

Gradient Descent

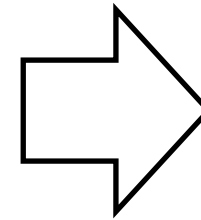
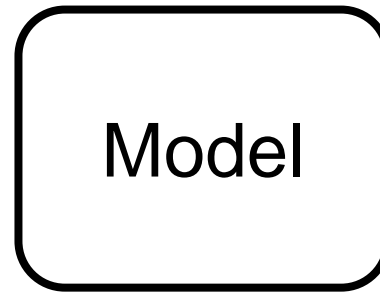
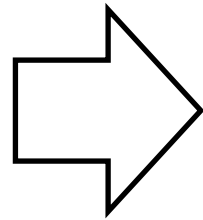
ML Lab

Table of Contents

- Machine learning
- Linear regression
- Gradient descent

Neural Network is Function (1)

- Classification

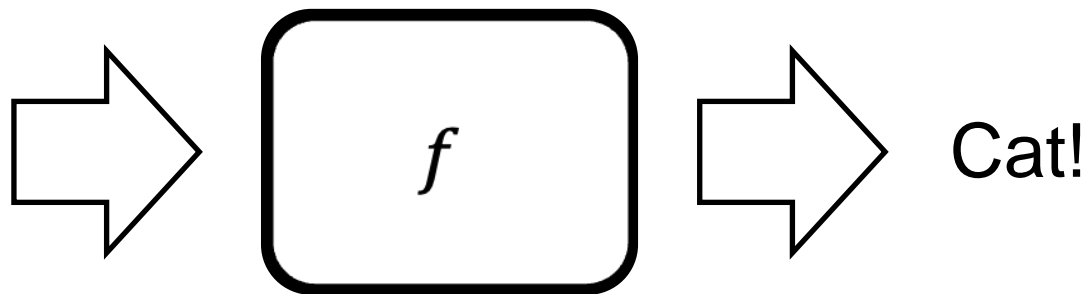


Cat!

- $f(\text{cat image}) = \text{cat}, f(\text{dog image}) = \text{dog}$

Neural Network is Function (2)

- Label된 이미지들 $\{(x_i, y_i)\}_{i=1}^N$ 가 주어졌을 때, 우리는 가장 좋은 neural network f 를 찾고 싶다!



How to Optimize Model? (1)

- 우리가 원하는 f 를 찾기 전에, 다음 부분들을 정의해야 합니다:
 - f 의 정확한 수식.
 - f 를 평가하는 함수.
- 예를 들어, 다음과 같이 정의할 수 있습니다:
 - $f(x) = Wx^T + b$ (linear regression model).
 - $l(f) = \sum_{i=1}^N (f(x_i) - y_i)^2$ (mean-squared error).

How to Optimize Model? (2)

$$f(x) = Wx^T + b$$

- 앞에서 정의한 식들을 가지고 문제를 구체적으로 정의하면 다음과 같다:
 - 주어진 data들에 대해서 $\sum_{i=1}^N (f(x_i) - y_i)^2$ 를 최소화하는 W, b 를 찾고 싶다!
 - Q) 어떻게 찾을까?

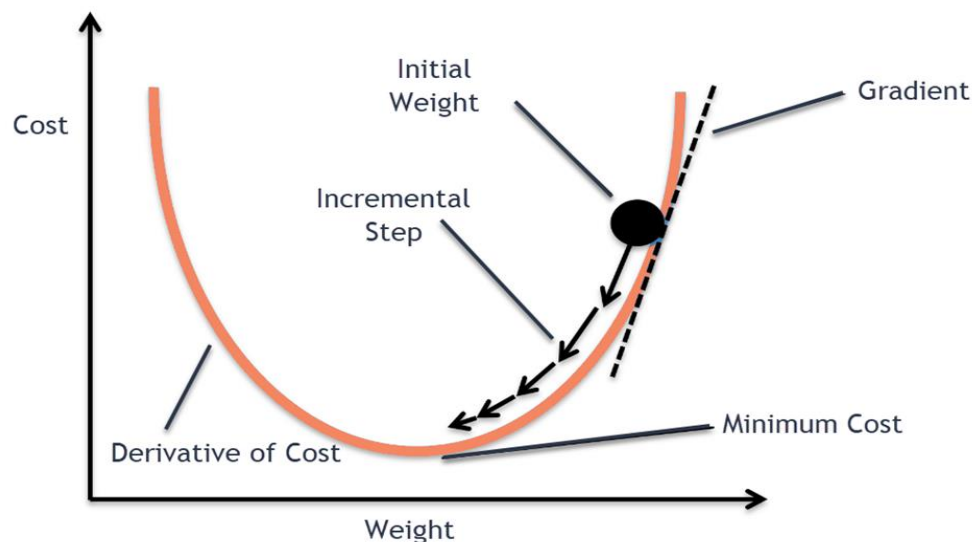
Gradient Descent (1)

$$l(f) = \sum_{i=1}^N (f(x_i) - y_i)^2$$

- 일반적으로 $l(W, b)$ 를 최소화하는 W, b 는 다음 성질을 만족한다:
 - $\frac{\partial l(W, b)}{\partial W} = \frac{\partial l(W, b)}{\partial b} = 0$.
 - Gradient descent는 이러한 성질을 만족하는 W, b 를 찾는 방법론 중 하나다!

Gradient Descent (2)

- Gradient descent는 다음 과정을 통해 우리가 원하는 W, b 를 찾는다:
 1. W, b 를 랜덤하게 초기화 한다.
 2. Gradient 를 계산한다.
 3. 구한 gradient를 가지고 업데이트한다: $W = W - \eta \frac{\partial l(W,b)}{\partial W}$.



Practice: Implement Gradient Descent

- Logistic regression을 gradient descent로 구현하기

$$y = \sigma(Wx + b)$$

$$z = w^T x + b$$

$$\hat{y} = a = \sigma(\underline{z})$$

$$\mathcal{L}(a, y) = -(y \log(a) + (1 - y) \log(1 - a))$$

Thank You :)