

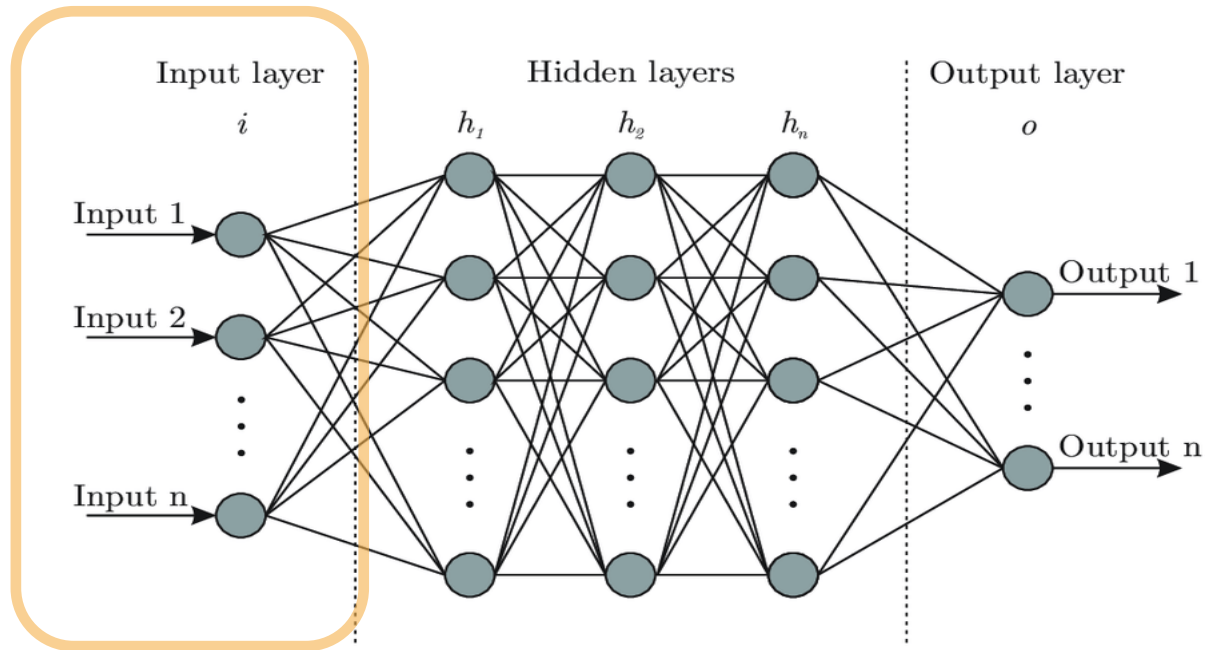
딥러닝팀

1팀

이수경
이승우
이은서
주혜인
홍현경

- 용어 정리

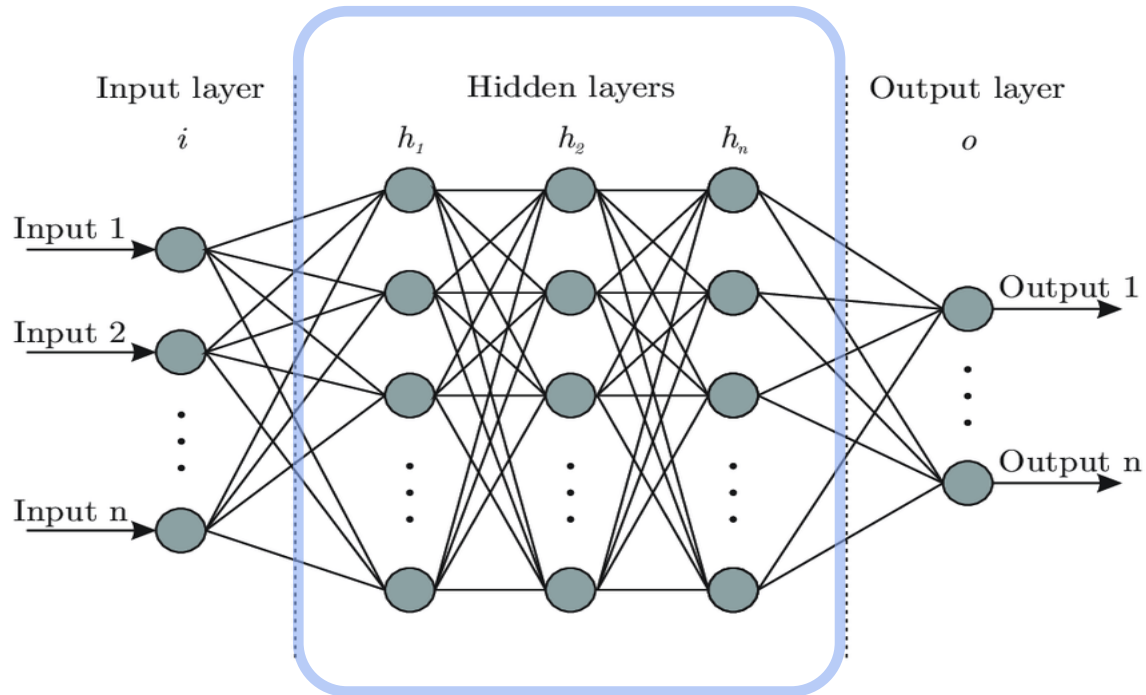
(2) 입력층



데이터의 feature가 입력값으로 들어옴

- 용어 정리

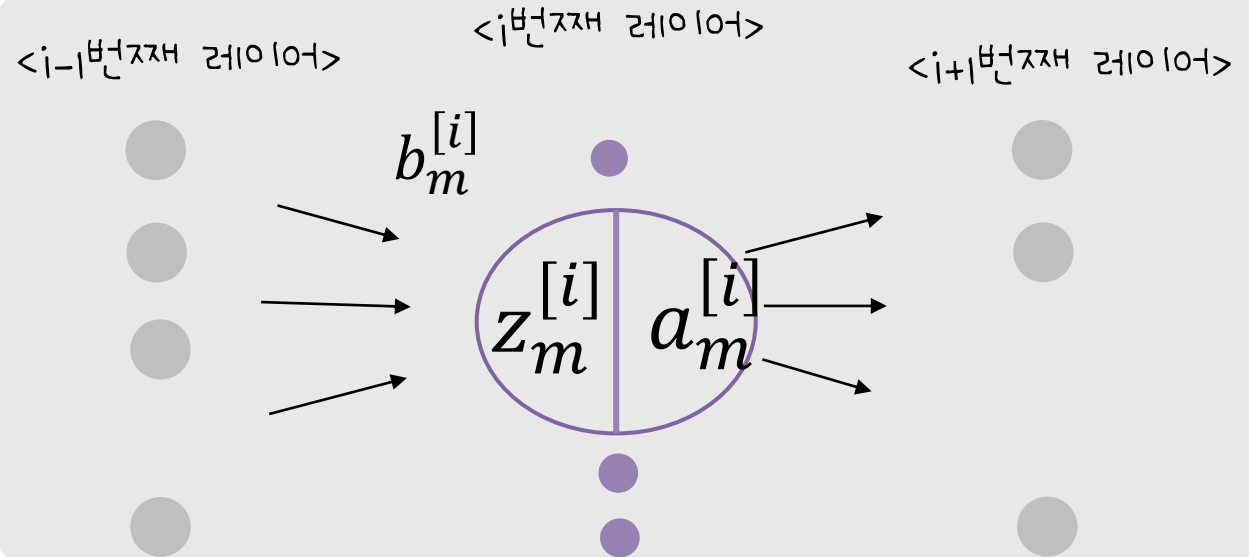
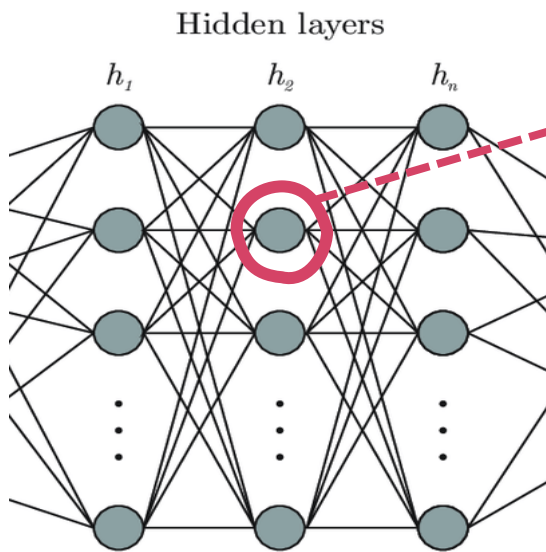
(3) 은닉층



입력층과 출력층 사이의 모든 layer

● 용어 정리

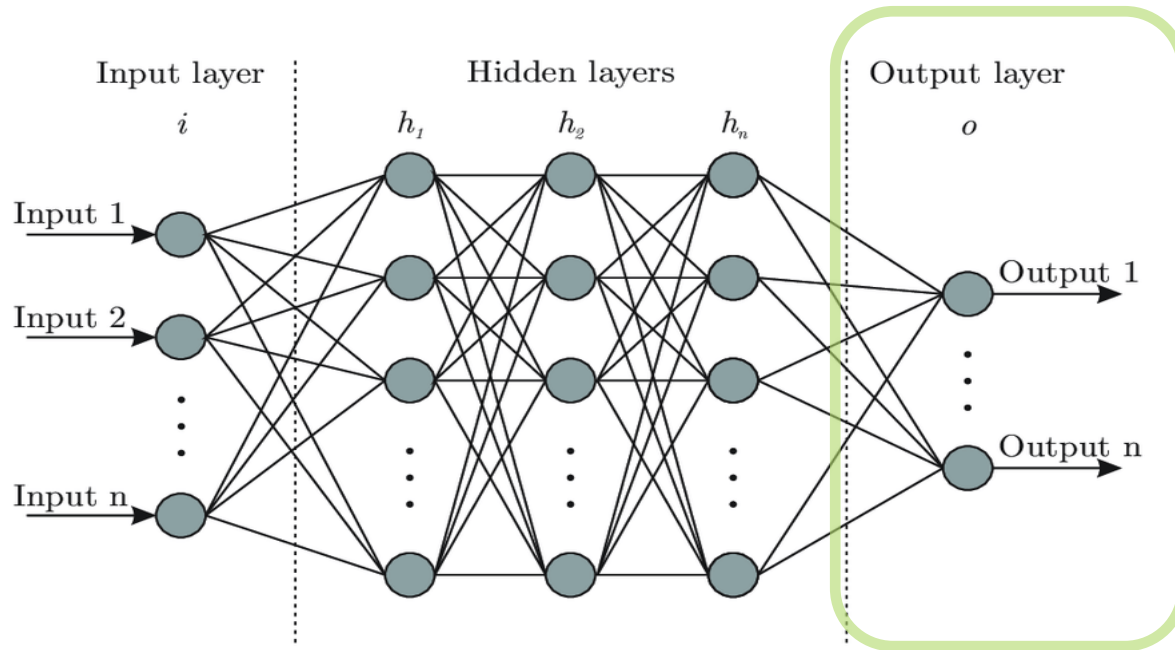
(3) 은닉층



1. 선형연산 $z^{[i]} = \sum_{m=1}^M X^{[i-1]} * W^{[i]} + b^{[i]}$
2. 활성화함수 통과 $\sigma(z^{[i]}) = a^{[i]} = y^{[i]}$
3. 전달 $y^{[i]} = x^{[i+1]}$

- 용어 정리

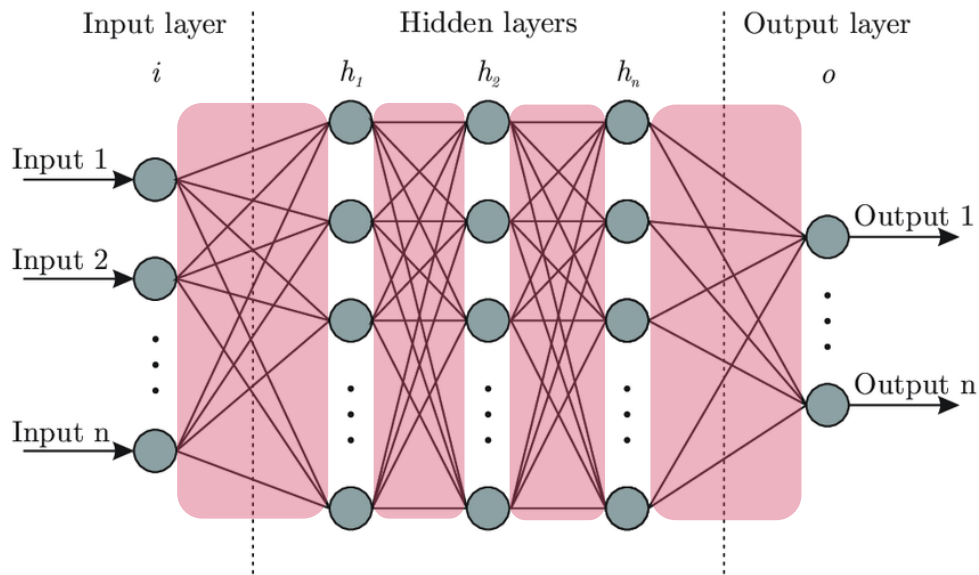
(4) 출력층



신경망 외부로 출력 신호 전달
출력층의 **활성화 함수**가 전체 신경망의 역할 결정

● 용어 정리

(5) 가중치



$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{bmatrix}$$

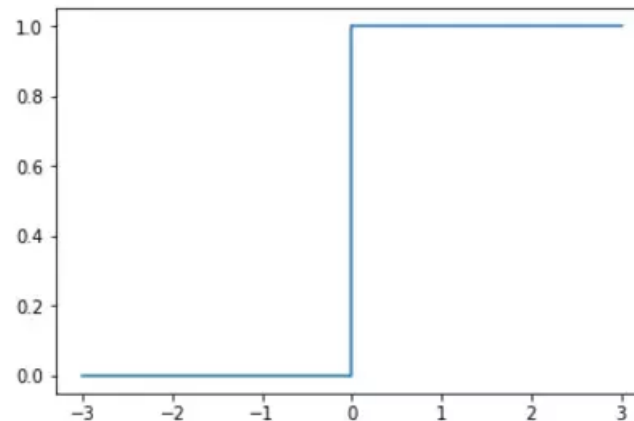
이전 층과 이후 층의 결합 강도 결정
즉, 입력 값의 영향력 결정함

- 활성화 함수

활성화 함수

입력 신호의 총합이 어떻게 활성화를 일으키는지를 결정

〈Step function〉



은닉층 – 선형결합을 거친 값을 통과시켜 다음 층 노드의 입력값이 됨

출력층 – 입력 신호의 총합을 어떤 모양으로 출력할지 결정

ex) sigmoid, tanh, ReLU, Softmax, Identity

- 손실함수

손실함수

y 와 \hat{y} 의 차이를 계산하는 함수
 y 와 \hat{y} 의 차이를 최소화하는 방향으로 학습 진행

모델의 목적 (문제)

분류 문제

교차 엔트로피 오차
Cross Entropy Error (CEE)

회귀 문제

오차 제곱합
Sum of Squares for Error(SSE)

3 딥러닝의 학습

- 역전파 (Back Propagation)

역전파란 무엇일까?

< i번째 sample >

“**학습**”

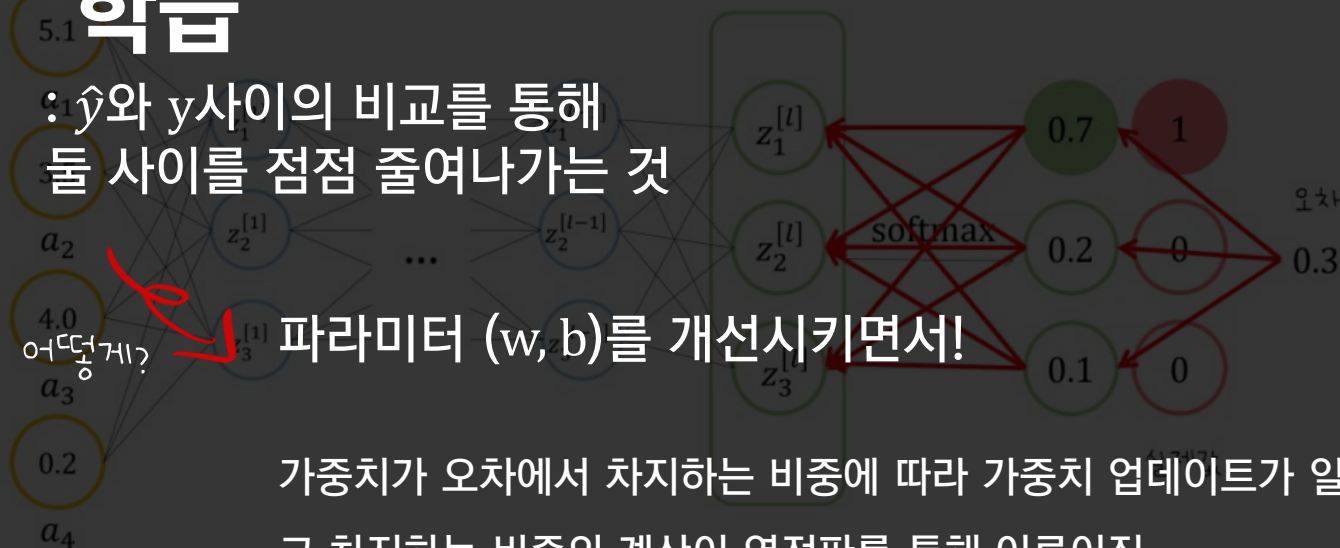
: \hat{y} 와 y 사이의 비교를 통해
둘 사이를 점점 줄여나가는 것

어떻게?

파라미터 (w, b)를 개선시키면서!

가중치가 오차에서 차지하는 비중에 따라 가중치 업데이트가 일어나는 데,
그 차지하는 비중의 계산이 역전파를 통해 이루어짐

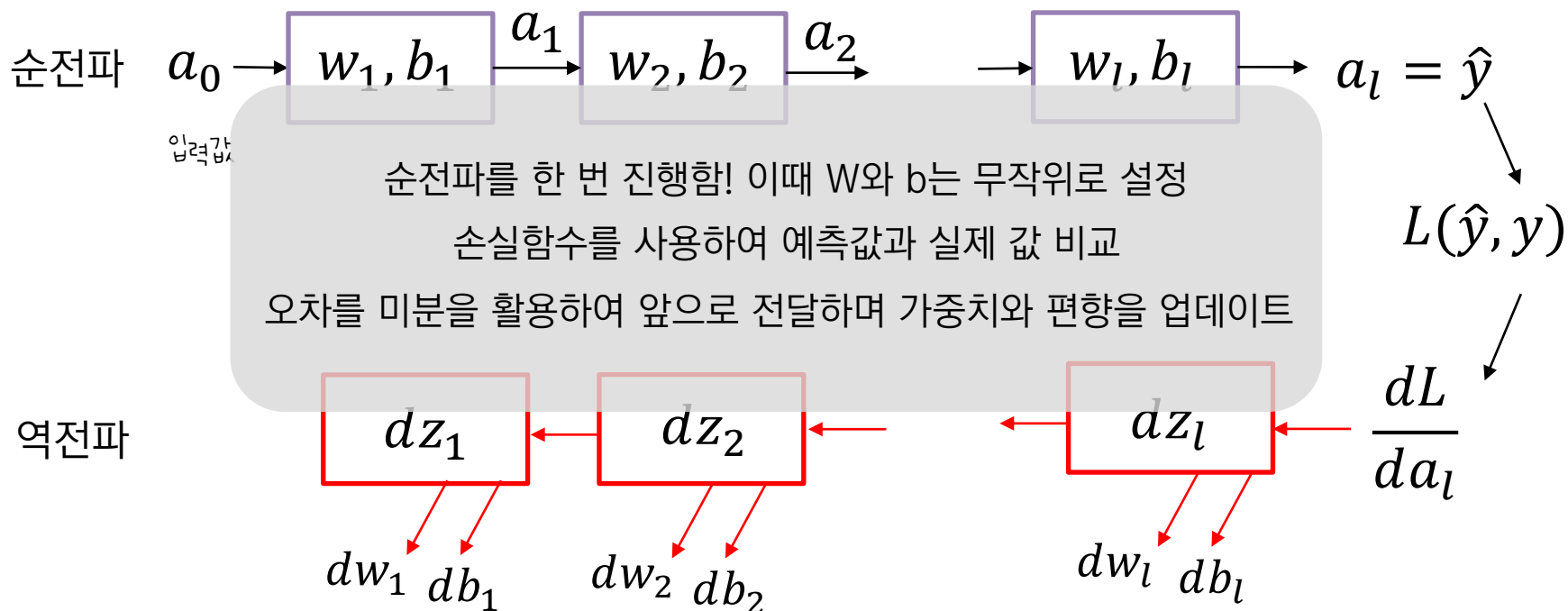
순전파 과정에서 얻은 오차를 줄이기 위해 앞 층으로 오차의 편미분 값을 전달해주며
모델이 파라미터를 학습할 수 있도록 하는 과정



3 딥러닝의 학습

● 역전파 (Back Propagation)

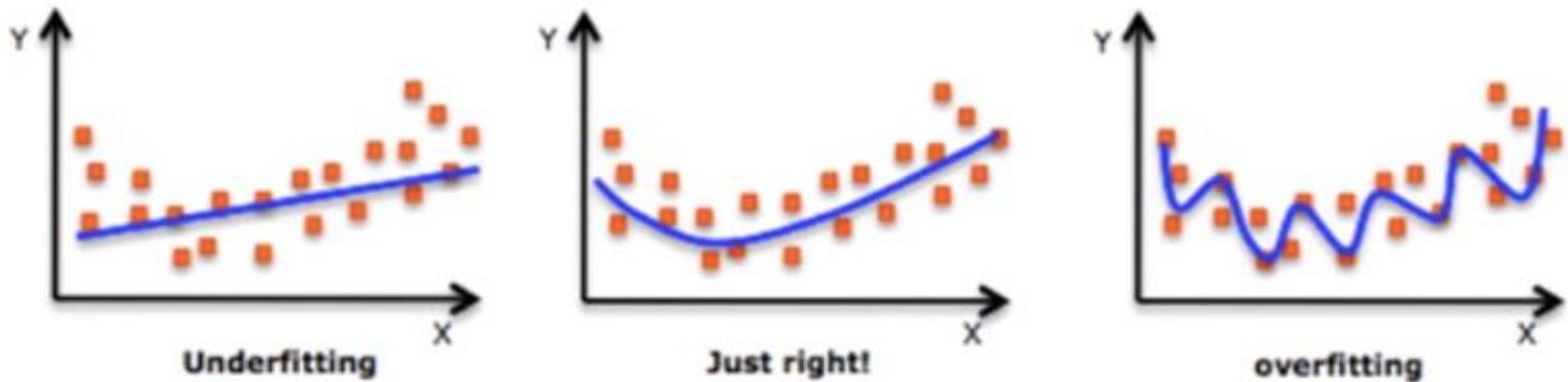
역전파란 무엇일까?



4 모델 성능 향상

- Overfitting이란?

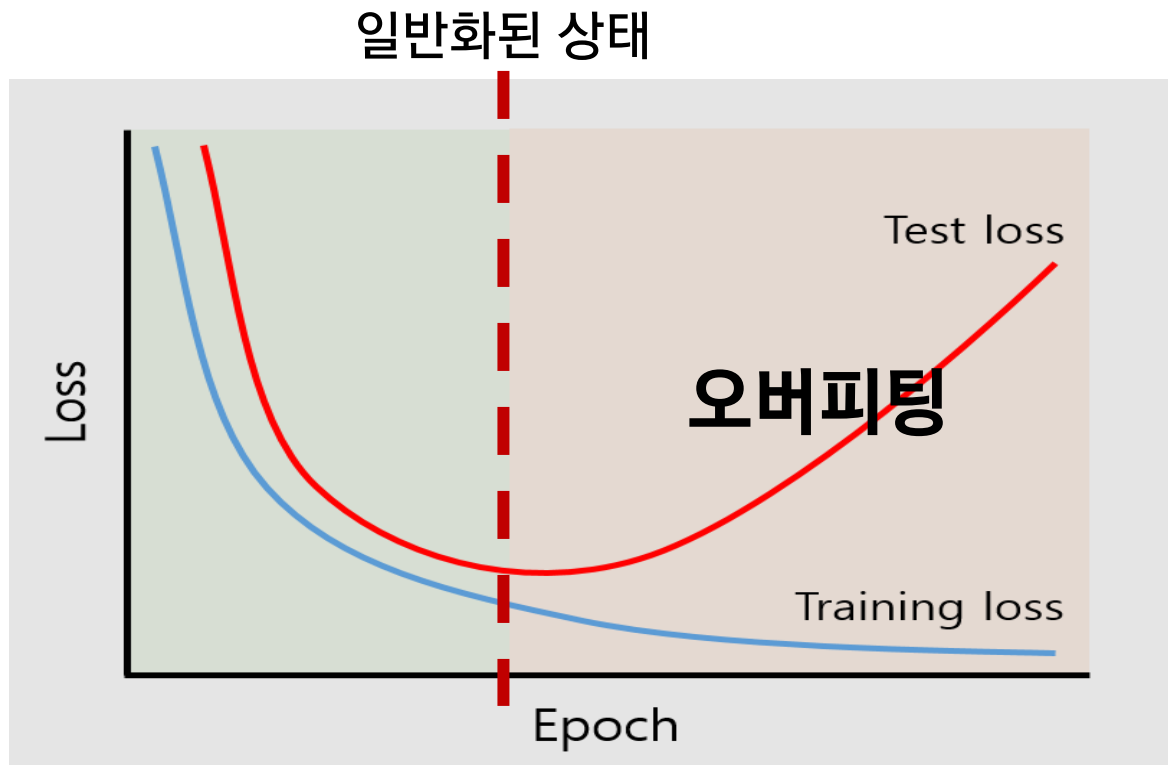
Overfitting



학습 데이터의 지엽적인 특성까지 학습하여
새로운 데이터 입력 시 제대로 반응하지 못하는 상태

4 모델 성능 향상

● Overfitting이란?

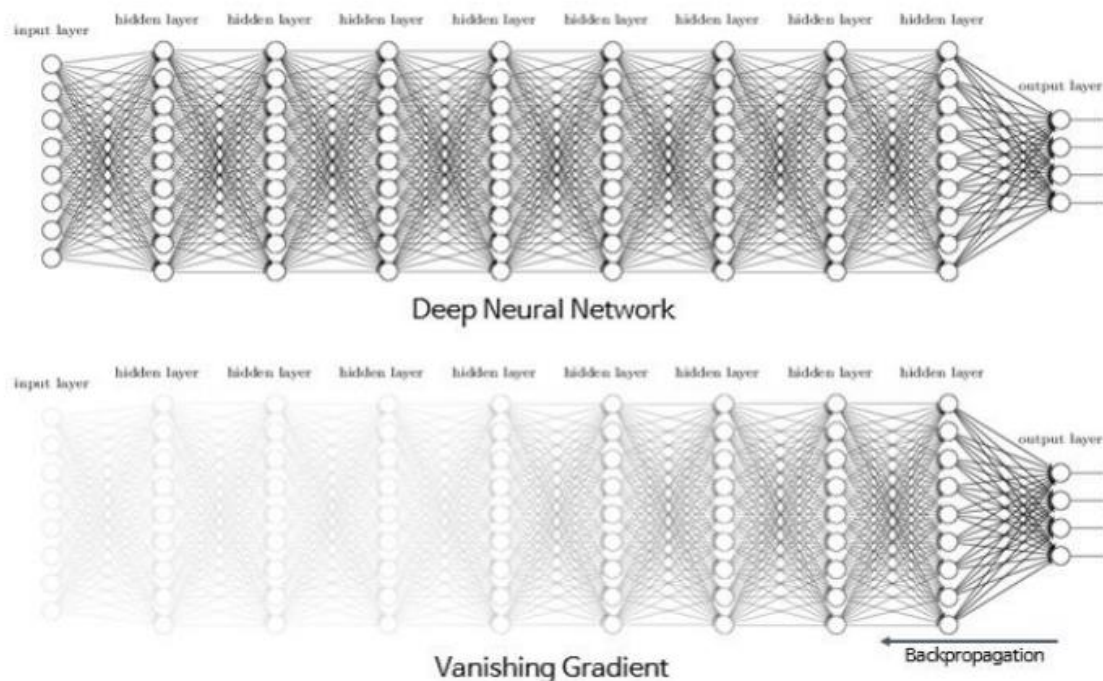


우리의 목표는 어떤 데이터를 주어도
균등한 성능을 내는 **일반화된** 모델을 만드는 것

4 모델 성능 향상

● 더 알아보기

기울기 소실 문제

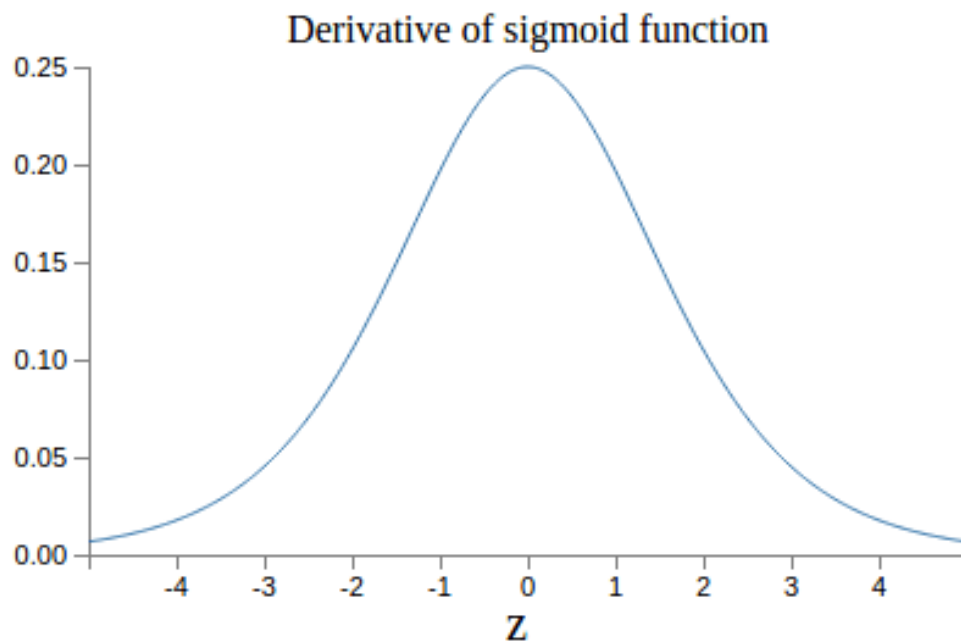


층이 깊은 심층신경망에서 역전파시 gradient가 입력층으로 전달됨에 따라 점점 작아져 가중치가 업데이트 되지 않는 것

4 모델 성능 향상

● 더 알아보기

기울기 소실 문제



기존에 사용하던 sigmoid 함수의 미분 값은 최대 0.30이 되지 않았고,
이 작은 값이 역전파를 하며 최소 수십 번 곱해지자 전해지는 값이 결국 사라지게 됨