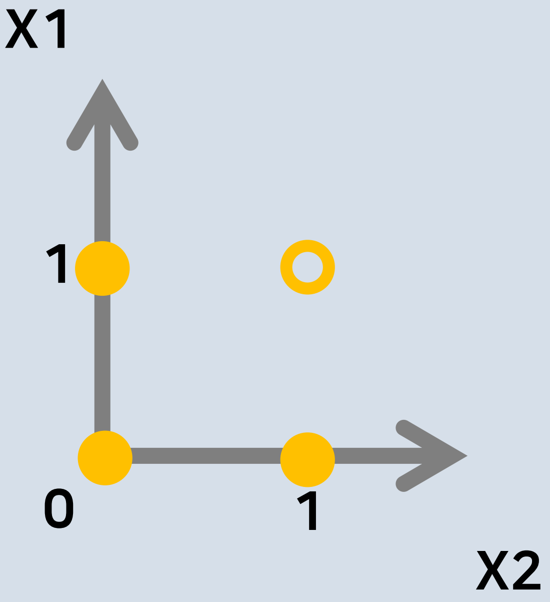
→ 0.00 + 0.70 = 0.70 → $0.70 > 0$ → 1

단층 퍼셉트론의 한계 – XOR 문제



NAND 게이트

X1	X2	NAND
0	0	1
0	1	1
1	0	1
1	1	0

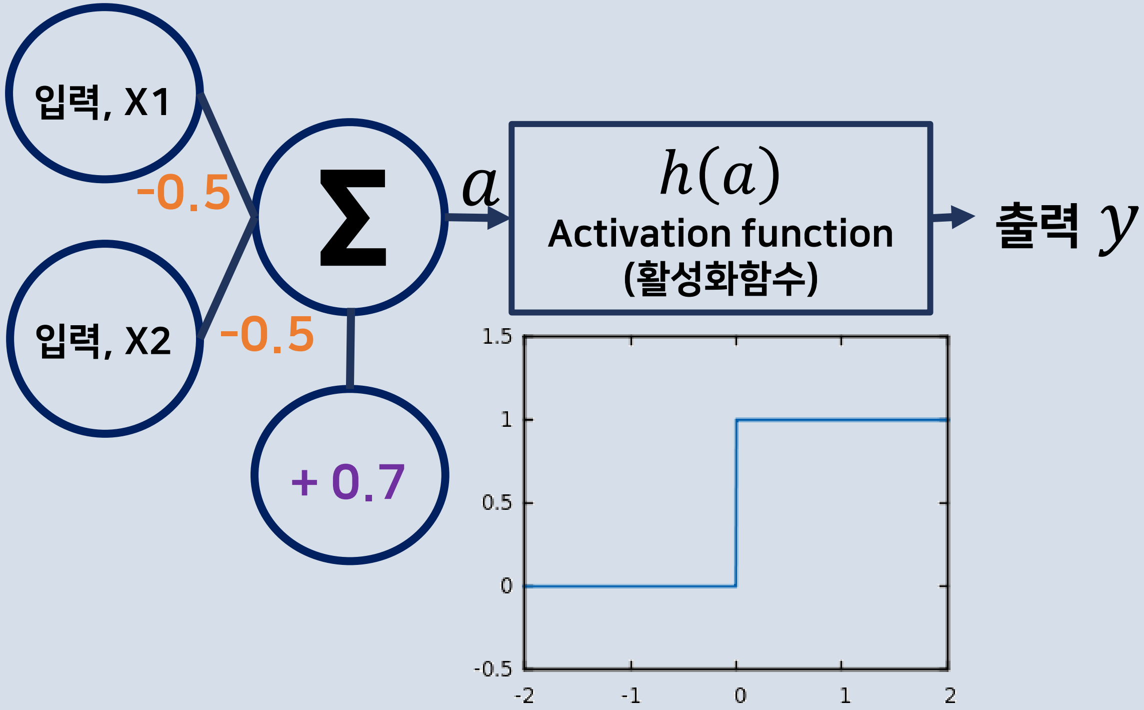


NAND

0.0 * -0.5 = 0.00
1.0 * -0.5 = -0.50 → -0.50 + 0.70 = 0.20 → $0.20 > 0$ → 1
+ 0.70

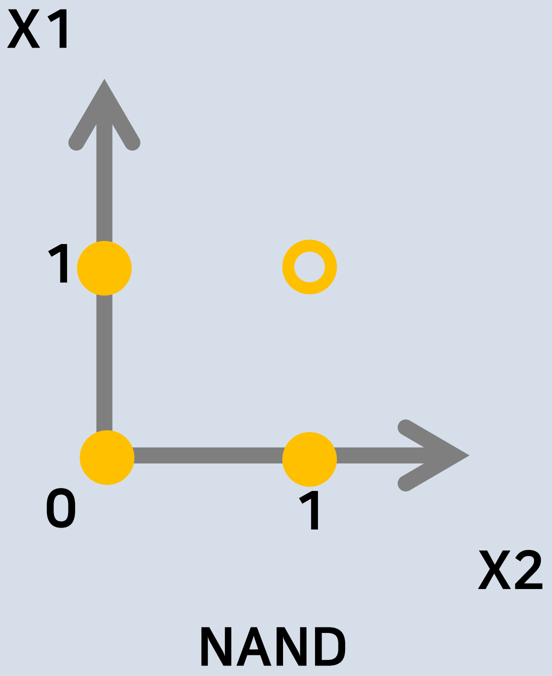
Deep Learning

단층 퍼셉트론의 한계 – XOR 문제



NAND 게이트

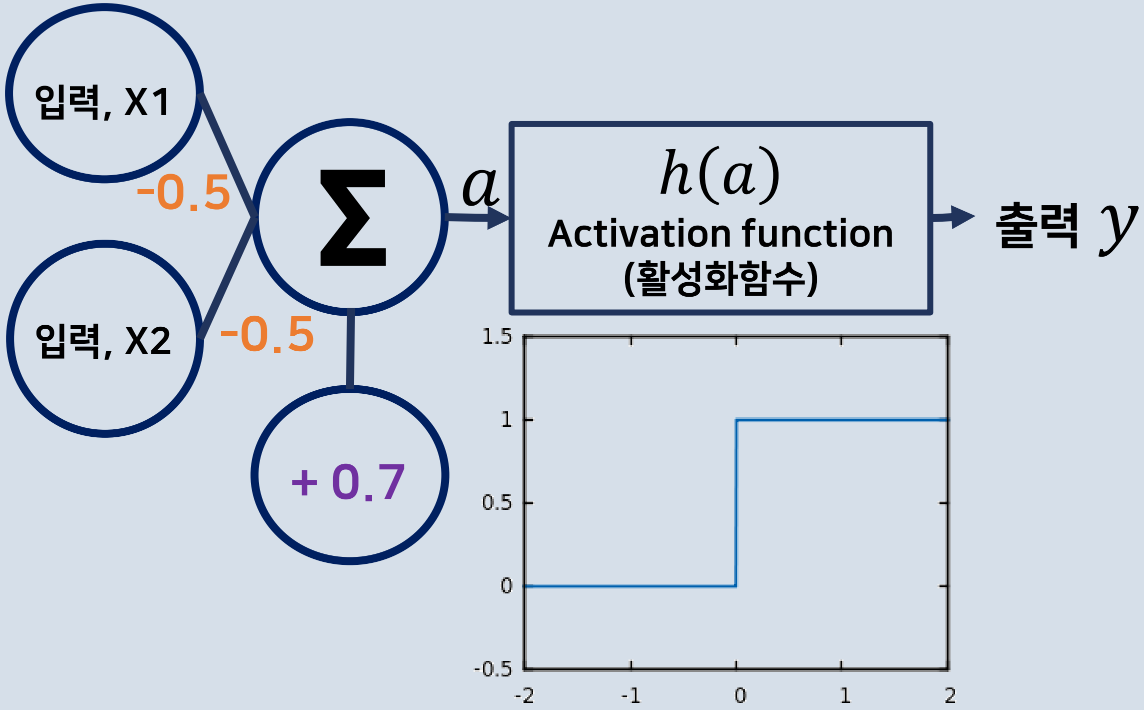
X1	X2	NAND
0	0	1
0	1	1
1	0	1
1	1	0



1.0 * -0.5 = -0.50
0.0 * -0.5 = 0.00
+ 0.70

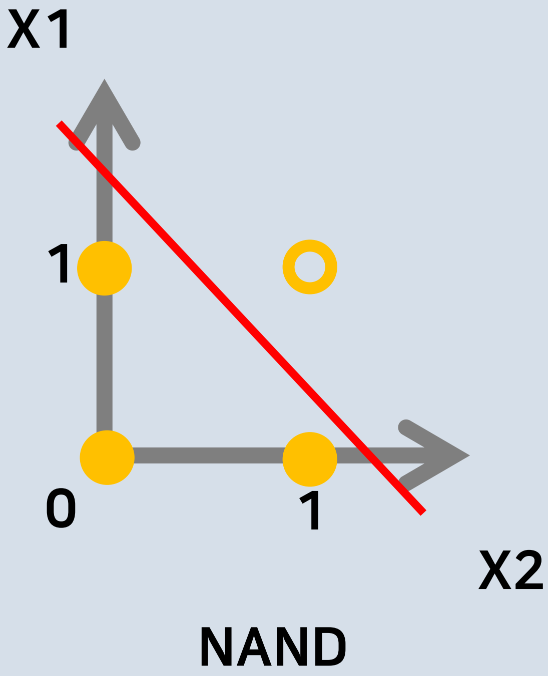
→ -0.50 + 0.70 = 0.20 → $0.20 > 0$ → 1

단층 퍼셉트론의 한계 – XOR 문제



NAND 게이트

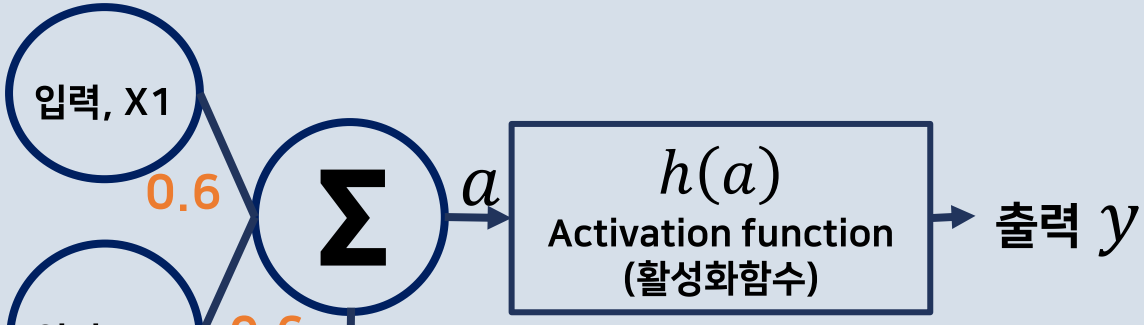
X1	X2	NAND
0	0	1
0	1	1
1	0	1
1	1	0

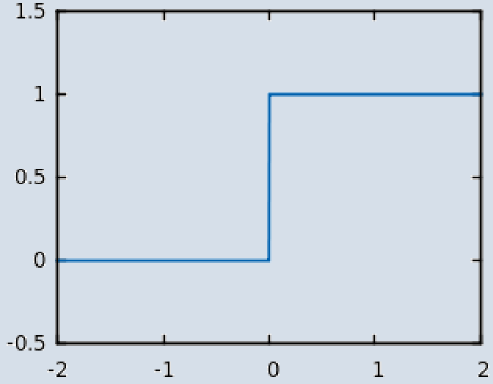


1.0 * -0.5 = -0.50
1.0 * -0.5 = -0.50
+ 0.70

→ -1.00 + 0.70 = -0.30 → $-0.30 \leq 0$ → 0

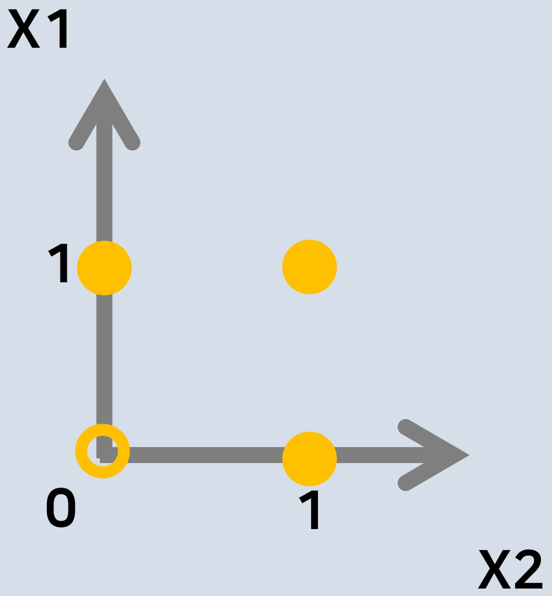
단층 퍼셉트론의 한계 – XOR 문제





OR 게이트

X1	X2	OR
0	0	0
0	1	1
1	0	1
1	1	1

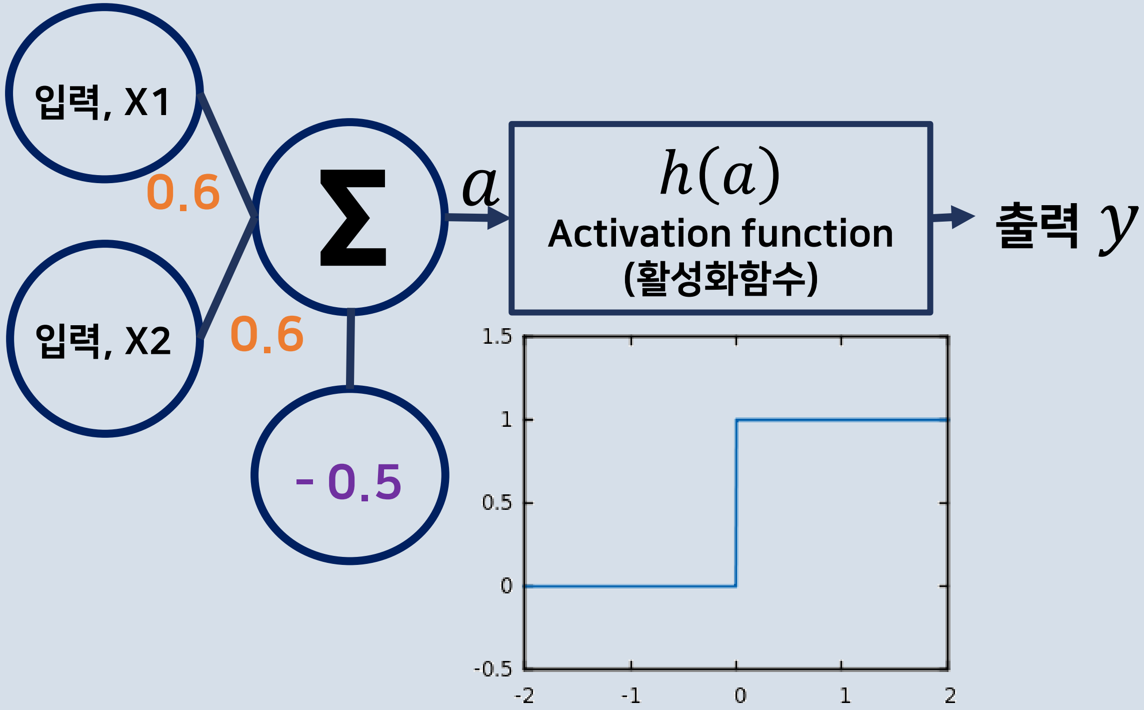


OR

0.0 * 0.6 = 0.00
0.0 * 0.6 = 0.00
- 0.50

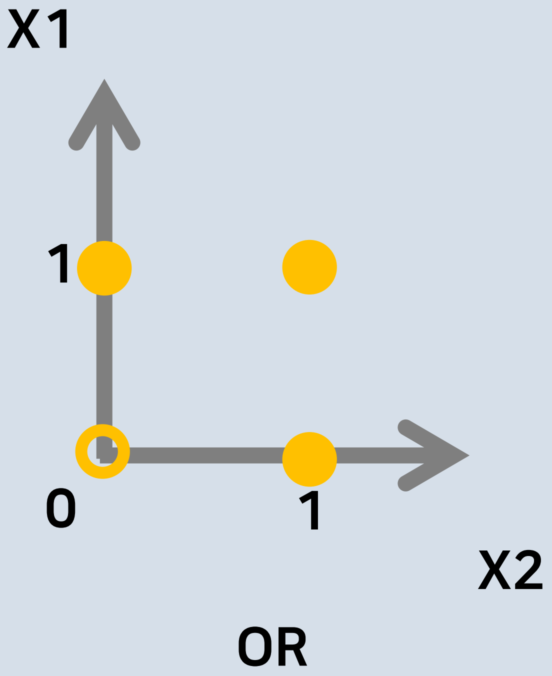
→ 0.00 - 0.50 = -0.50 → $-0.50 \leq 0$ → 0

단층 퍼셉트론의 한계 – XOR 문제



OR 게이트

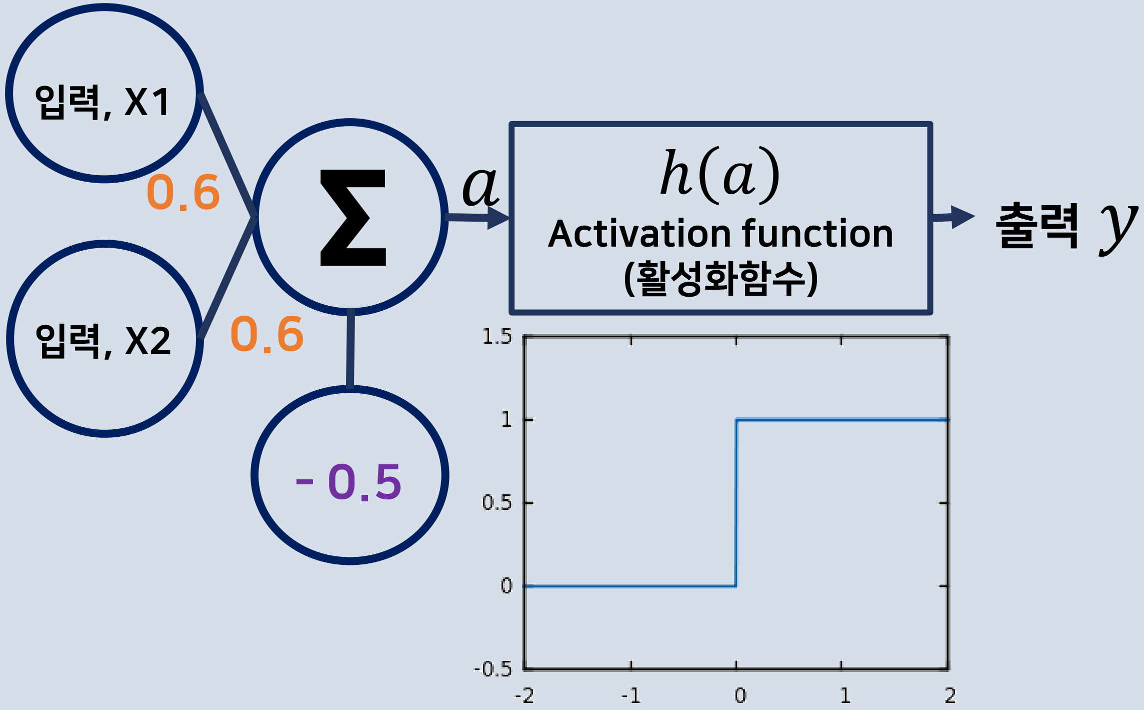
X1	X2	OR
0	0	0
0	1	1
1	0	1
1	1	1



0.0 * 0.6 = 0.00
1.0 * 0.6 = 0.60
- 0.50

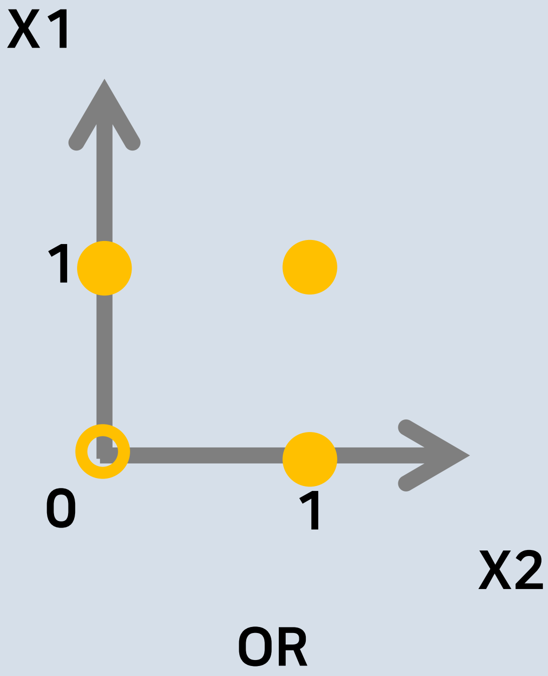
→ 0.60 - 0.50 = 0.10 → $0.10 > 0$ → 1

단층 퍼셉트론의 한계 – XOR 문제



OR 게이트

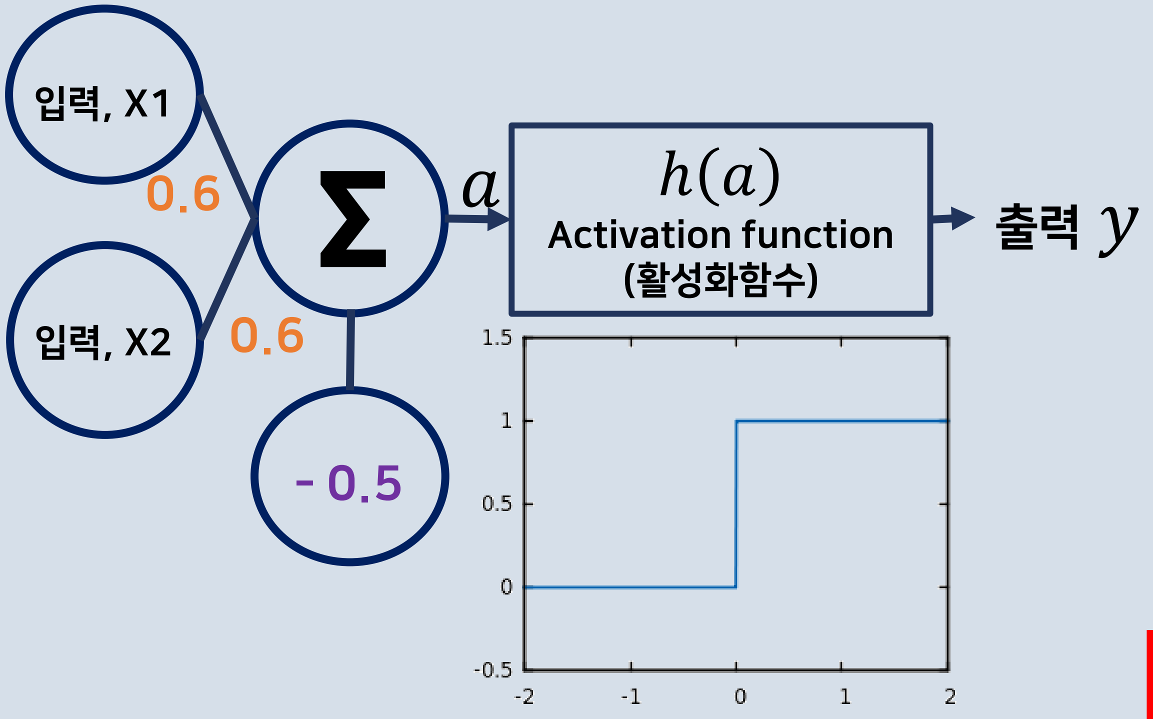
X1	X2	OR
0	0	0
0	1	1
1	0	1
1	1	1



1.0 * 0.6 = 0.60
0.0 * 0.6 = 0.00
- 0.50

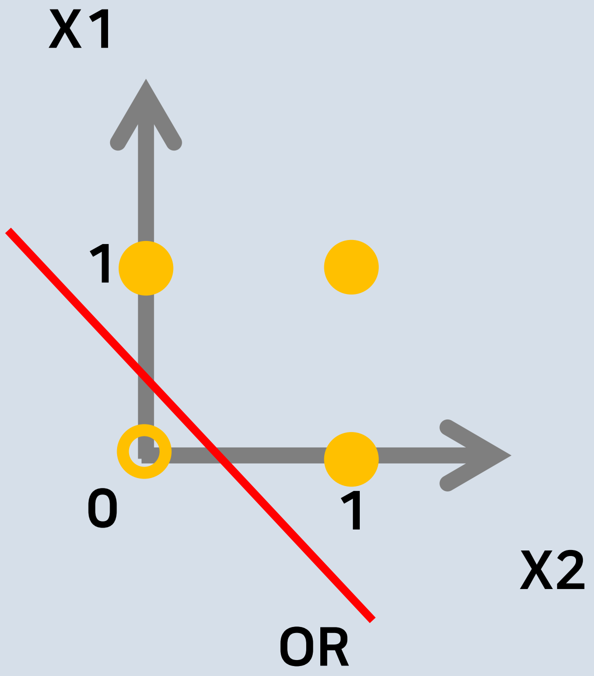
→ 0.60 - 0.50 = 0.10 → $0.10 > 0$ → 1

단층 퍼셉트론의 한계 – XOR 문제



OR 게이트

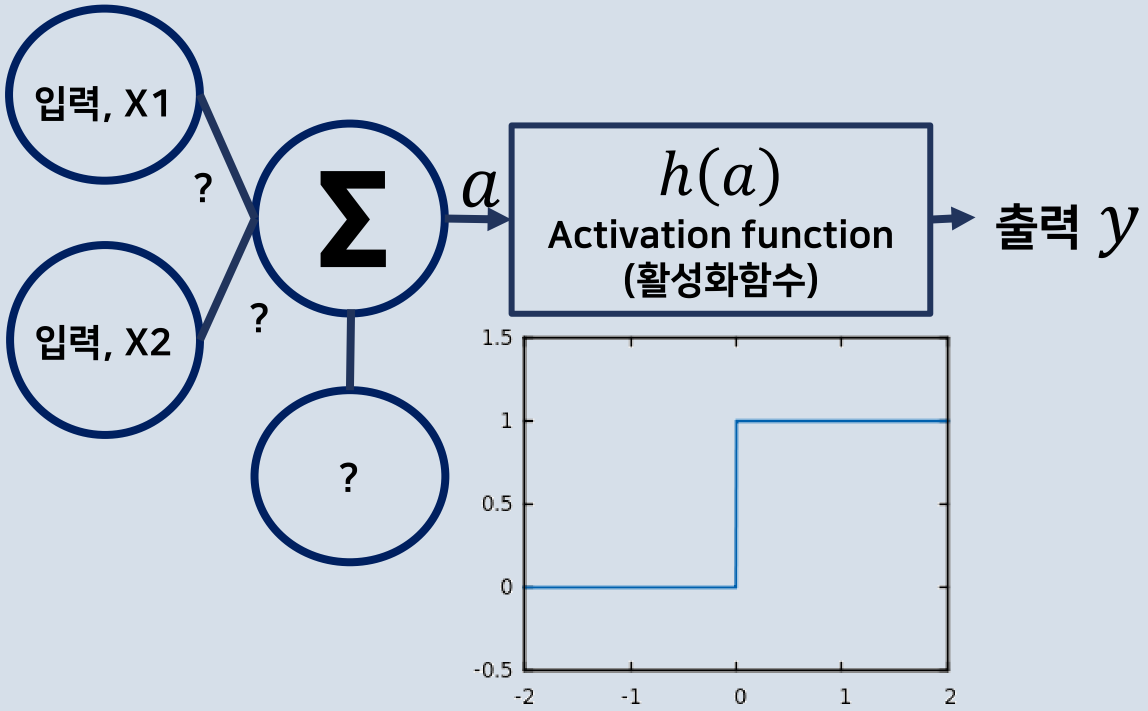
X1	X2	OR
0	0	0
0	1	1
1	0	1
1	1	1



1.0 * 0.6 = 0.60
1.0 * 0.6 = 0.60
- 0.50

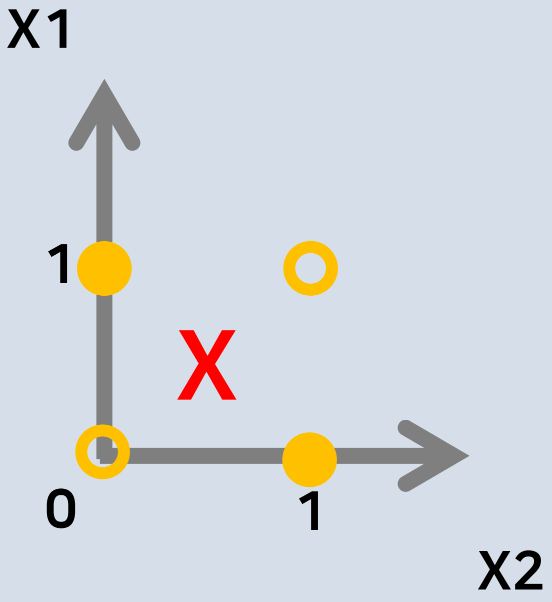
→ 1.20 - 0.50 = 0.70 → $0.70 > 0$ → 1

단층 퍼셉트론의 한계 – XOR 문제



XOR 게이트

X1	X2	OR
0	0	0
0	1	1
1	0	1
1	1	0

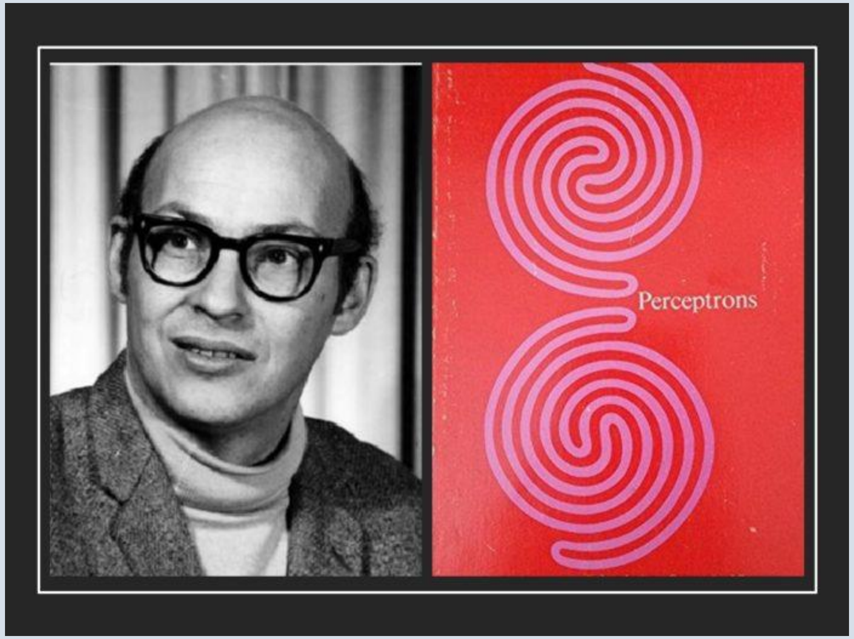


XOR

단층 퍼셉트론의 한계 – XOR 문제

AND,OR는 해결이 가능하지만

간단한 XOR 문제를 해결 할 수 없었다.



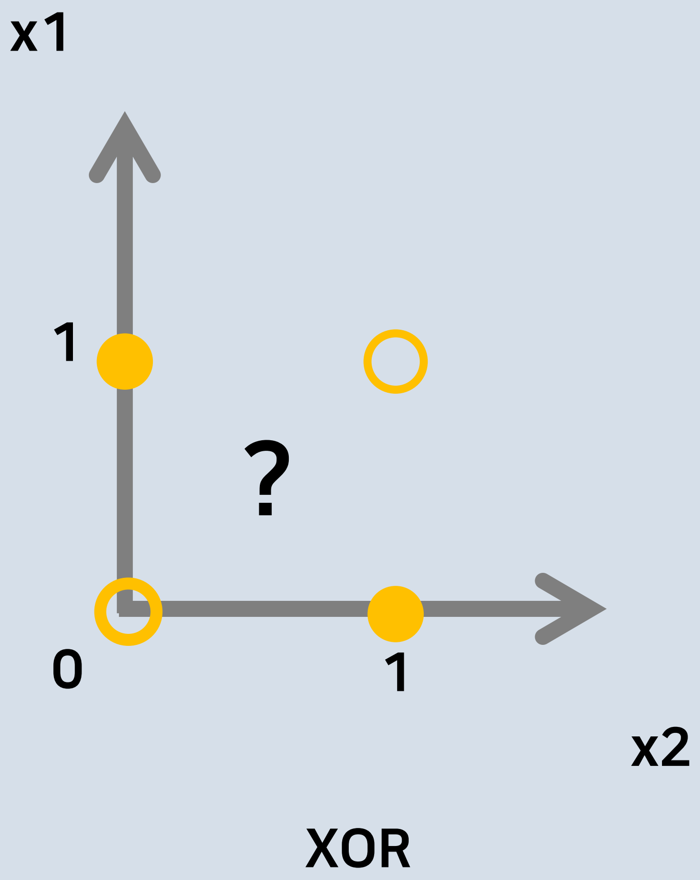
출처 : 마빈 민스키와 퍼셉트론의 문제를 지적한 저서 <Perceptrons>

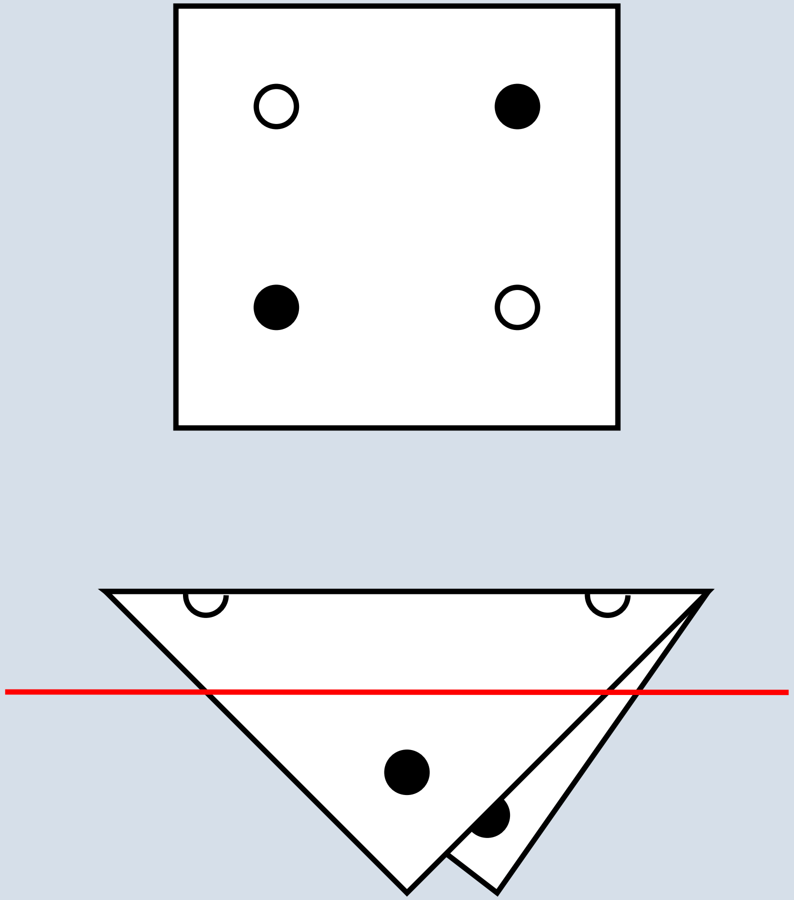
Deep Learning

딥러닝 역사

- 1943년: ANN 개념 도입
- 1958년: Perceptron 제안
- 1969년: XOR 문제 지적
- 1986년: XOR 문제 해결 (MLP)
- 1986년: 파라미터 개수 증가 → 가중치, 편향 학습 어려움 → 오차역전파 알고리즘
- 1997년: LSTM
- 1998년: CNN
- 1990대: Vanishing Gradient/과적합/진동·발산 문제 대두
- 2000년 초: 머신러닝의 전진 (SVM, Random forest)
- 2006년: 과적합 문제 해결 →비지도학습-RBM→ DBN (딥러닝)제안
- 2007년: 자기부호화기
- 2013년: Dropout / Maxout (과적합 문제 해결)
- 2014년: ReLU 적용 (Vanishing Gradient 문제 해결)
- 2014년: GRU (Vanishing Gradient 문제 해결)
- 2015년: DQN (Alphago)
- 2023년: ChatGPT(MS Bing)

딥러닝 역사 – XOR 문제 해결(MLP)



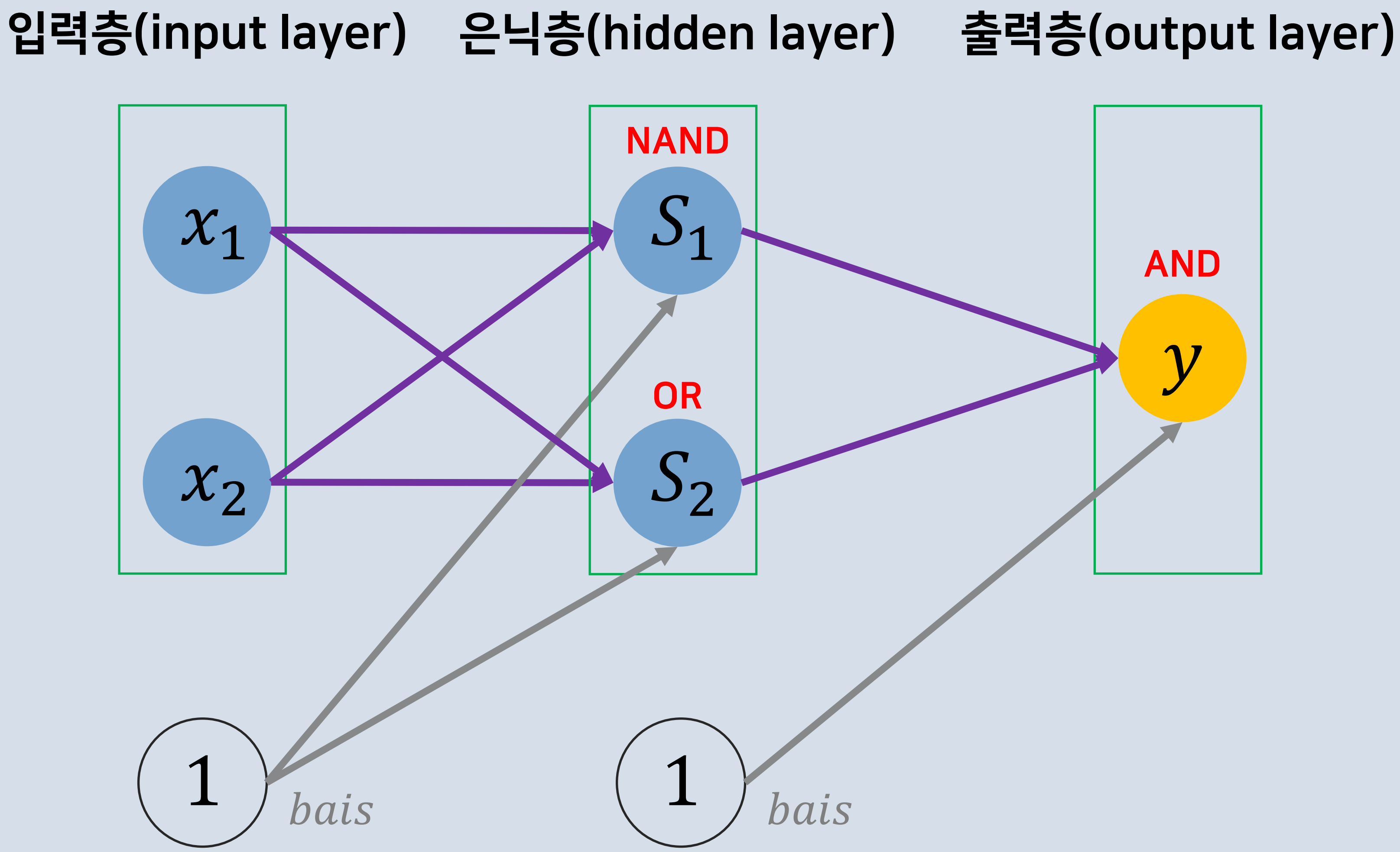


딥러닝 역사 – XOR 문제 해결(MLP)

다층 퍼셉트론(Multi Layer Perceptron)

단층 퍼셉트론의 차원 수를 확장시켜
여러 개의 층으로 구성하여 만든 신경망.

딥러닝 역사 – XOR 문제 해결(MLP)



딥러닝 역사 – XOR 문제 해결(MLP)

x1	x2	NAND
0	0	1
0	1	1
1	0	1
1	1	0

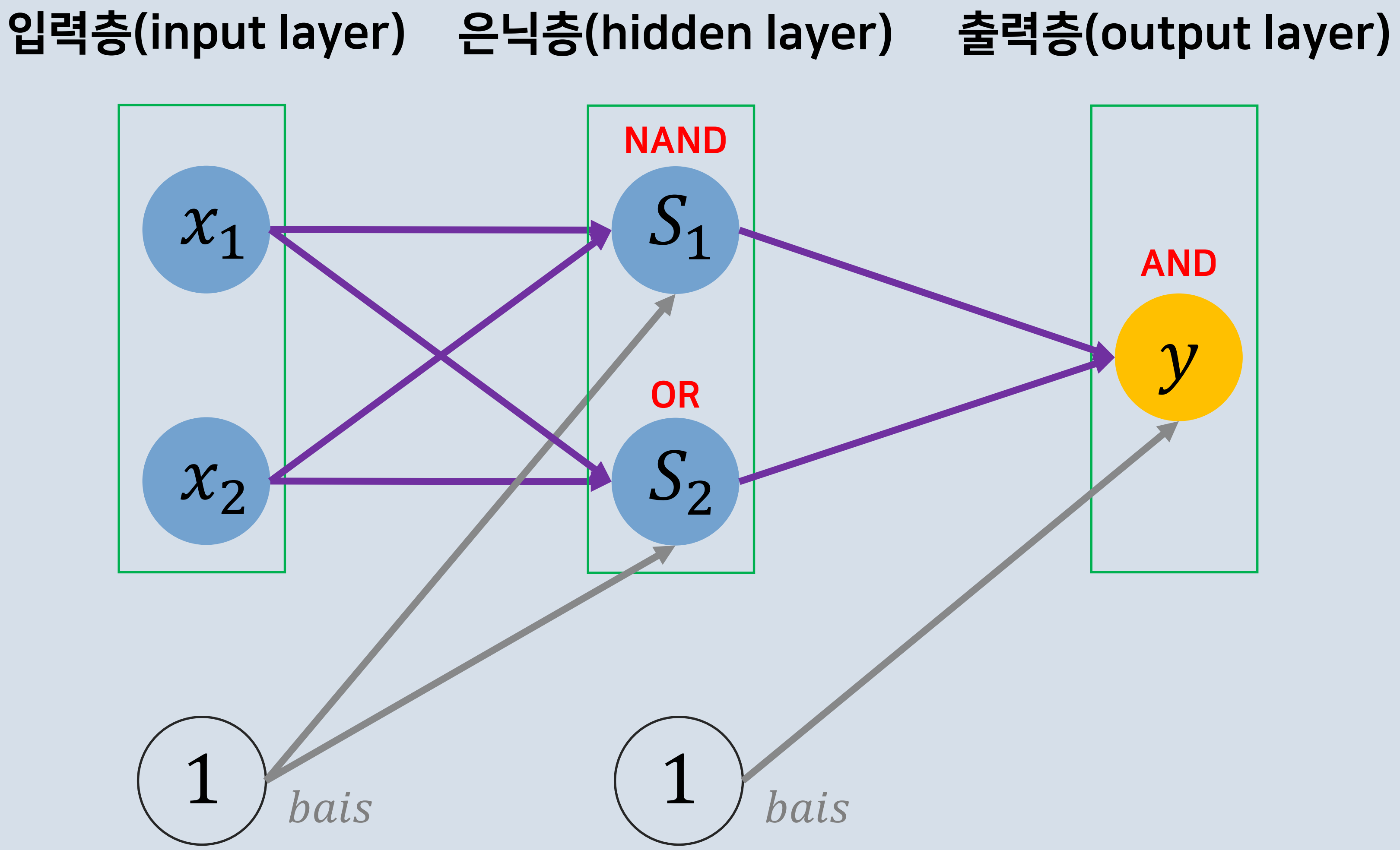
x1	x2	OR
0	0	0
0	1	1
1	0	1
1	1	1

NAND	OR	AND
1	0	0
1	1	1
1	1	1
0	1	0

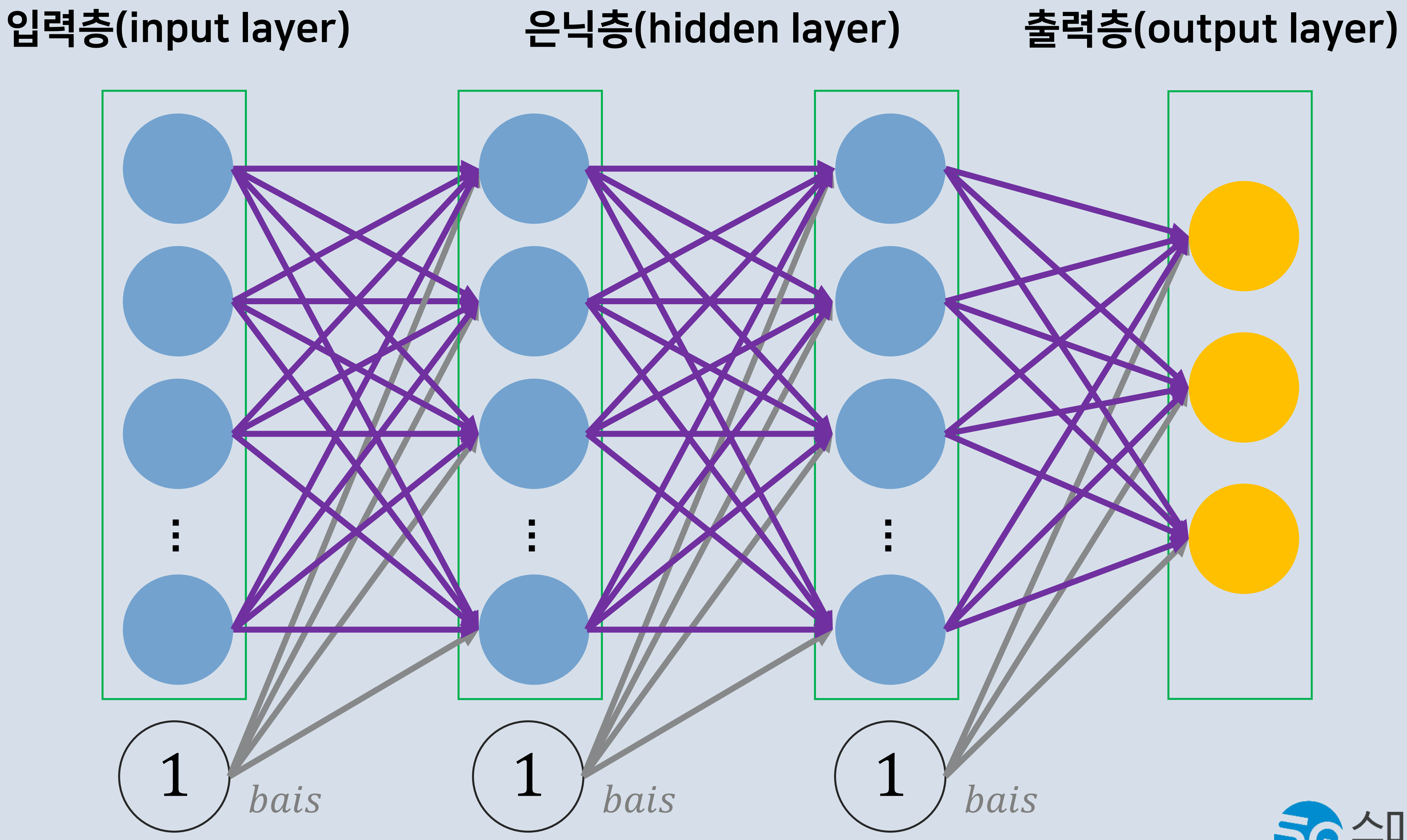
==

XOR
0
1
1
0

다층 퍼셉트론(Multilayer Perceptron)



다층 퍼셉트론(Multilayer Perceptron)



다층 퍼셉트론(Multilayer Perceptron)

- 한 번의 연산으로 해결되지 않는 문제를 해결할 수 있다.
- 단층에 비해 학습시간이 오래 걸린다.
- 모델(신경망)이 복잡해지고 가중치 파라미터가 많아 학습 시 과대적합되기 쉽다.

keras 맛보기 : 유방암 데이터 예측 실습