

ADVANCED TEXT MINING

by FINGEREDMAN (fingeredman@gmail.com)

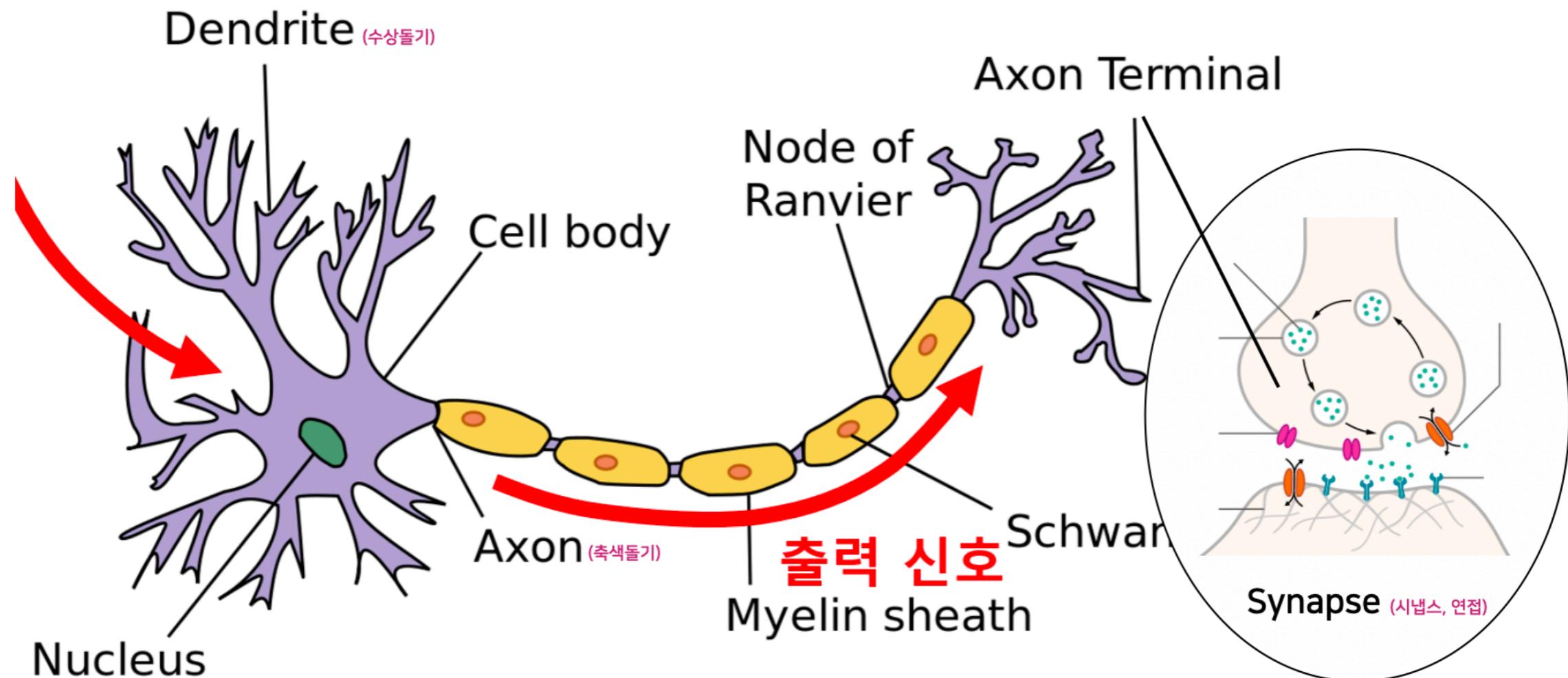
WEEK 08

Deep-learning for Text Mining

딥러닝 (Deep Learning)

인공신경망 (Artificial Neural Networks, ANN)

- 인간의 뇌를 구성하는 **뉴런(neuron)**의 구조를 모방한 연결 구조의 연산을 통해 문제를 해결하는 머신러닝 방법
- 뉴런은 수상돌기를 통해 입력신호를 받아 임계치 이상의 신호로 식별되는 경우 축색돌기를 따라 출력신호를 생성하여 다음 연결된 뉴런으로 신호를 전달하는 역할을 함



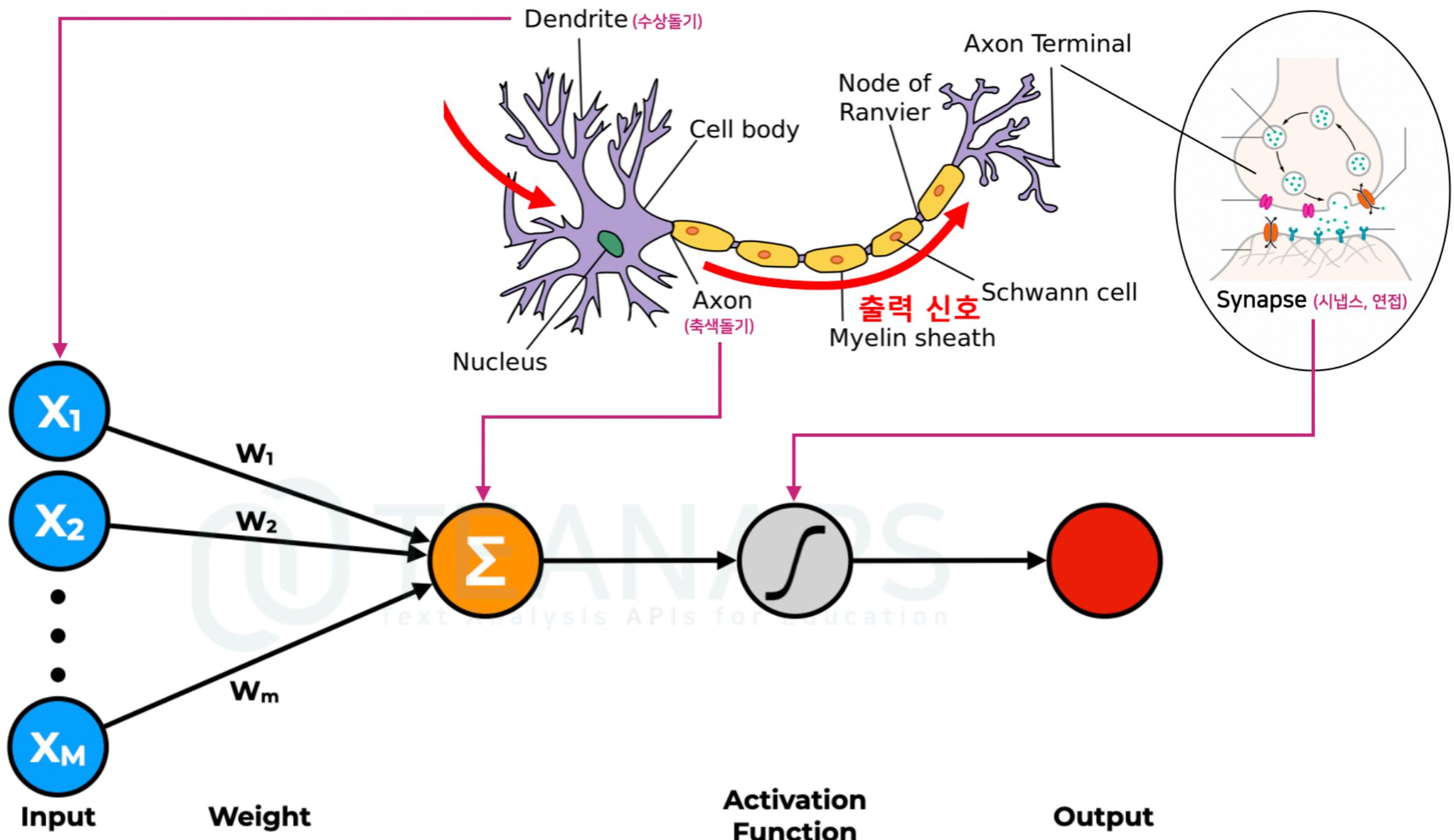
What is the difference?

Artificial Neural Network (ANN)

- ≈ Multi-Layer Perceptron (MLP)
- ≈ Deep Neural Network (DNN)
- ≈ Deep Feed-forward Networks
- ≈ Feed-forward neural Networks

딥러닝 (Deep Learning)

퍼셉트론 (Perceptron)



* 미프로박재우의원, 유전성 말초신경병증 시술, https://stems88.cafe24.com/?page_id=14798/.

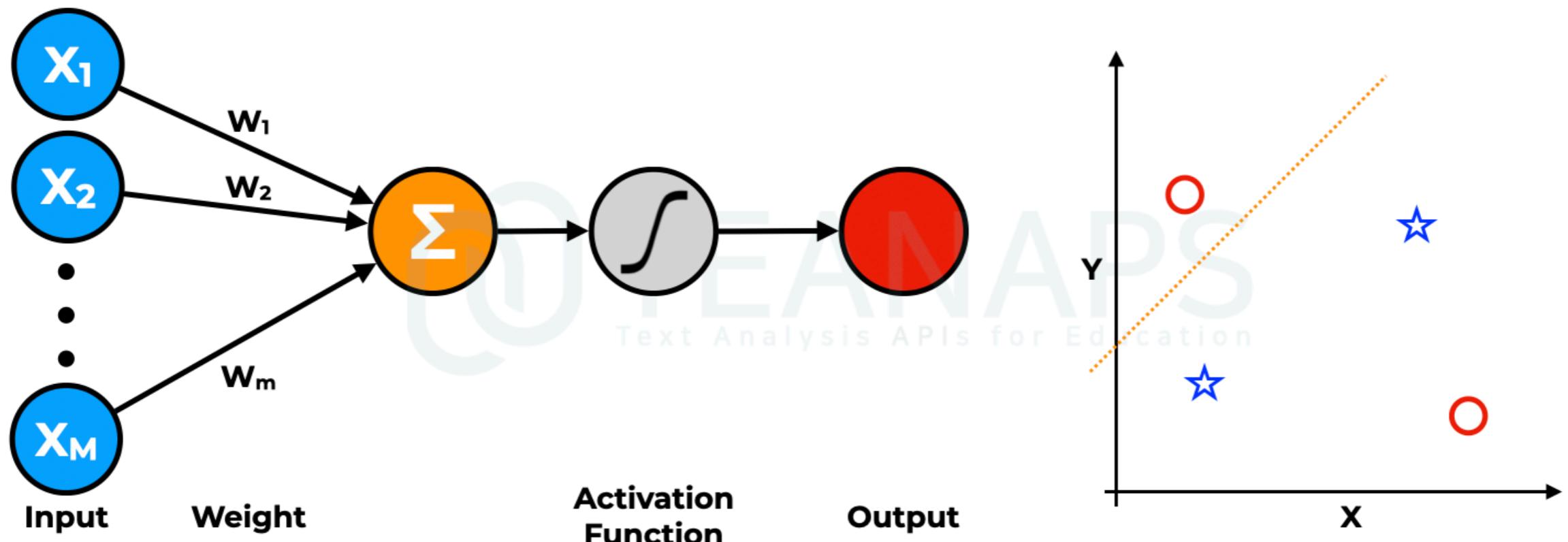
** 위키피디아, Neurotransmission, <https://en.wikipedia.org/wiki/Neurotransmission/>.

*** references

딥러닝 (Deep Learning)

다중 퍼셉트론 (Multi-layer Perceptron)

- 퍼셉트론은 분류문제 해결을 위한 알고리즘 중에 하나로 제안되어 좋은 평가를 받았으나^(Franky, 1957), 적용대상이 1차 선형분류 문제에 한정되어 있다는 한계점이 있음^(Minsky & Papert, 1969)

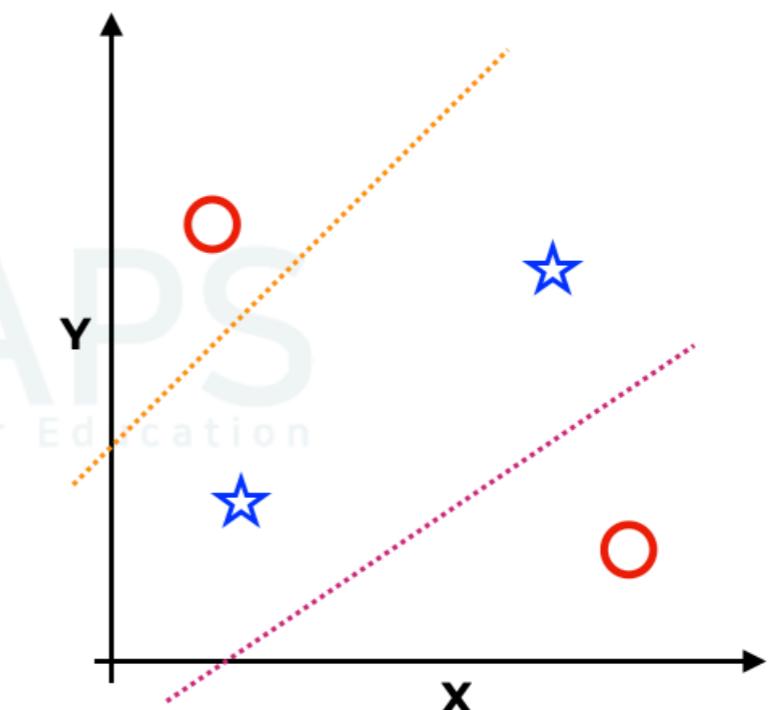
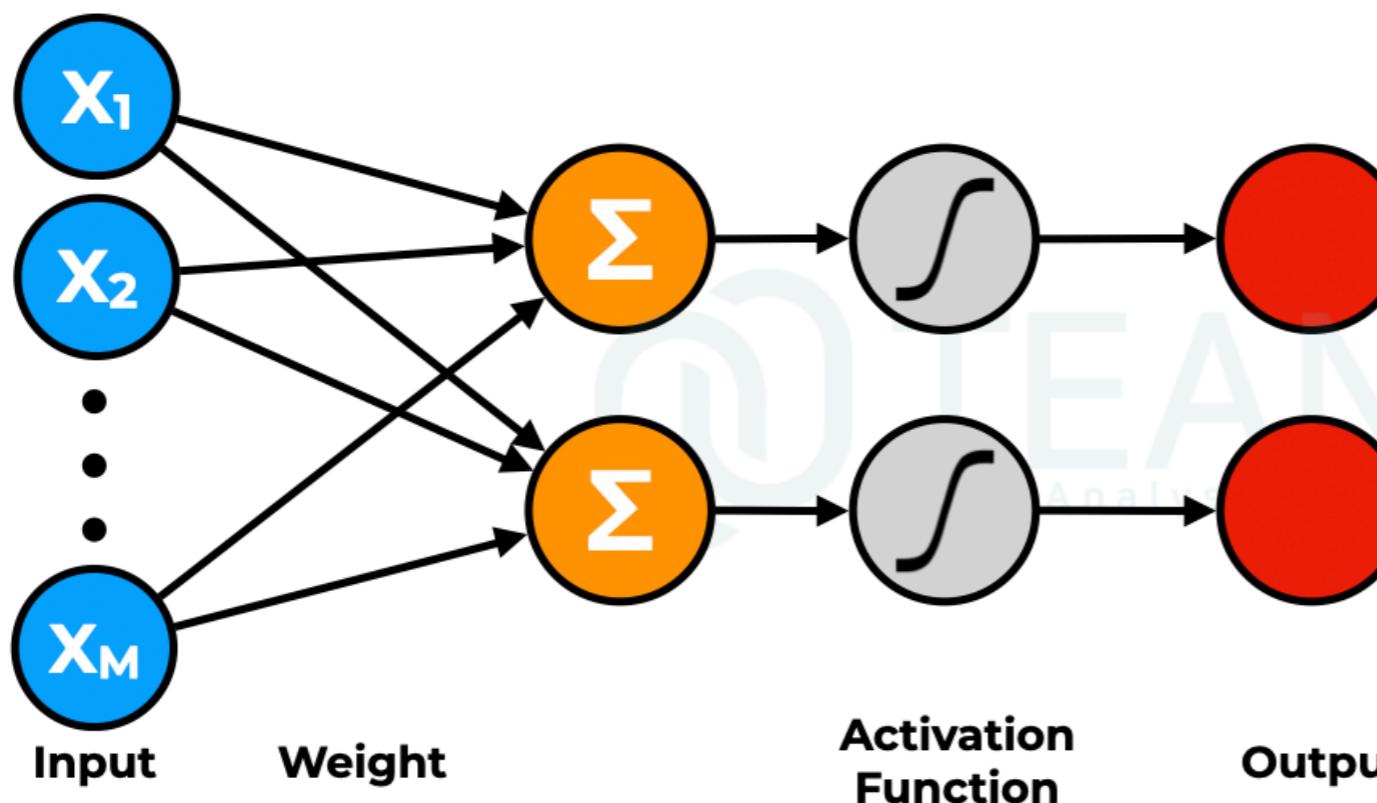


$$Y_n = W_1 \cdot X_1 + W_2 \cdot X_2 + \cdots + W_m \cdot X_m \text{ (직선의 방정식)}$$

딥러닝 (Deep Learning)

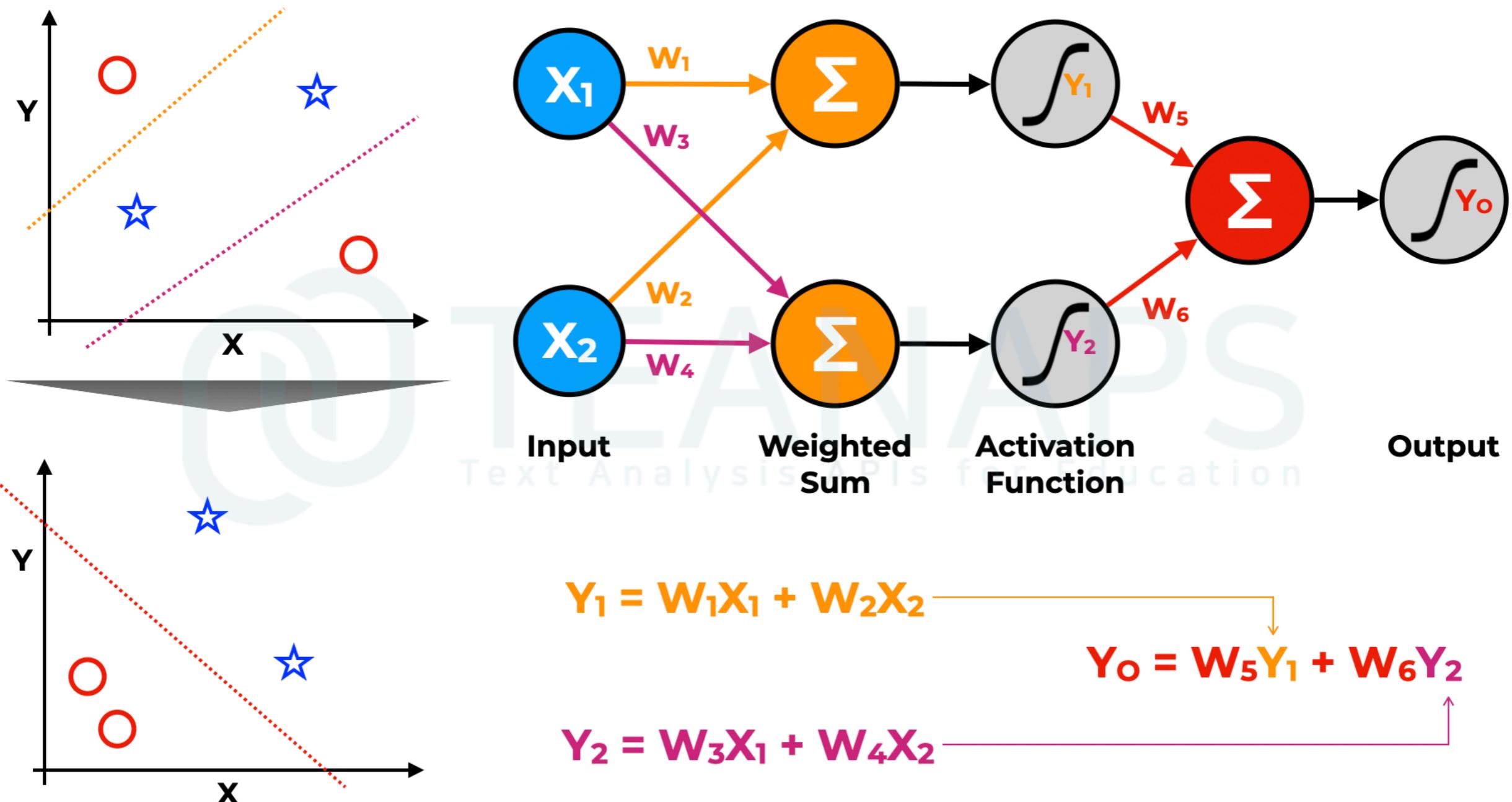
다중 퍼셉트론 (Multi-layer Perceptron)

- 퍼셉트론은 분류문제 해결을 위한 알고리즘 중에 하나로 제안되어 좋은 평가를 받았으나 (Franky, 1957), 적용대상이 1차 선형분류 문제에 한정되어 있다는 한계점이 있음 (Minsky & Papert, 1969)
- 퍼셉트론의 층을 여러개 쌓아 하나의 직선으로 해결할 수 없는 분류문제를 다수의 직선을 활용해 해결하는 다중 퍼셉트론 알고리즘이 제안됨 (Rumelhart, Hinton & Williams, 1986)



딥러닝 (Deep Learning)

다중 퍼셉트론 (Multi-layer Perceptron)

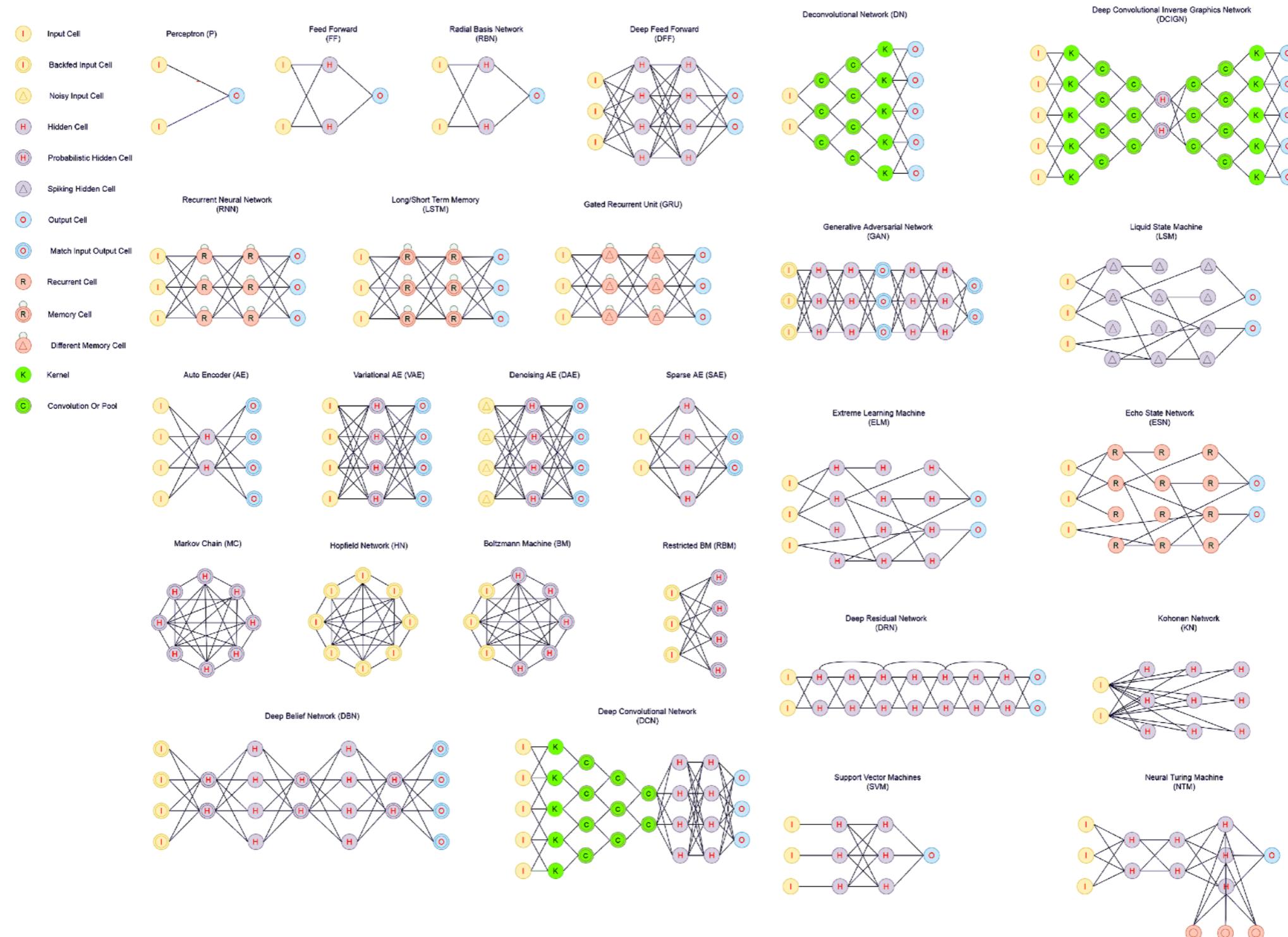


What is the difference?

Artificial Neural Network (ANN)

- ≈ Multi-Layer Perceptron (MLP)
- ≈ Deep Neural Network (DNN)
- ≈ Deep Feed-forward Networks
- ≈ Feed-forward neural Networks

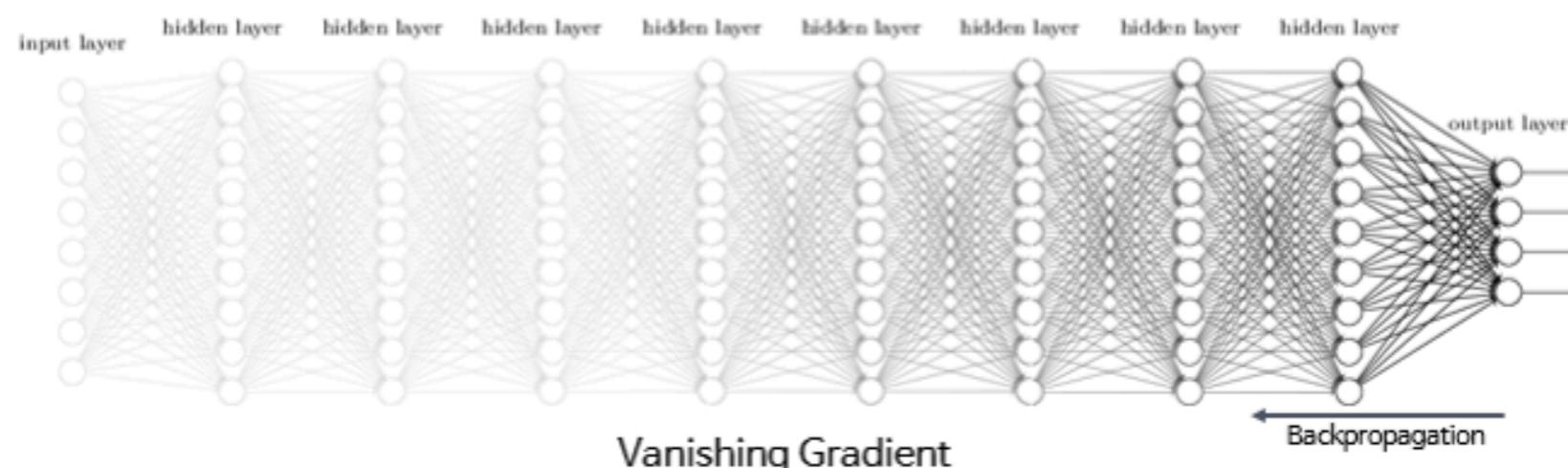
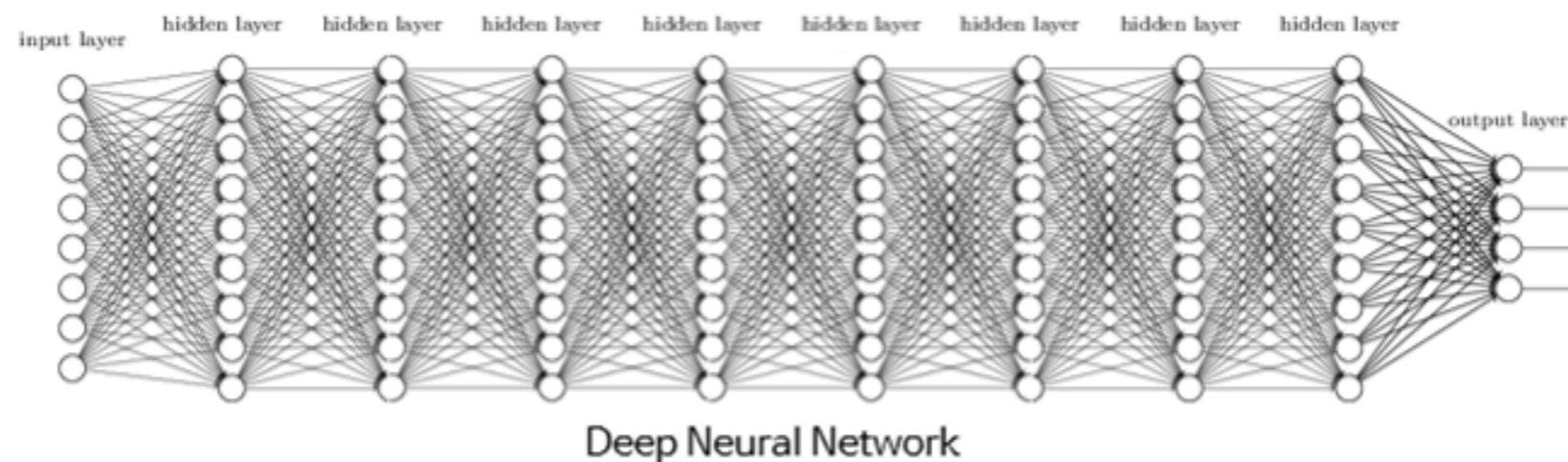
수많은 종류의 ANN 구조 등장



딥러닝 (Deep Learning)

경사 소실 문제 (Vanishing Gradient Problem)

- 여러 층으로 구성된 인공신경망의 학습과정에서 층에서 층으로 학습 파라미터가 전달됨에 따라 경사가 점점 작아져 결국 가중치(w)가 업데이트 되지 않아 학습이 되지 않는 현상
- ↪ **그래디언트 폭주 문제 (Exploding Gradient Problem)** :
경사가 점점 커져 가중치(w)가 기하급수적으로 커지게 됨에 따라 학습이 되지 않는 현상



What is the difference?

Artificial Neural Network (ANN)

≈ **Multi-Layer Perceptron (MLP)**

≈ **Deep Neural Network (DNN)**

≈ **Deep Feed-forward Networks**

≈ **Feed-forward neural Networks**

What is the difference?

Artificial Neural Network (ANN)

≈ Multi-Layer Perceptron (MLP)

≈ Deep Neural Network (DNN)

≈ Deep Feed-forward Networks

≈ Feed-forward neural Networks

기계학습 절차: 학습

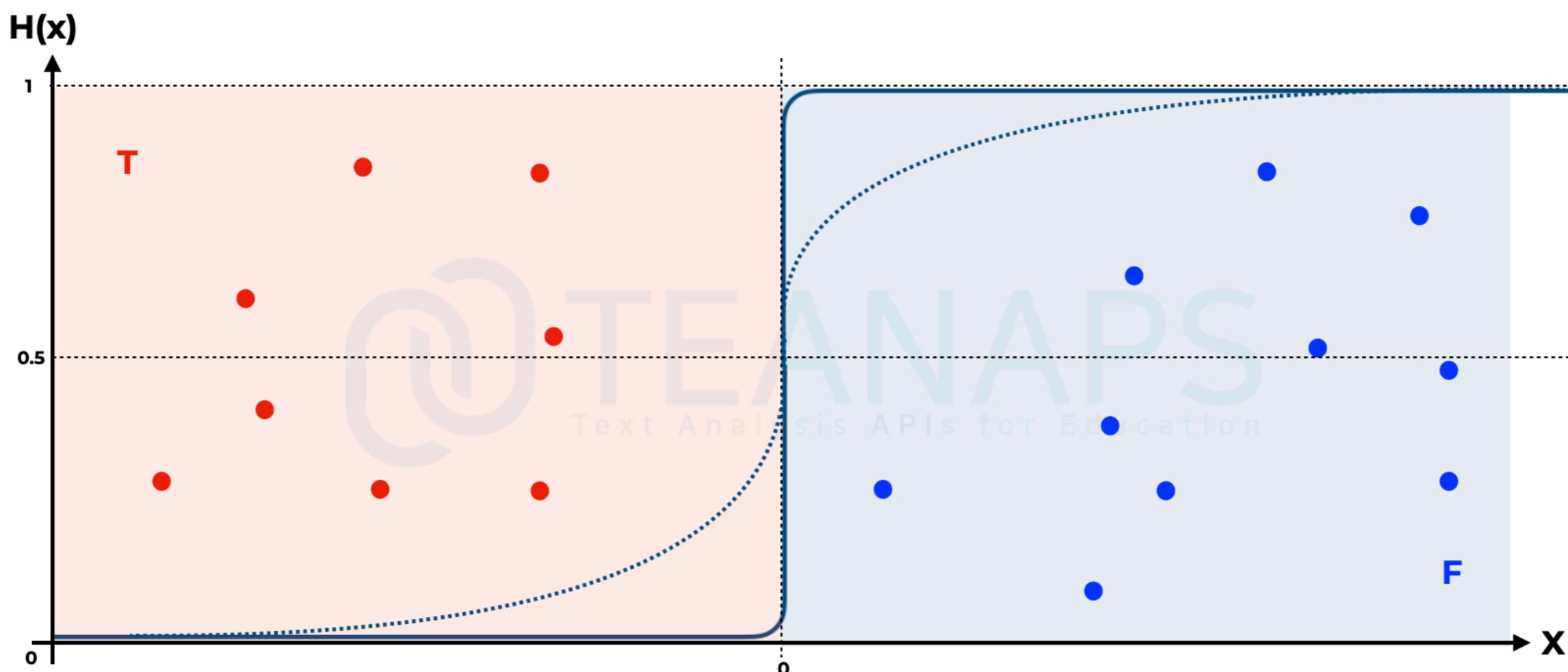
(Training)

Review

시그모이드 함수를 통한 로짓변환

$$\text{Sigmoid}(g) = \frac{1}{(1 + e^{-g})}$$

$$H(x) = Wx + b \rightarrow \text{Sigmoid}(H(x)) = \frac{1}{(1 + e^{-H(x)})}$$

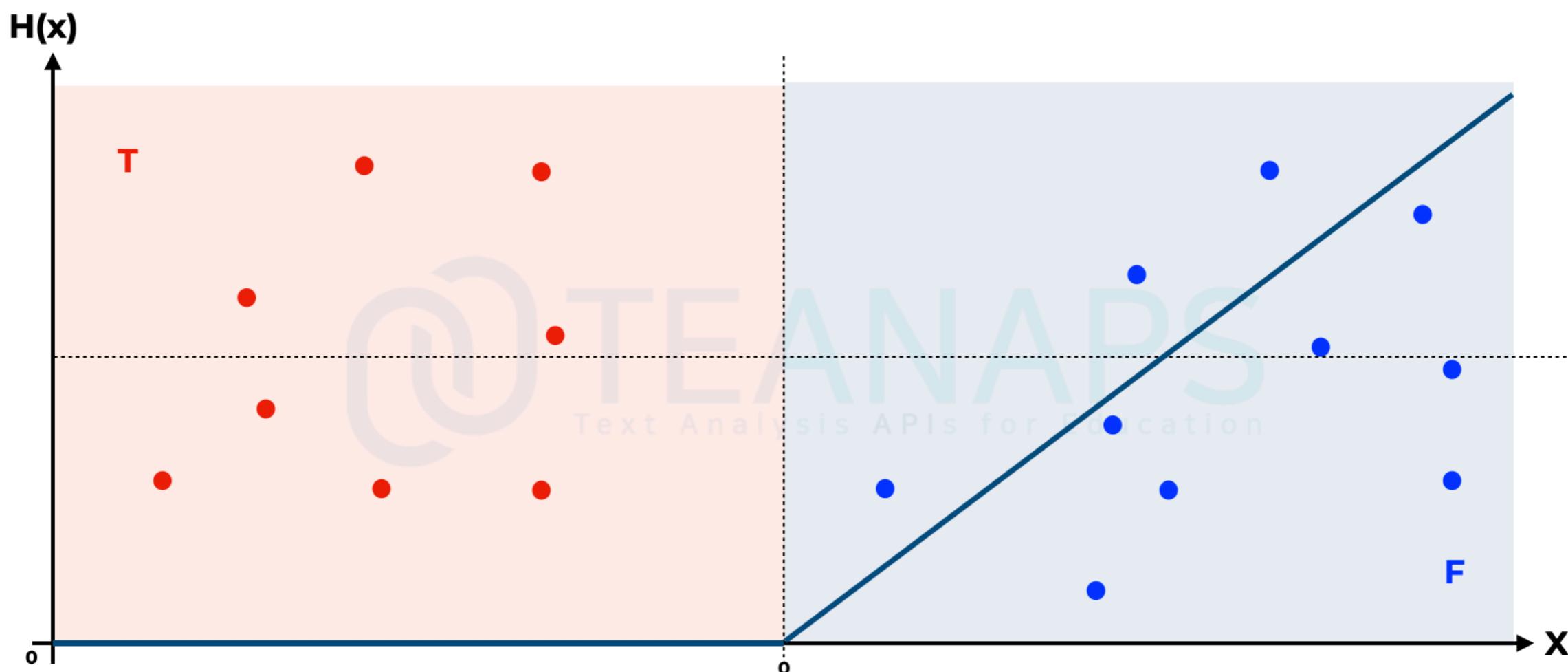


딥러닝

(Deep Learning)

ReLU (Rectified Linear Unit) 함수를 통한 로짓변환

$$\text{Sigmoid}(g) = \frac{1}{(1 + e^{-g})}$$
 \Rightarrow
$$\text{ReLU}(x) = \begin{cases} 0 & (x < 0) \\ x & (x \geq 0) \end{cases}$$



딥러닝 (Deep Learning)

활성화 함수 (Activation Function)

Name	Plot	Equation	Derivative
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		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2 + 1}$
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parameteric Rectified Linear Unit (PReLU) [2]		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Exponential Linear Unit (ELU) [3]		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1 + e^x)$	$f'(x) = \frac{1}{1 + e^{-x}}$

기계학습 절차: 데이터 준비

Review

자질추출 (Feature Extraction)

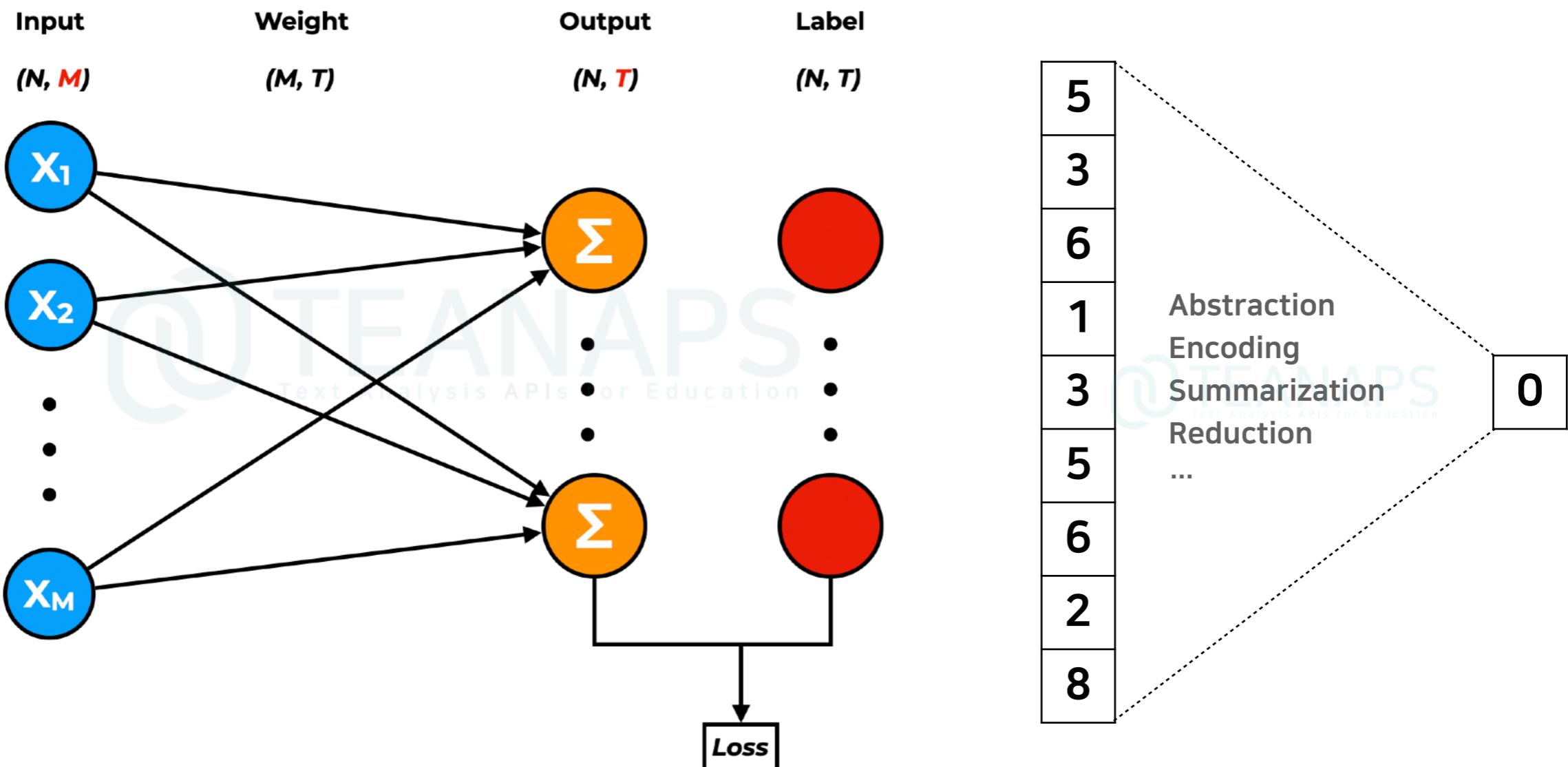
- 기계학습에 필요한 자질(변수, feature)을 추출하고 이를 수치로 표현하는 방법
- 과거 기계학습의 성능은 동일한 알고리즘에 대해 특징을 추출하는 방법에 의해 좌우되었으나, 딥러닝(deep learning)의 등장으로 자질을 추출하는 전처리 과정이 자동화되어 거의 사라짐

구분	메시지	특징 (Feature)					
		메시지 길이	URL 여부	특수문자 개수	해외수신 여부	의심단어 개수	광고문자 표시여부
1	[국제발신] 하루 30만 달 1천만 만원으로 이렇게! http://bit.ly/3f~	40	1	6	1	1	0
2	팀장님 안윤현 대리입니다. 출근하시면 결재 부탁드립니다.	20	0	1	0	0	0
3	(광고)웰컴박하라 vc⑤47③.co 코드 wc1004 무료수신거부 01084510000	45	1	5	0	1	1
4	(광고)신한과 함께하는 소중한 미래 따뜻한 금융 [신한]입니다. 2019년에 힘들었던 모든 일들은 다 잊어버리시고, ~	80	0	4	0	0	1
5	[WEB발신] 갤럭시 노트20/노트20 울트라 사전예약 오늘이 마지막날입니다-!! 구매를 망설이고 ~	75	0	6	0	0	0
6	(광고)등촌역스톤힐 ★더블역세권 9호선 ~ ★선착순으로 동호수 지정분양 가능 ★인근주변 아파트 시세보다 4~5억 저렴 ★~	120	0	8	0	0	1
7	[한진택배] 상품 배송 안내 안녕하세요 고객님. ★상품 수령이 편하신 장소를 선택 ~ ①직접수령 ②경비실 ③문앞 ~	110	0	9	0	0	0

딥러닝 (Deep Learning)

데이터 변환과 자질추출 (Data Transformation & Feature Extraction)

- 여러개의 입력변수를 적은 수의 변수로 요약하여 표현하는 방법
- 요약하는 방법에 따라 원래 입력변수의 특징과 출력에 영향을 미치는 변수의 특징을 표현할 수 있음



딥러닝 (Deep Learning)

데이터 변환과 자질추출 (Data Transformation)

Data

5		
3		
6		
1		
3		
5		
6		
2		
8		

Center

5	3	6
1	3	5
6	2	8

Average

5	3	6
1	3	5
6	2	8

4.3

Median

X	X	X
X	3	X
X	X	X

Weighted Sum/Average

5	3	6
1	3	5
6	2	8

3

0	0	0
1	0	0
0	1	0

합성곱 신경망

(Convolution Neural Network, CNN)

합성곱 (Convolution)

Filter

1	0	0
0	1	0
1	0	1

Data

1	0	0	1	1	0	0
0	0	1	1	0	1	0
1	0	0	1	1	1	1
0	0	1	1	1	1	0
0	1	1	0	0	0	0

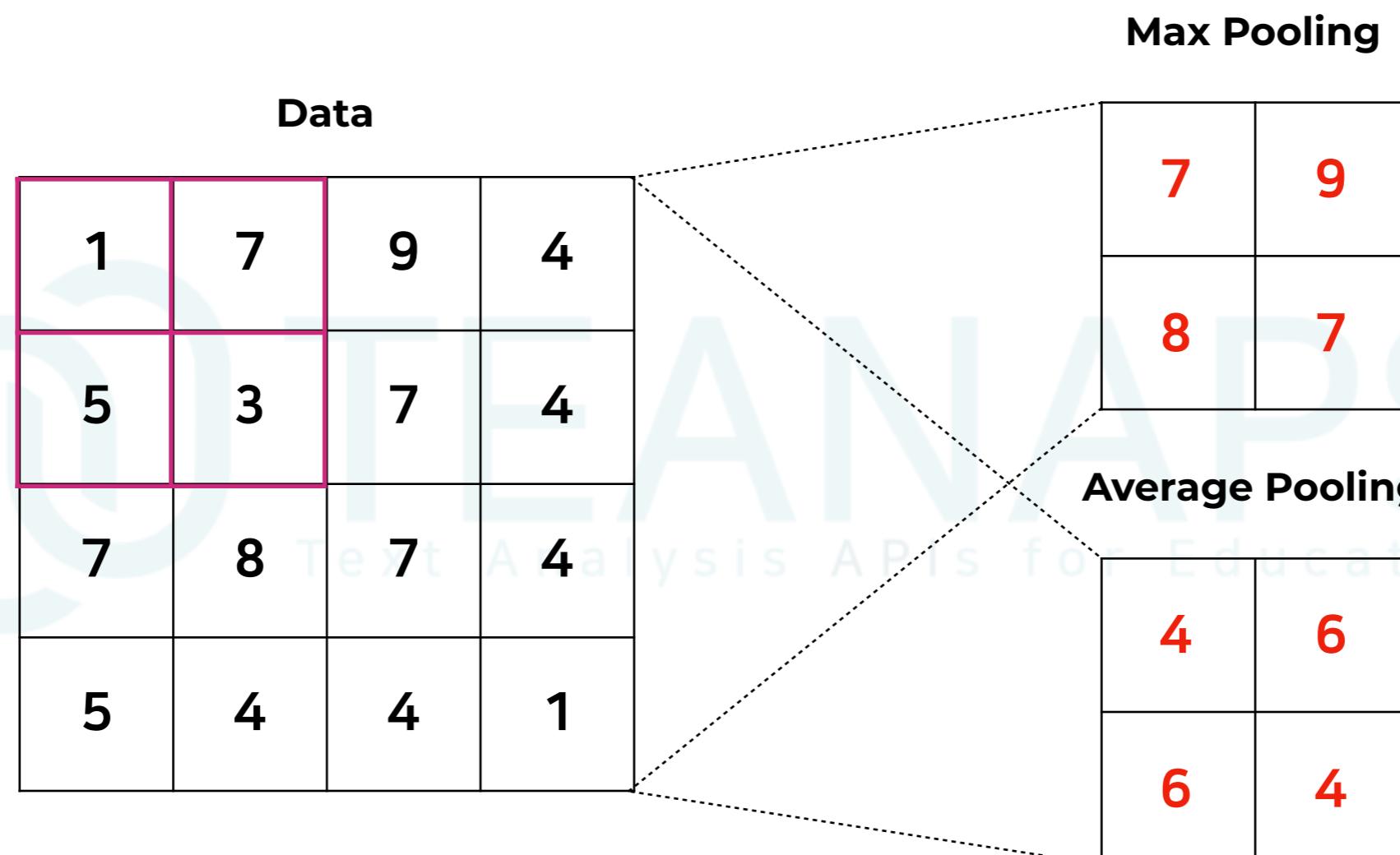
Convolved Data

1	2	3
4	5	6
7	8	9

합성곱 신경망

(Convolution Neural Network, CNN)

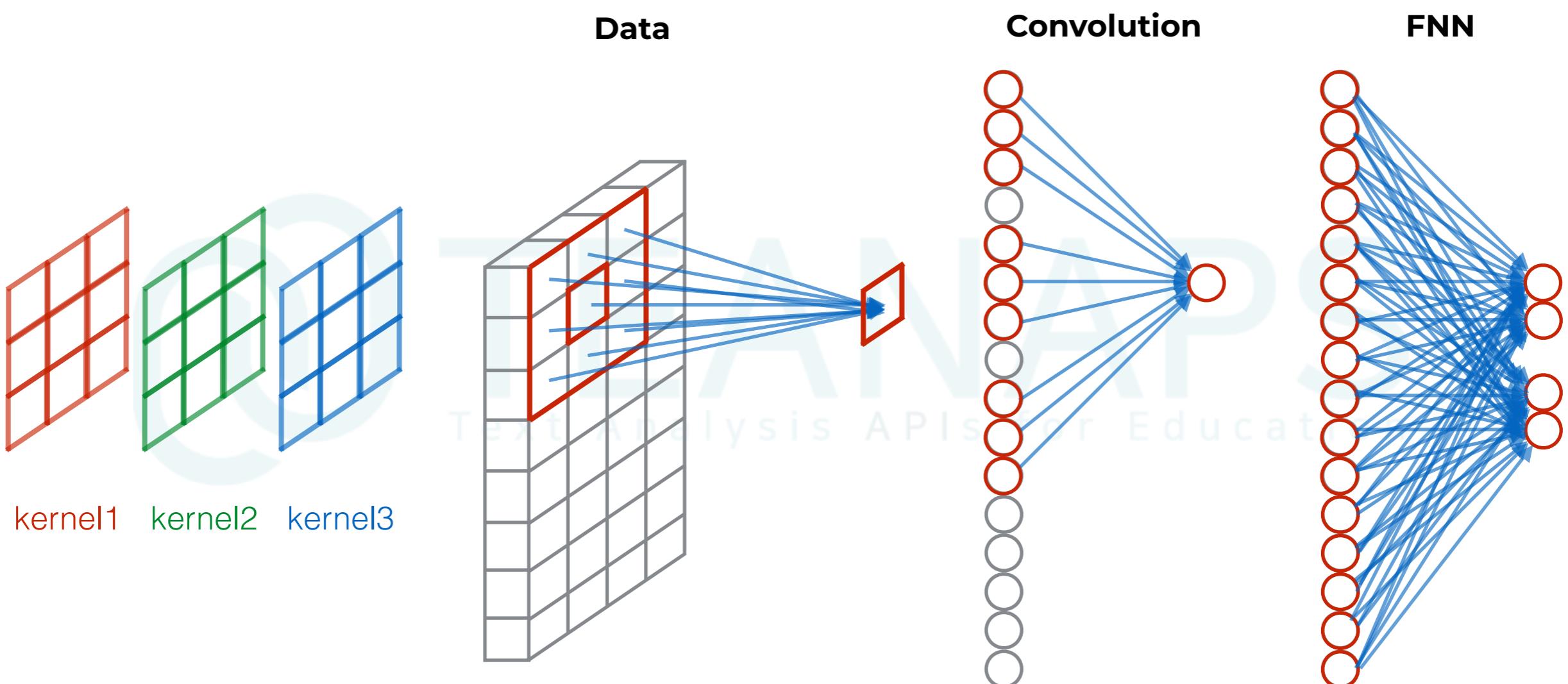
풀링 (Pooling)



합성곱 신경망

(Convolution Neural Network, CNN)

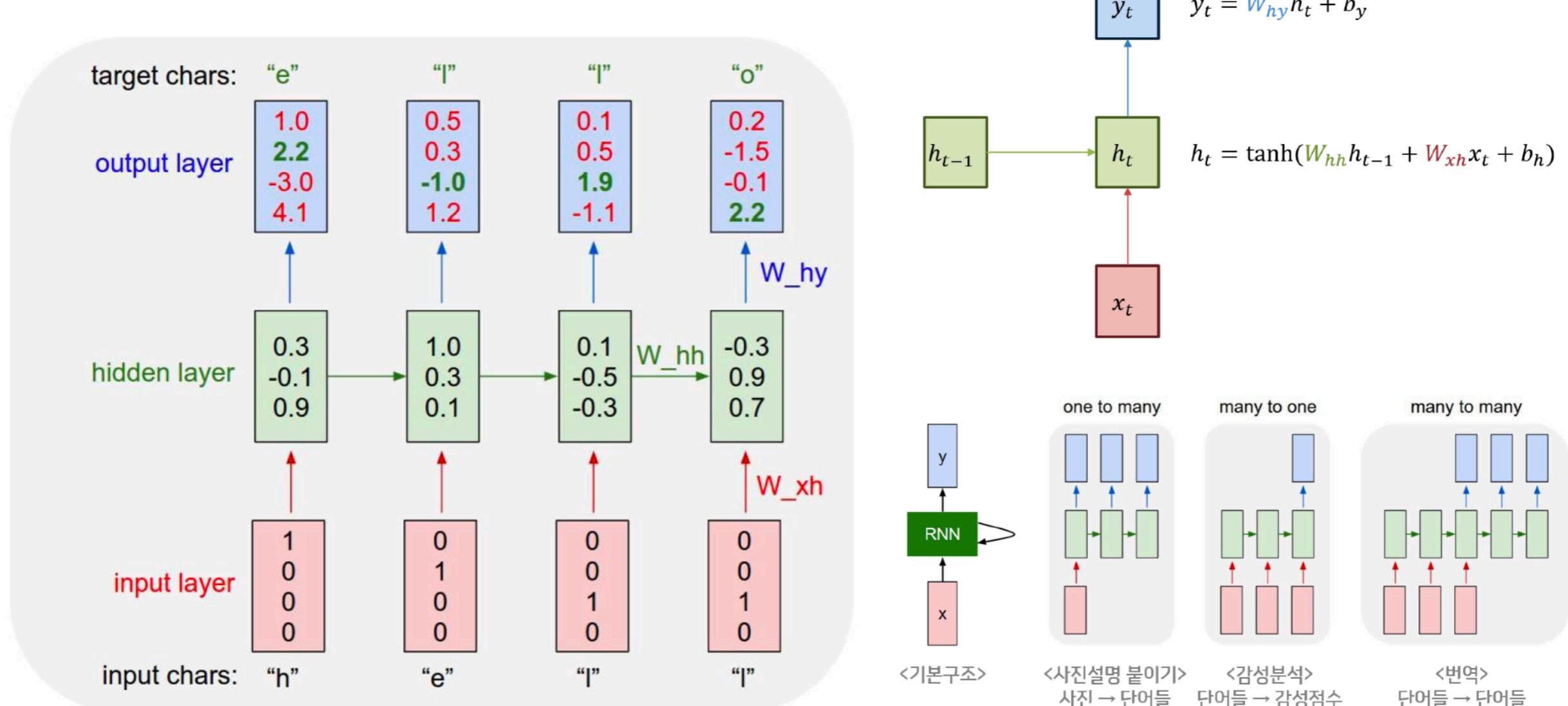
합성곱과 풀링 (Convolution & Pooling)



순환 신경망

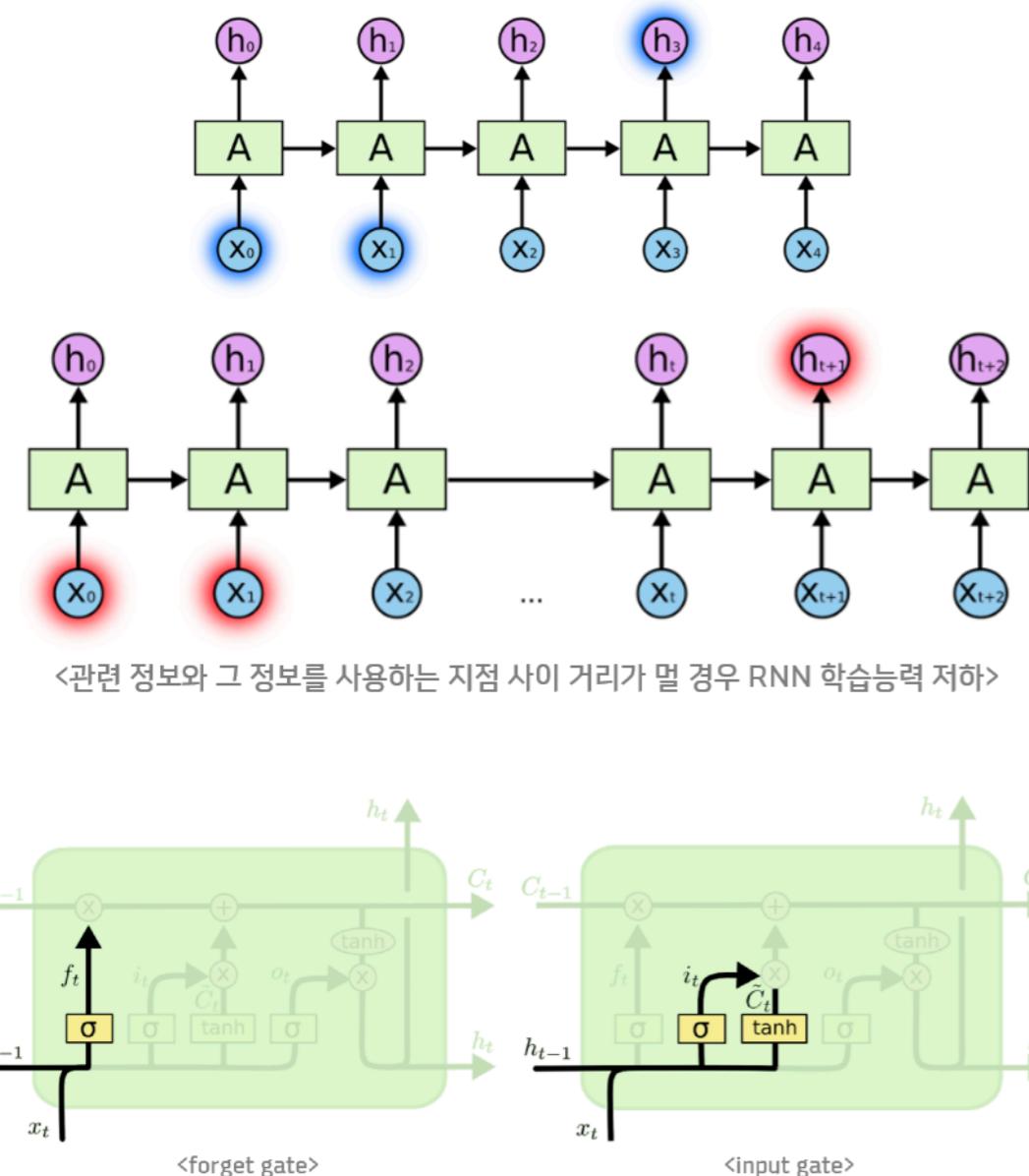
(Recurrent Neural Network, RNN)

인공신경망과 재귀신경망 (ANN & RNN)

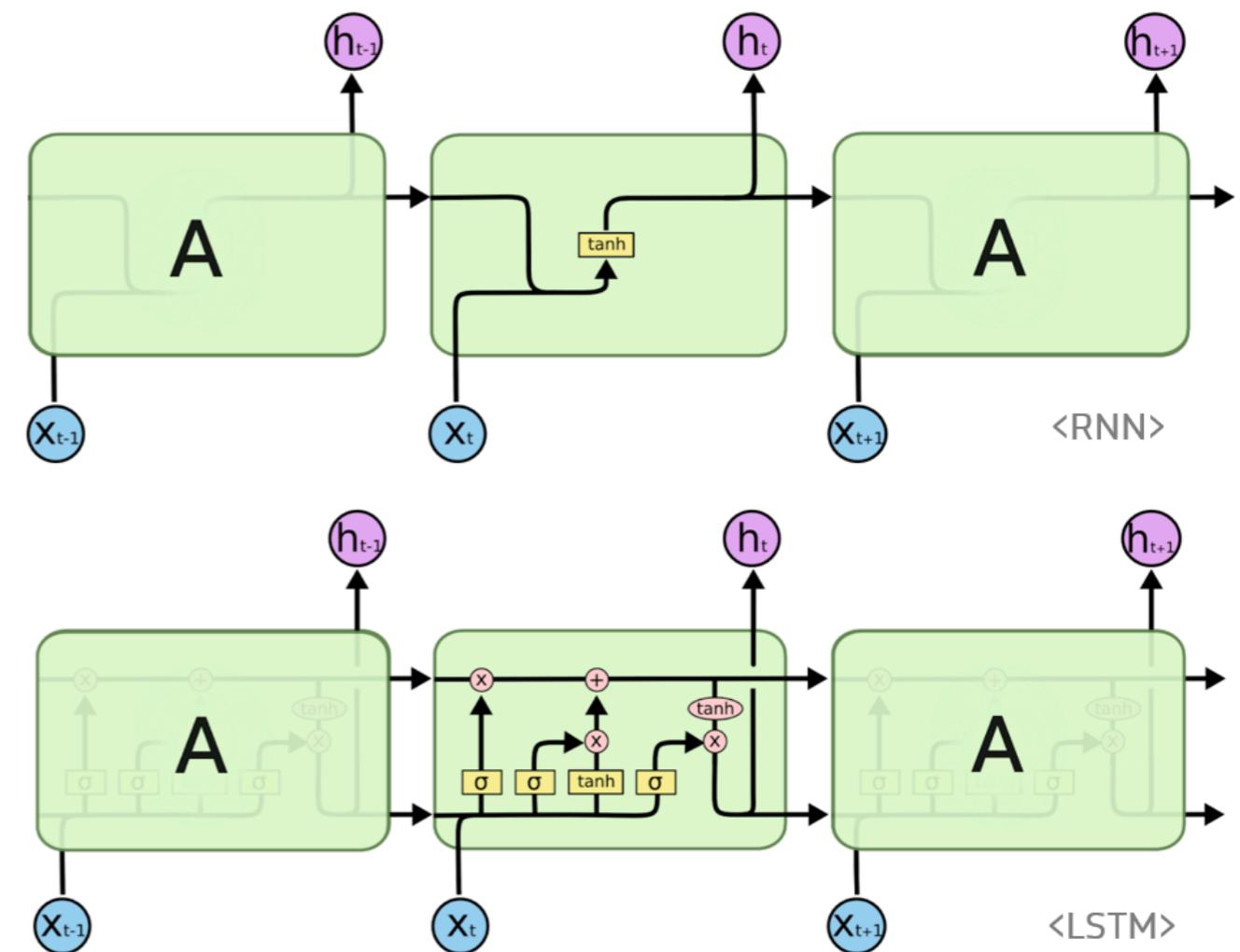


순환 신경망: LSTM (Long Short-Term Memory)

LSTM (Long Short-Term Memory)



<관련 정보와 그 정보를 사용하는 지점 사이 거리가 멀 경우 RNN 학습능력 저하>



순차 레이블링: RNN

RNN을 활용한 순차 레이블링

Preview

구분

내용 (B: Beginner, I: Inner, O: Outer)

원문 손 흥 민 이 골 을 작 렬 하 며 토 트 넘 홓 스 퍼 의 승 리 를 이 꼴 었 다 .



E.O.D

Contact

-  <http://www.teanaps.com>
-  fingeredman@gmail.com