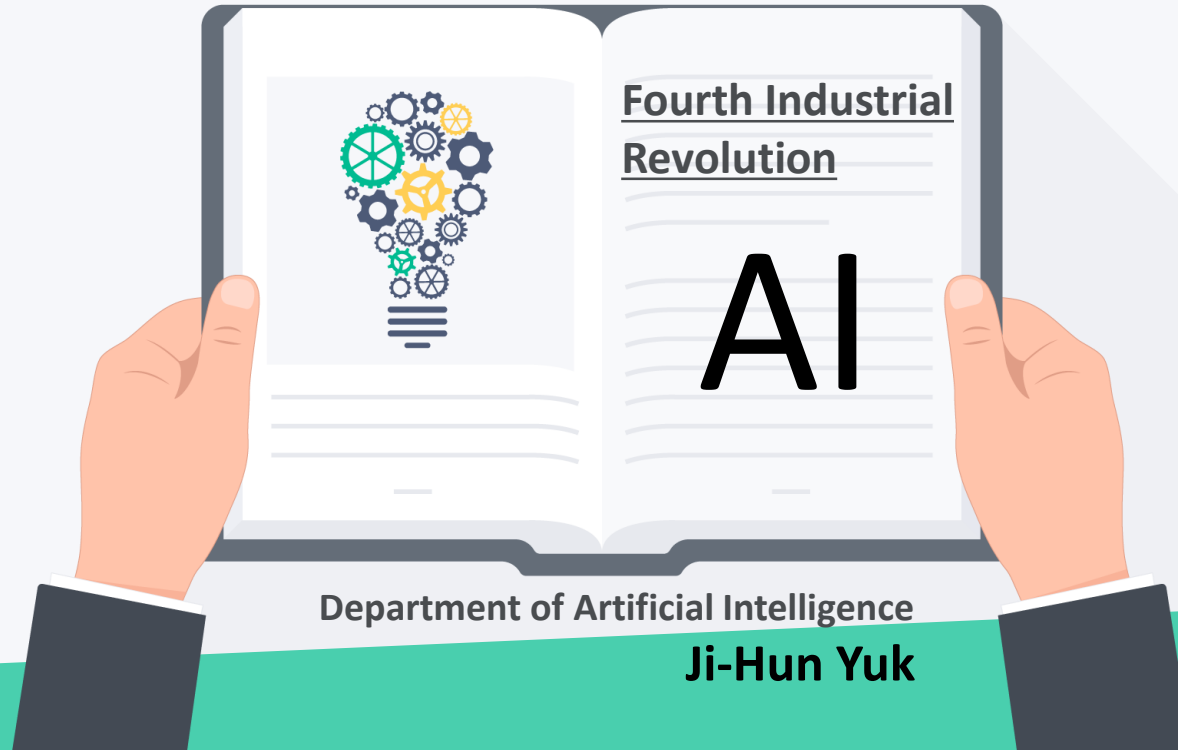


# What is Artificial Intelligence(AI)

affective.AI Lab



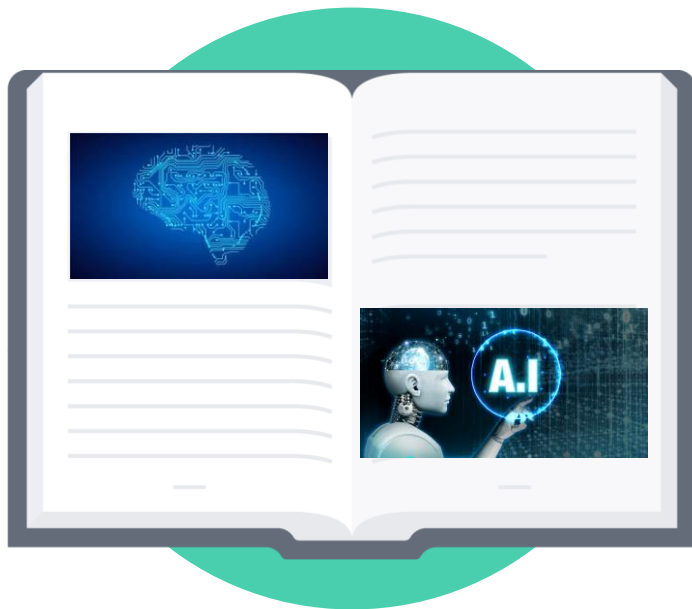
Department of Artificial Intelligence

Ji-Hun Yuk

6 July 2023

# Table of contents

목 차



**1. Definition of AI**

**2. What is Machine Learning?**

**3. Weekly study**

**4. Summary**

**5. Q&A**

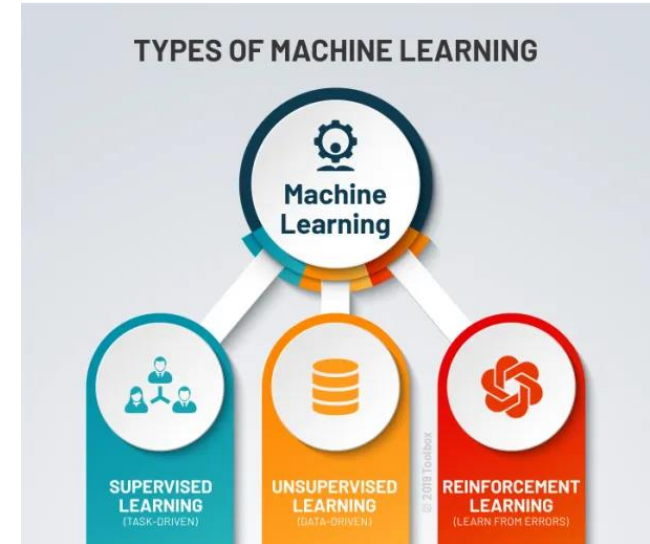
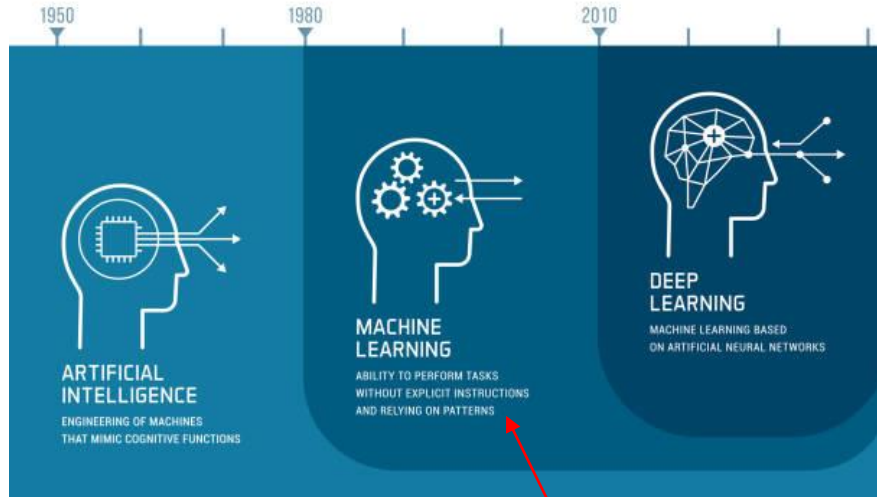
# 1.

## What is AI?

### **Definition:**

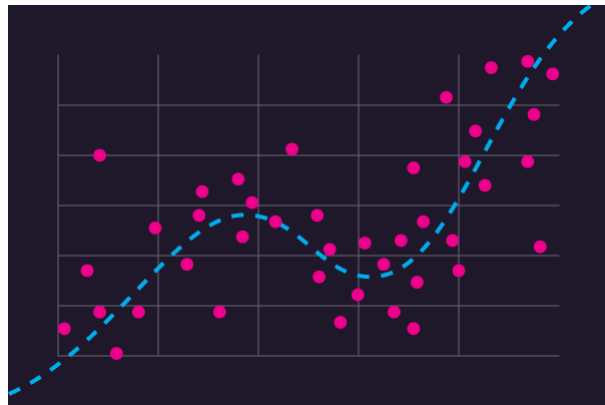
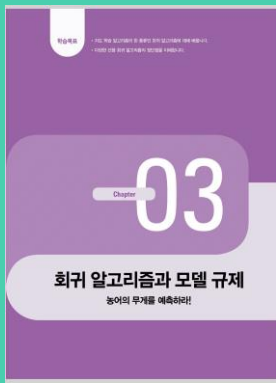
**Artificial intelligence (AI) is intelligence—perceiving, synthesizing, and inferring information—demonstrated by machines, as opposed to intelligence displayed by humans or by other animals.**

# 2. What is Machine Learning?

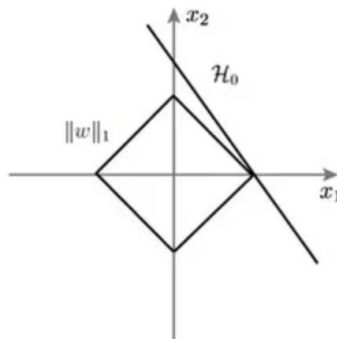


# 3-0. Weekly study

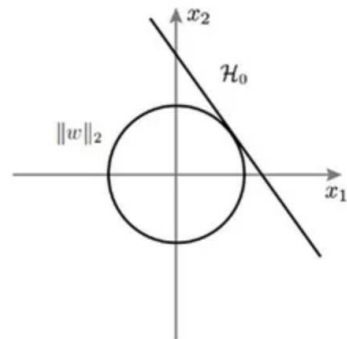
## Ch3. 회귀 알고리즘과 모델 규제



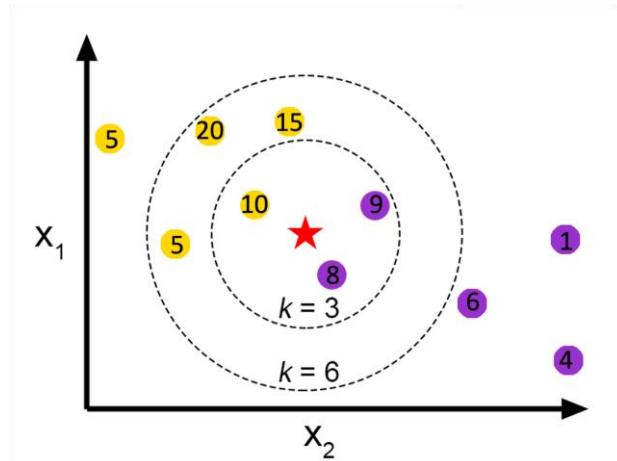
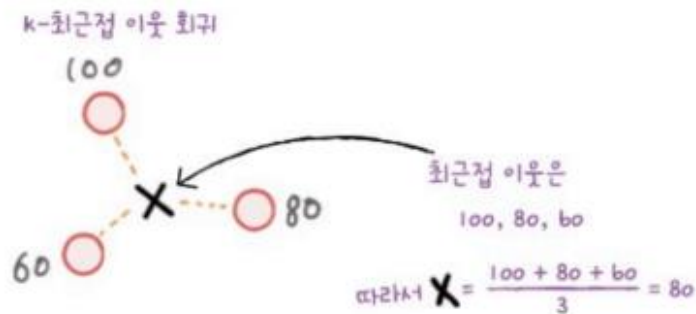
A L1 regularization



B L2 regularization



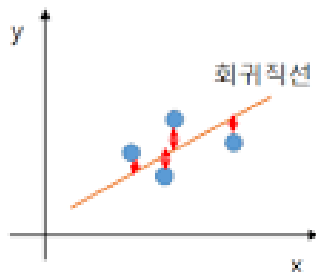
# 3-1. k-최근접 이웃 회귀



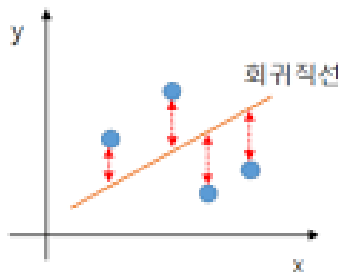
# 3-1. 결정계수( $R^2$ )

- 회귀 모델의 설명력( $0 \leq R^2 \leq 1$ ).
- 구해진 회귀식을 “얼마나 납득할 수 있는가”를 말함.

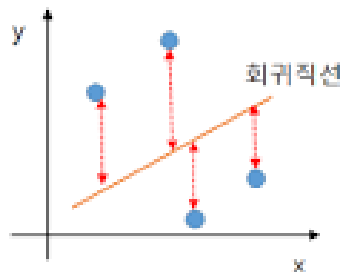
$$R^2 = 1 - \frac{(\text{타깃} - \text{예측})^2 \text{의 합}}{(\text{타깃} - \text{평균})^2 \text{의 합}}$$



$R^2$ 가 1에 가까운 경우

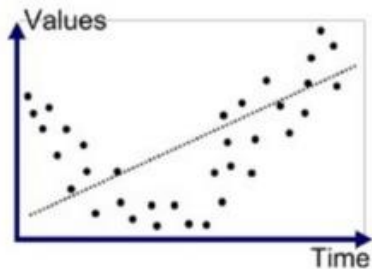


$R^2$ 가 0.5에 가까운 경우

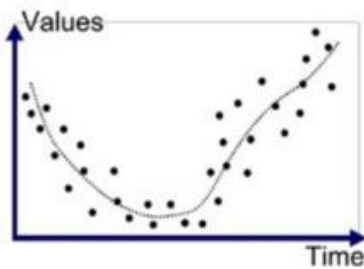


$R^2$ 가 0에 가까운 경우

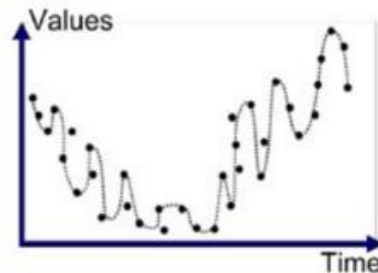
# 3-1. 과대적합 vs 과소적합



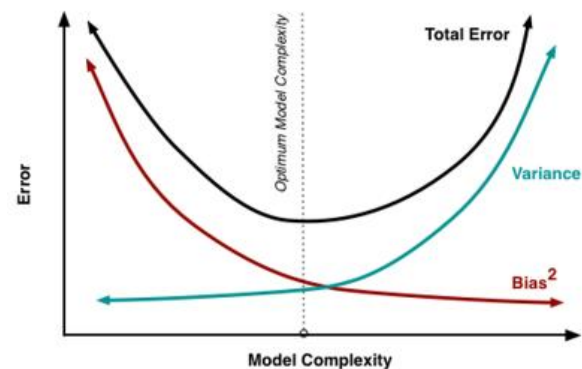
Underfitted



Good Fit/Robust

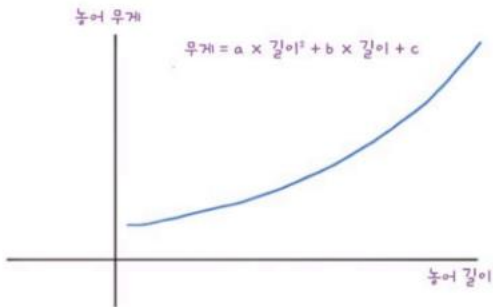
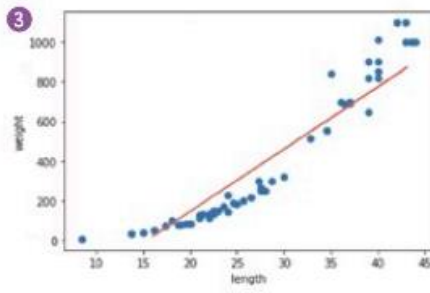
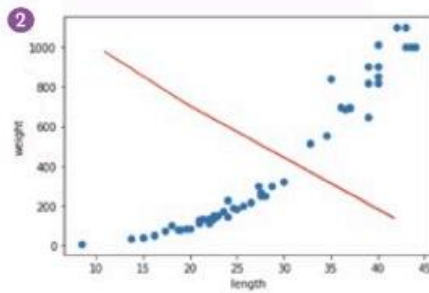
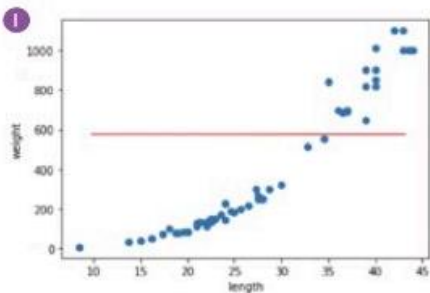


Overfitted



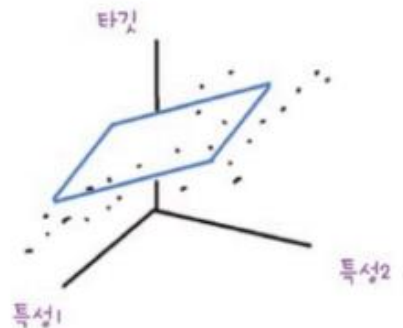
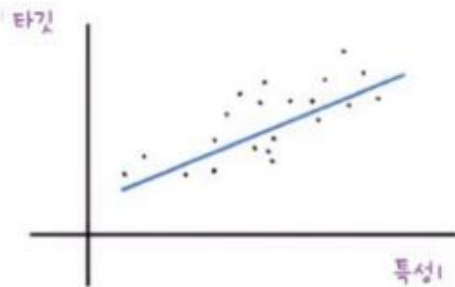


# 3-2. 선형 회귀



# 3-3. 특성 공학

- 다중 회귀



- 특성 공학(feature engineering)  
:기존의 특성을 사용해 새로운 특성을 뽑아내는 작업.

# 3-3. Norm

$$\|\mathbf{x}\|_p := \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}.$$

- 벡터의 크기를 측정하는 방법.
- 두 벡터 사이의 거리를 측정하는 방법.
- P는 Norm의 차수.
- P = 1 이면 L1 Norm, P = 2 이면 L2 Norm.

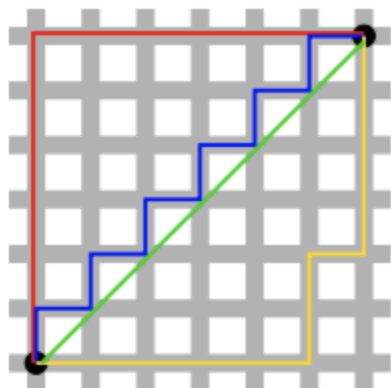
## • L1 Norm

## • L2 Norm

$$d_1(\mathbf{p}, \mathbf{q}) = \|\mathbf{p} - \mathbf{q}\|_1 = \sum_{i=1}^n |p_i - q_i|, \quad \|\mathbf{x}\|_2 := \sqrt{x_1^2 + \cdots + x_n^2}.$$

- L1 Norm은 벡터 p, q의 각 원소들의 차이의 절대값의 합.
- L2 Norm은 벡터 p, q의 유클리디안 거리(직선거리).

# 3-3. L1 Norm과 L2 Norm의 차이



- 검정색 두 점 사이의 L1 Norm은 R,B,Y 선 등 여러 표현 가능.
- L2 Norm은 오직 초록색 선으로만 표현.

# 3-3. Loss

- L1 Loss

$$L = \sum_{i=1}^n |y_i - f(x_i)|$$

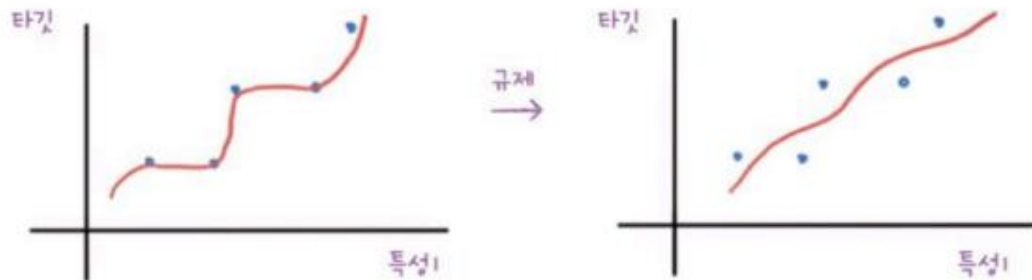
- L2 Loss

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

- $y_i$ 는 실제값,  $f(x_i)$ 는 예측값.
- L1 Loss는 실제값과 예측값 사이의 오차의 절대값의 합.
- L2 Loss는 오차의 제곱의 합.
- L2 Loss는 L1 Loss보다 Outlier(이상치)에 더 큰 영향을 받는다.
- Outlier가 적당히 무시되길 원한다면 L1 Loss를 사용.
- But L1 Loss는 0인 지점에서 미분이 불가능하다는 단점을 가지고 있다.

# 3-3. Regularization (규제)

- 규제는 머신러닝 모델이 훈련 세트를 너무 과도하게 학습하지 못하도록 뒤편하는 것
- 즉 모델이 훈련 세트에 과대적합되지 않도록 만드는 것



# 3-3. 라쏘 회귀

- Least Absolute Shrinkage and Selection Operater(Lasso)
- L1 Regularization 사용
- Cost function

$$C = C_0 + \frac{\lambda}{n} \sum_w |w|$$

- $w$ 에 대해 편미분

$$w \rightarrow w - \frac{\eta \lambda}{n} \text{sgn}(w) - \eta \frac{\partial C_0}{\partial w}$$

$$\text{sgn}(x) = \begin{cases} -1 & \text{for } x < 0 \\ 0 & \text{for } x = 0 \\ 1 & \text{for } x > 0. \end{cases}$$

# 3-3. 릿지 회귀

- L2 Regularization 사용
- Cost function

$$C = C_0 + \frac{\lambda}{2n} \sum_w w^2$$

- $w$ 에 대해 편미분

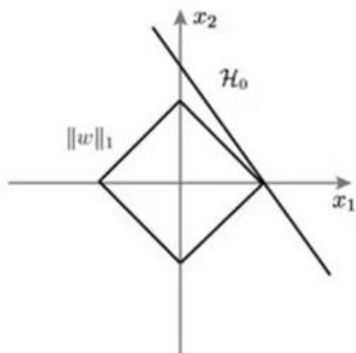
$$w \rightarrow w - \eta \frac{\partial C_0}{\partial w} - \frac{\eta \lambda}{n}$$

$$= \left(1 - \frac{\eta \lambda}{n}\right) w - \eta \frac{\partial C_0}{\partial w}$$

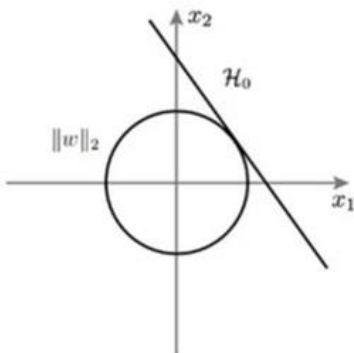


# 3-3. L1 규제와 L2 규제의 차이

A L1 regularization



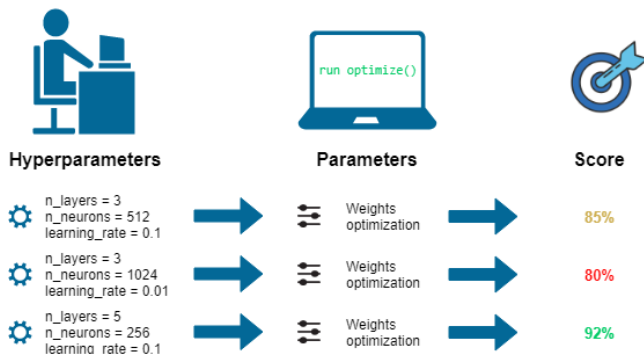
B L2 regularization



- L1 규제는 가중치 업데이트 시, 가중치의 크기에 상관 없이 상수값을 빼면서 진행되기 때문에 작은 가중치들은 거의 0으로 수렴 되어 중요한 가중치들만 남게 된다.
- 하지만 L1 규제는 위 그림처럼 미분 불가능한 점이 있기 때문에 기울기 연산에 위험이 있다.
- L2 규제는 가중치 업데이트시, 가중치의 크기가 직접적인 영향을 미치므로 L1 보다 가중치 규제에 효과적이고, 미분 연산의 위험성이 낮다.

# 3-3. 하이퍼 파라미터

- 하이퍼파라미터는 머신러닝 알고리즘이 학습하지 않는 파라미터로 모델 설계자가 사전에 지정해야한다.
- 최적의 모델을 찾기 위해서 실험적으로 알아내야한다.
- (e.g. 학습률, 손실함수, 알파&람다 파라미터, 미니배치 크기, 에포크 수, 가중치 초기화, 은닉층의 개수, Dropout 비율, K값.)



$\lambda$

$\mu$

$\alpha$

# 4.

## Summary



# 5.

## Q&A



# Thanks for listening

