

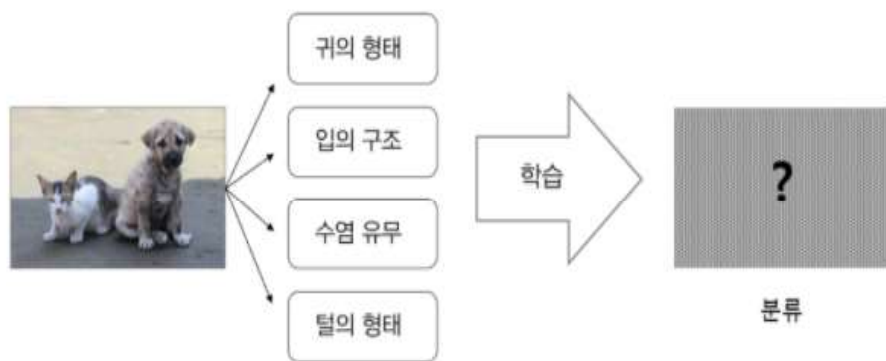
INDEX

1. 딥러닝 Basic
2. 신경망과 딥러닝
3. 딥러닝의 학습
4. 모델 성능 향상

1 딥러닝 배경지식

- 특징 추출 (Feature Extraction)

머신러닝



어떤 특징을 추출할 지 개발자가 알려줌
→ Feature extraction 필요

딥러닝

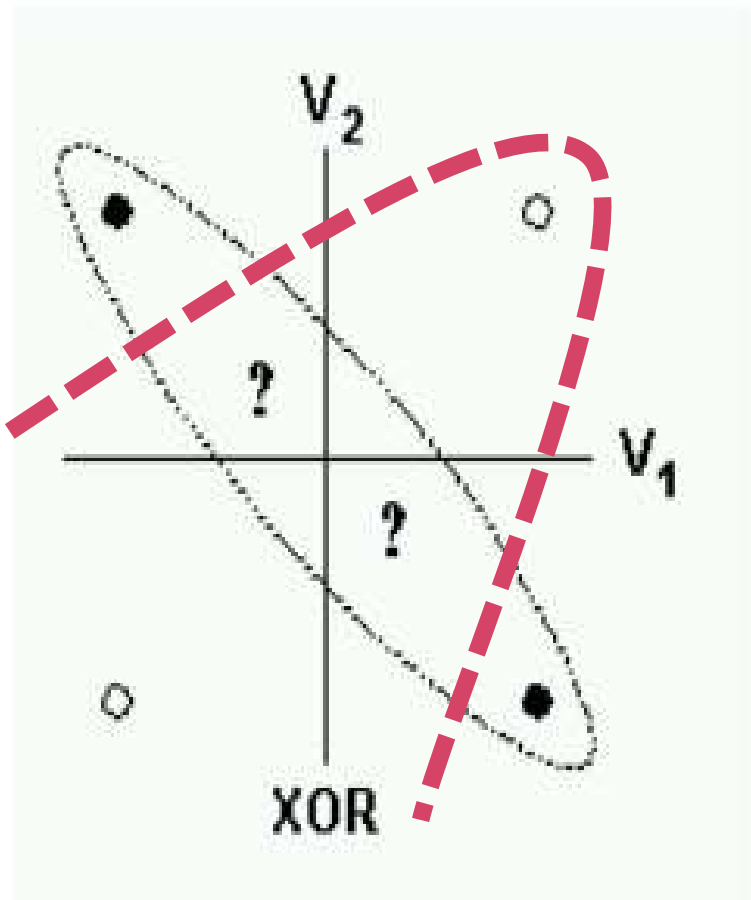


알려주지 않아도 알아서 학습함
→ Feature extraction 필요 X

1 딥러닝 배경지식

- 딥러닝 기본 단위, perceptron

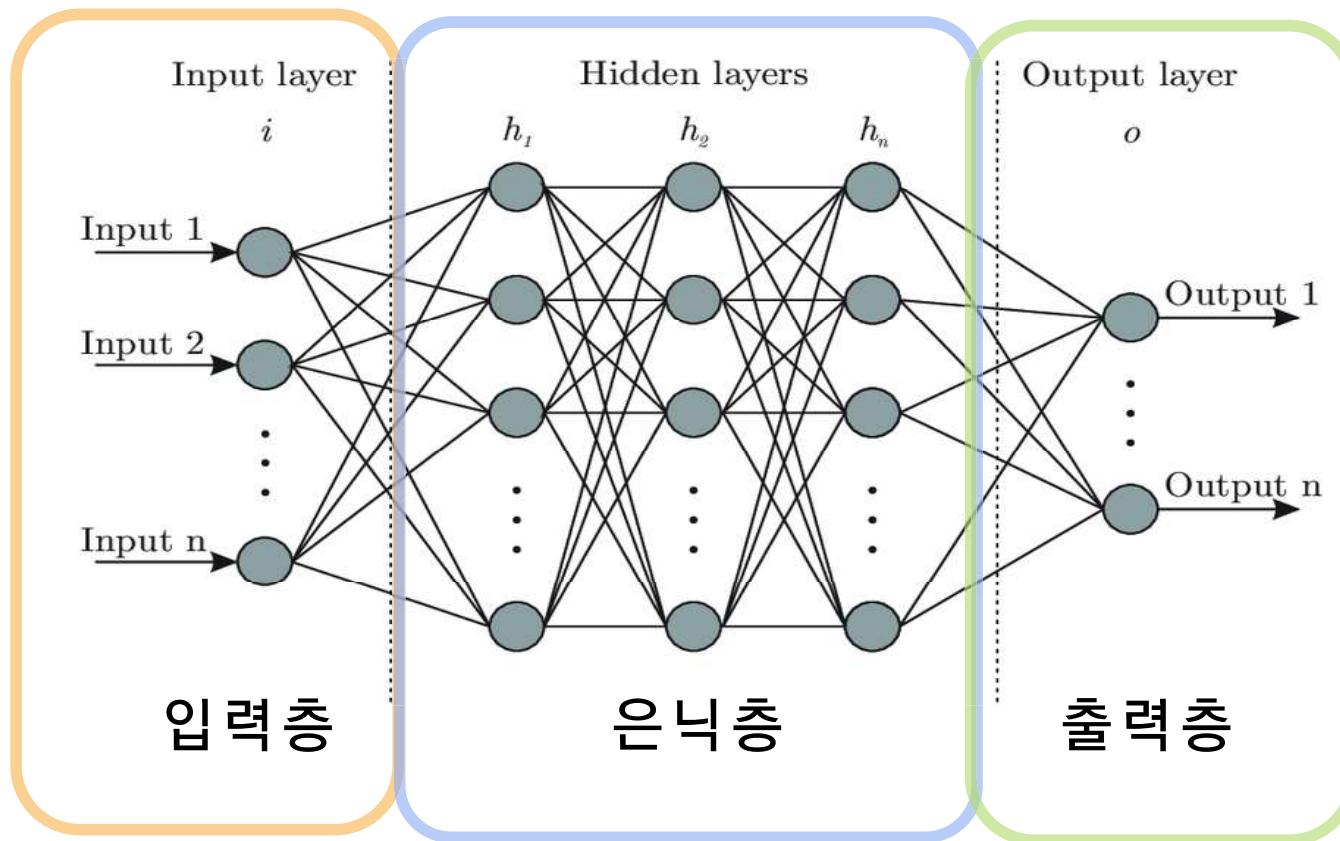
XOR 문제



비선형적인 모델을 활용해
해결할 수 있을 것 같다!

- 용어 정리

(1) 신경망

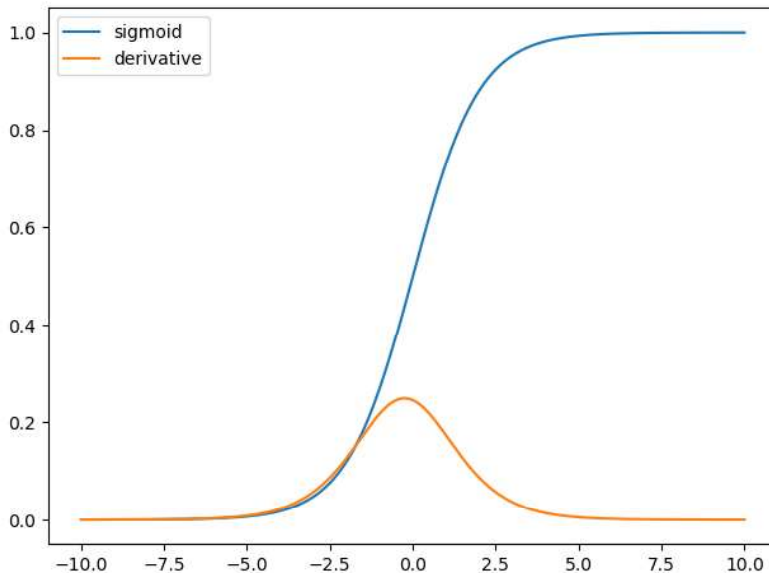


다층 퍼셉트론과 마찬가지로, 3가지 층으로 구성됨

● 활성화 함수

(1) Sigmoid Function

0부터 1사이의 모든 실수를 출력값으로 가짐



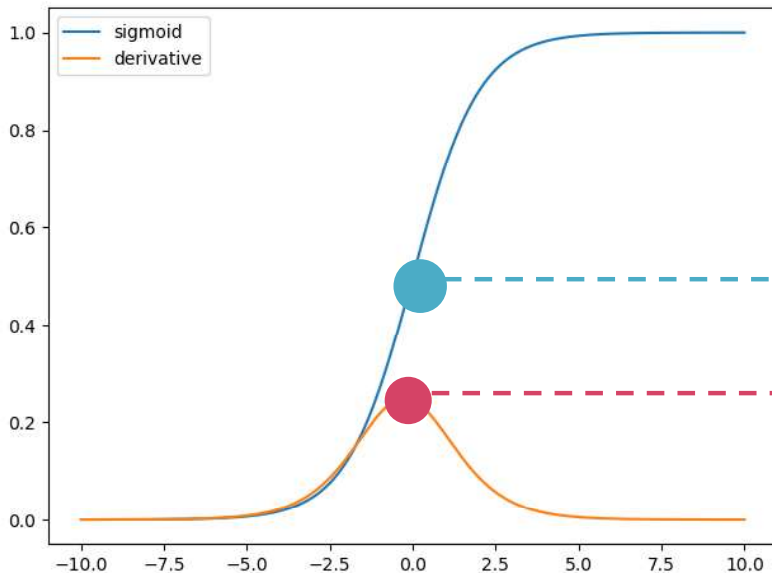
$$S(x) = \frac{1}{1 + e^{-x}} = \frac{e^x}{e^x + 1}$$

- 미분 가능 - - - - - 역전파 가능
- 비선형 함수 - - - - - 변수 간의 복잡한 관계 설명 가능

● 활성화 함수

(1) Sigmoid Function

0부터 1사이의 모든 실수를 출력값으로 가짐



But,

$$S(x) = \frac{1}{1 + e^{-x}} \rightarrow \text{학습이 느려짐}$$

미분 최대값 0.25
→ 기울기 소실 문제

- 미분 가능 ----- 역전파 가능
- 비선형 함수 ----- 변수 간의 복잡한 관계 설명 가능

2 신경망과 딥러닝

- 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

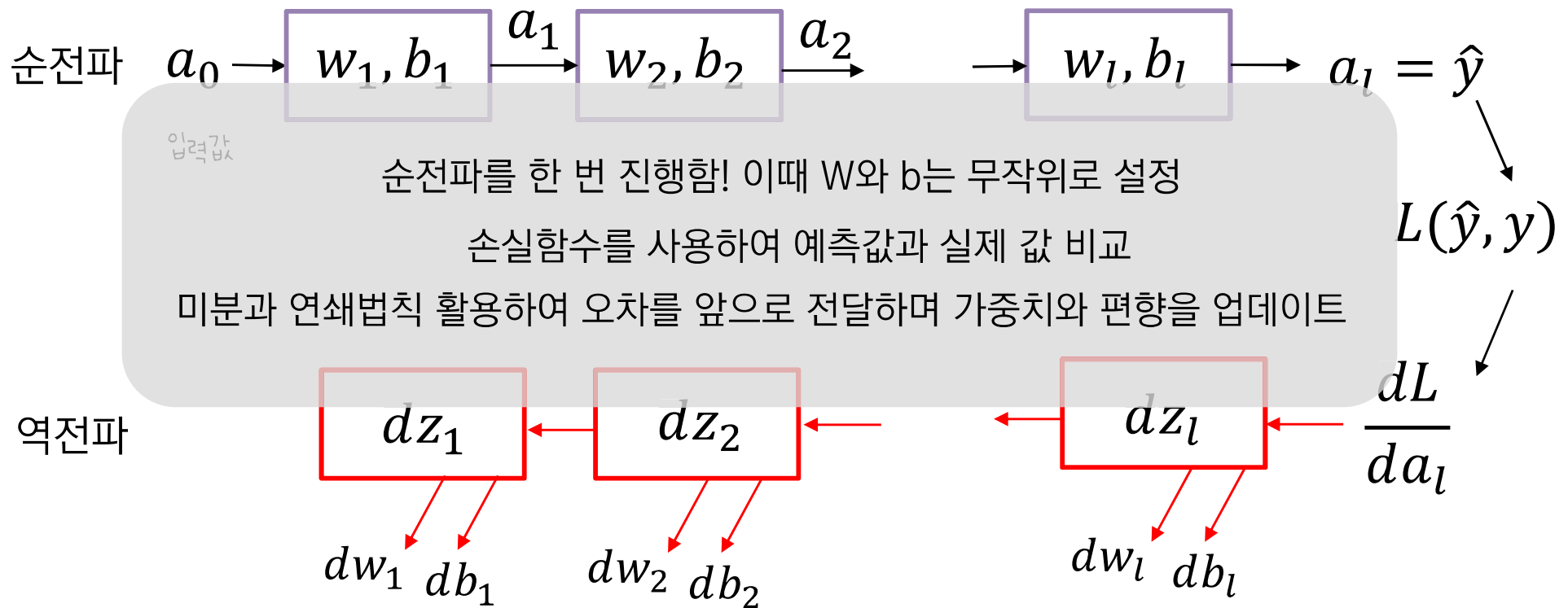
회귀 문제

오차 제곱합
Sum of Squares Error, SEE

3 딥러닝의 학습

- 역전파 (Back Propagation)

역전파란 무엇일까?

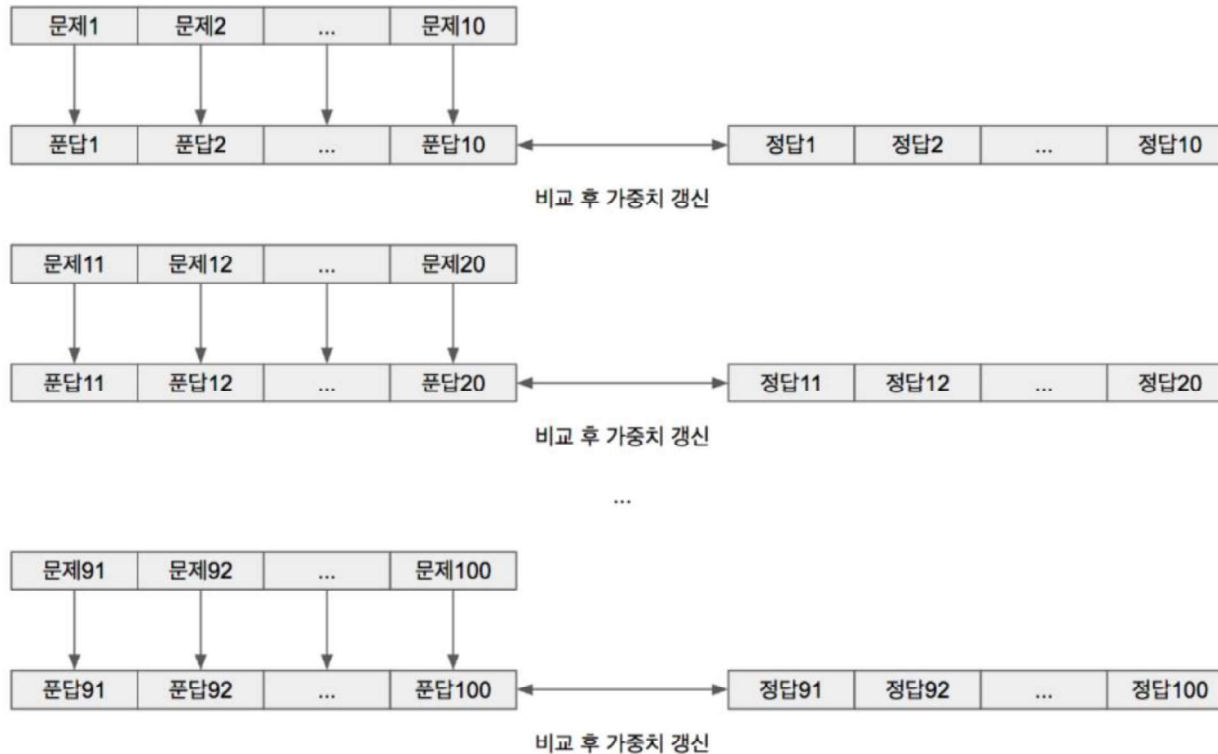


3 딥러닝의 학습

● 최적화 (Optimizer)

미니배치 경사 하강법

(Mini-batch gradient descent, MSGD)

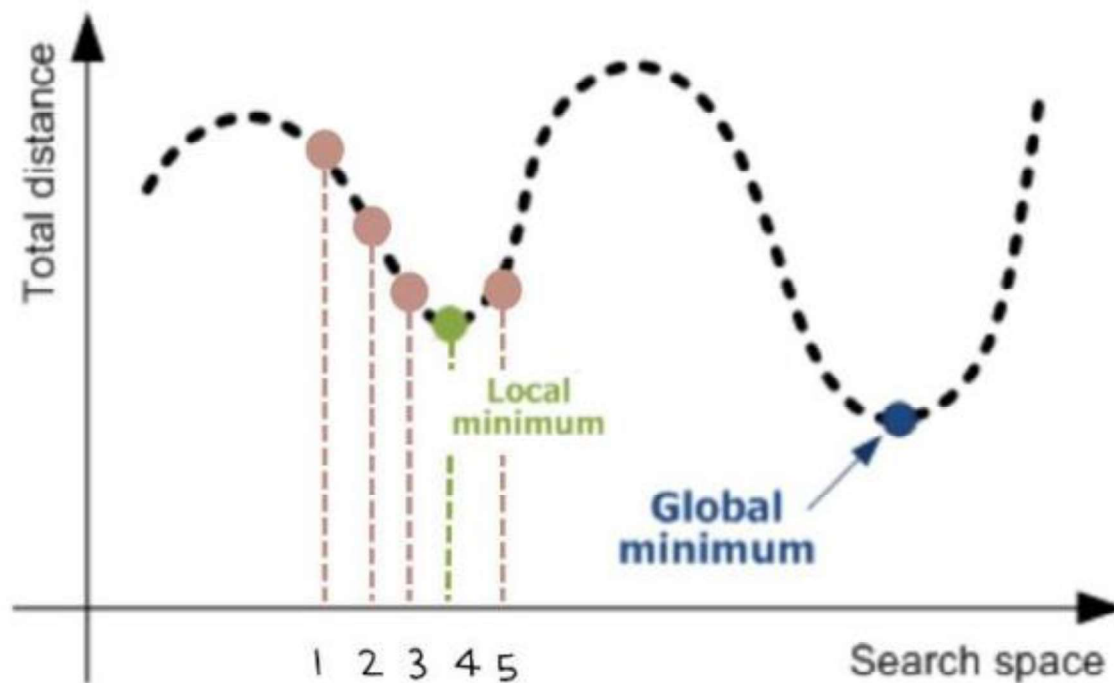


배치 사이즈 = m 인 경우,
1 epoch당 # 's of rows / m 만큼의
파라미터 업데이트가 발생함
옆의 경우 $100 / 10 = 10$ 번의
파라미터 업데이트 발생

3 딥러닝의 학습

- 문제점

국소 최적해 (local optima)



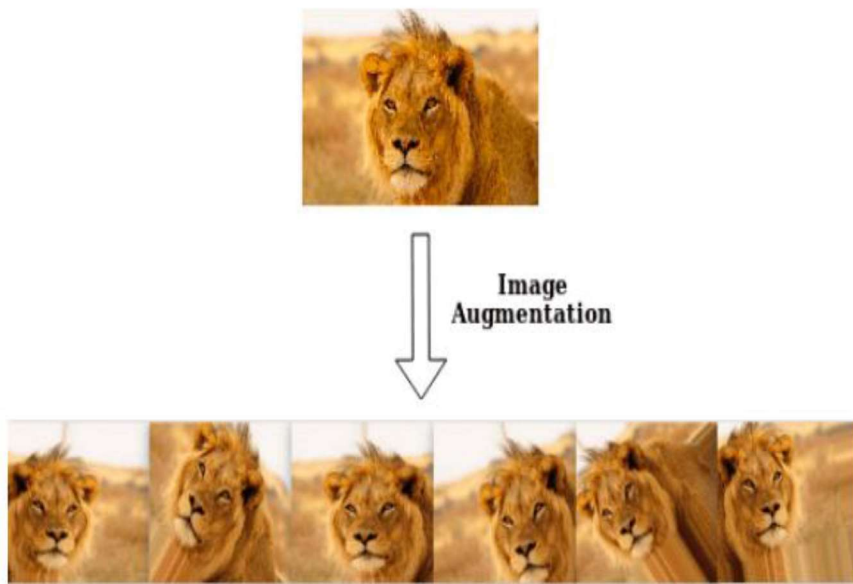
하지만 현실에서는 잘 일어나지 않는 문제임.

대부분의 데이터는 고차원의 데이터이므로, 모든 W 들이 국소 최적해에 동시에 있어야 하기 때문

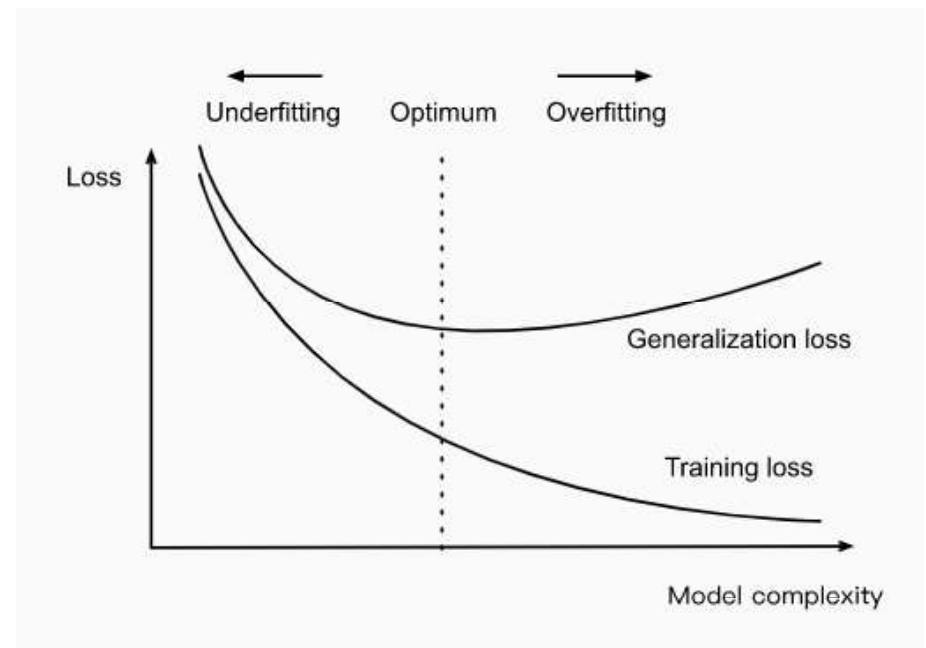
4 모델 성능 향상

● 모델 성능 향상

1. 데이터 수, 모델 구조



데이터 증강(augmentation) : 데이터 수 확보
Ex) 좌우반전, 자르기, 밝기 조절 등



모델 구조 (파라미터 줄이기) 단순화
→ 노드 개수, layer 개수 조절

4 모델 성능 향상

● 더 알아보기

가중치 초기화

:가중치의 초기값 범위를 최적의 가중치 범위에 근사하게 설정

$$r = \sqrt{\frac{3}{n_{in}}} [LeCun1998]$$
$$r = \sqrt{\frac{6}{n_{in} + n_{out}}} [Glorot2010]$$
$$r = \sqrt{\frac{6}{n_{in}}} [KaimingHe2015]$$

Uniform distribution

$$r = \sqrt{\frac{1}{n_{in}}} [LeCun1998]$$
$$r = \sqrt{\frac{2}{n_{in} + n_{out}}} [Glorot2010]$$
$$r = \sqrt{\frac{2}{n_{in}}} [KaimingHe2015]$$

gaussian distribution

가중치 초기화를 최적의 가중치 범위에 가깝게 하면 더욱 빠른 학습이 이루어짐

∴ 위의 r을 이용하여 $(-r, r)$ 범위로 초기화를 하는 것이 권장됨