

머신러닝

.



# 목차

1. AI, 머신러닝, 딥러닝
2. 지도학습, 비지도학습, 강화학습
3. 지도학습 - 선형 회귀(Linear Regression)
4. 지도학습 - 의사결정 트리(Decision Tree)
5. 지도학습 - K-nearest neighbor classifier(k-최근접 이웃 분류기)
6. 비지도학습 - K-means Clustering(평균 군집)

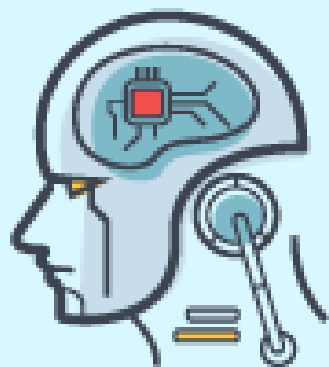


# AI, 머신러닝, 딥러닝

Artificial Intelligence

## 인공지능

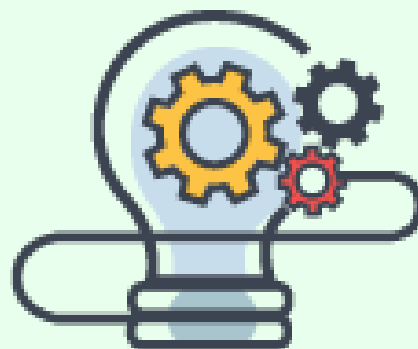
사고나 학습 등 인간이 가진  
지적 능력을 컴퓨터를 통해  
구현하는 기술



Machine Learning

## 머신러닝

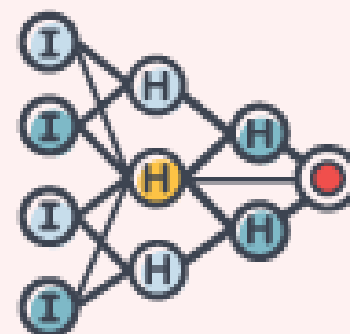
컴퓨터가 스스로 학습하여  
인공지능의 성능을  
향상 시키는 기술 방법



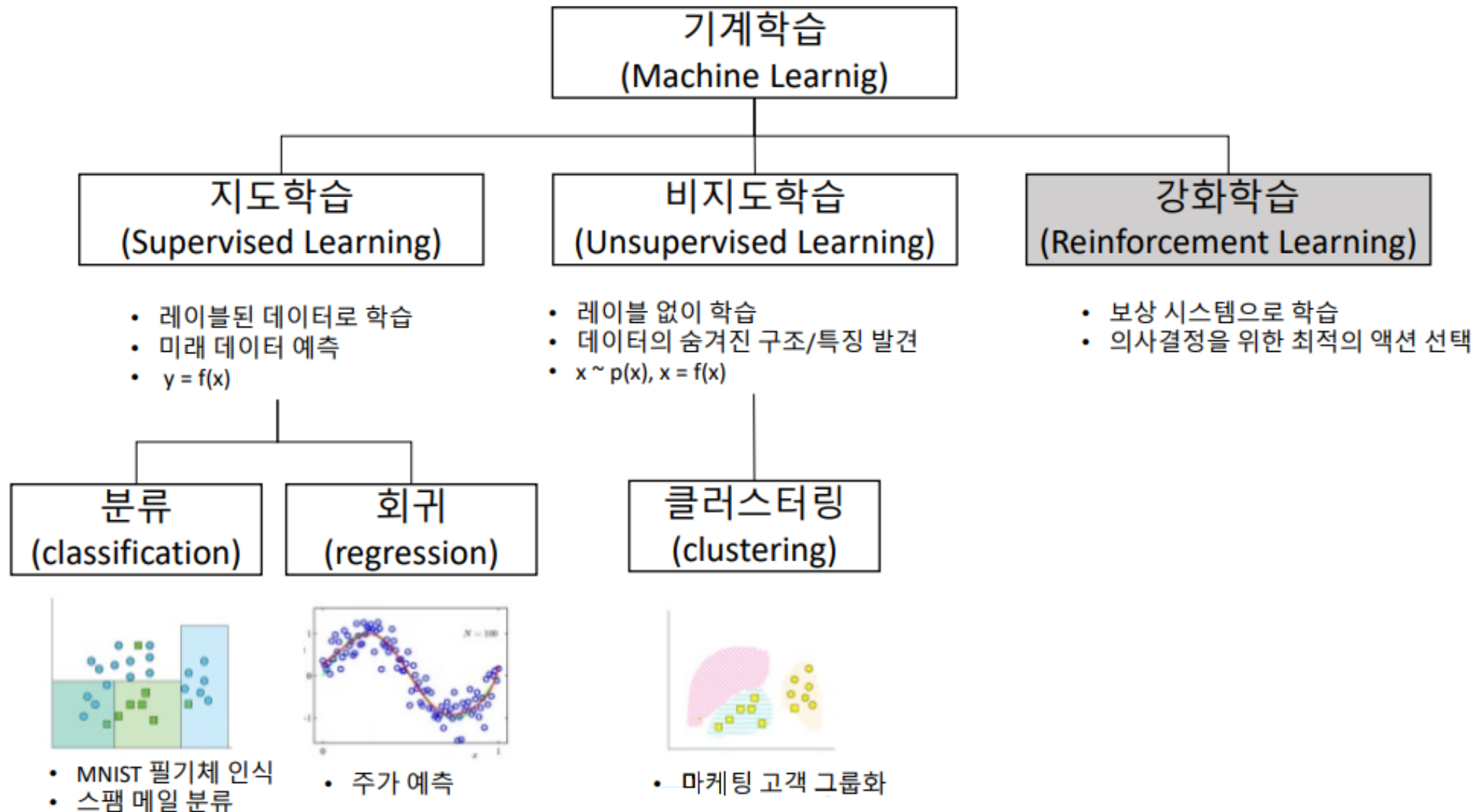
Deep Learning

## 딥러닝

인간의 뉴런과 비슷한  
인공신경망 방식으로  
정보를 처리



# 지도학습, 비지도학습, 강화학습

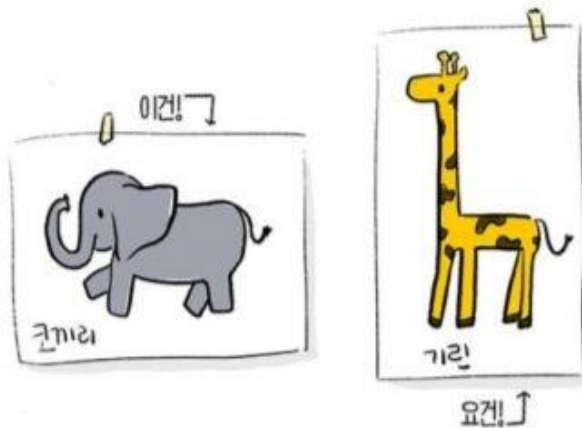


# 지도학습, 비지도학습, 강화학습

## 머신 러닝

### 지도 학습 (Supervised Learning)

문제와 정답을 모두 알려주고  
공부시키는 방법

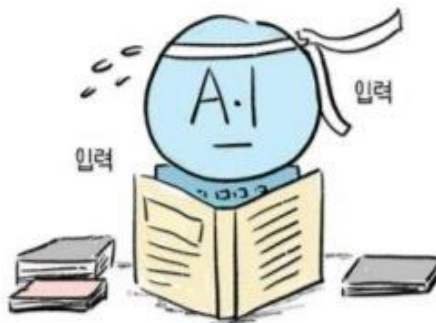


예측, 분류

### 비지도 학습 (Unsupervised Learning)

답을 가르쳐주지 않고  
공부시키는 방법

비지도학습은 답을 가르쳐주지 않고 공부를 시키는거야.



연관 규칙, 군집

### 강화 학습 (Reinforcement Learning)

보상을 통해  
상은 최대화, 벌은 최소화하는  
방향으로 행위를 강화하는 학습

강화학습은 일종의 게임 같이 보상해주는거야

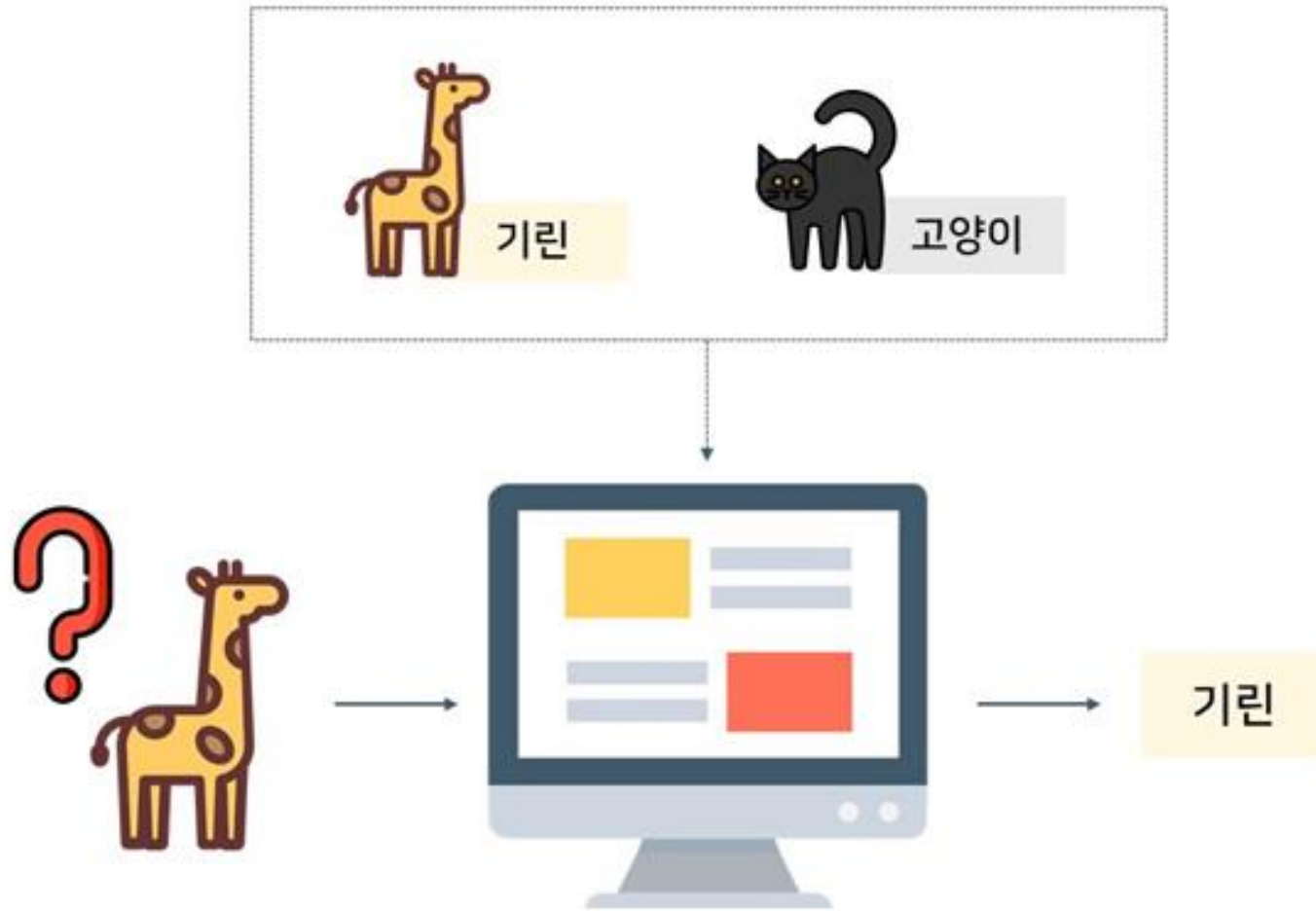


보상

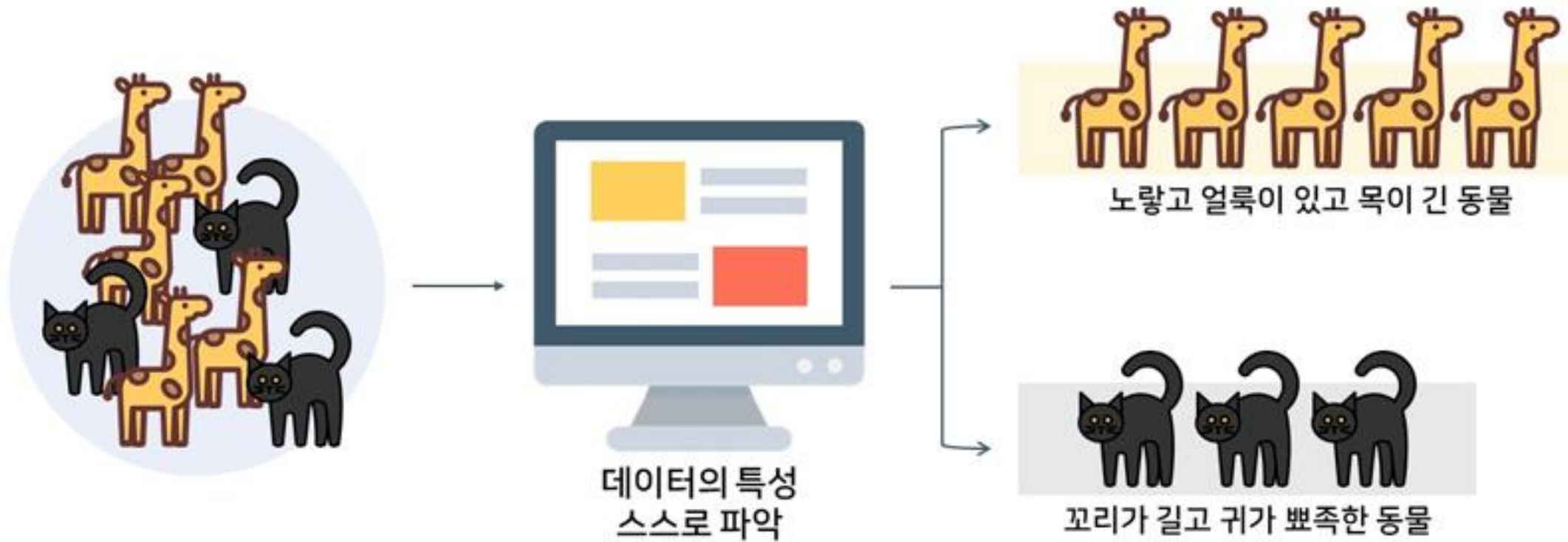


# 지도학습

정답 미리 학습



# 비지도학습



# 지도학습(예측) – 선형 회귀(Linear Regression)

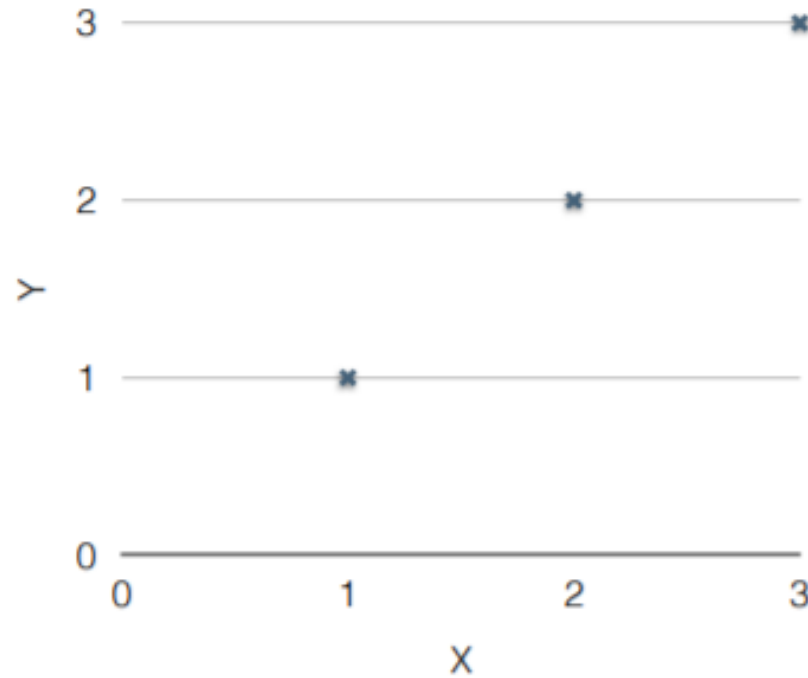
주어진 데이터를 이용해 일차방정식을 수정해 나가는 것

- 학습을 거쳐서 가장 합리적인 선을 찾아내는 것
- 학습을 많이 해도 '완벽한' 식을 찾아 내지 못할 수 있다.
- 가장 근사치의 값을 찾는 것



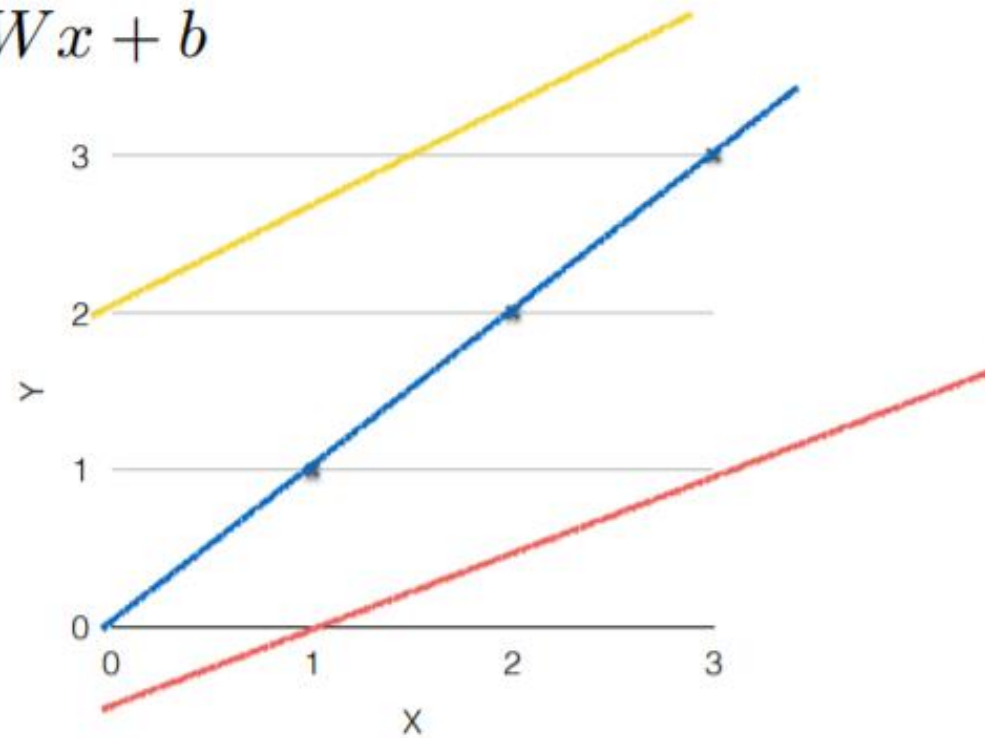
# 지도학습 – 선형 회귀(Linear Regression)

| X | Y |
|---|---|
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |

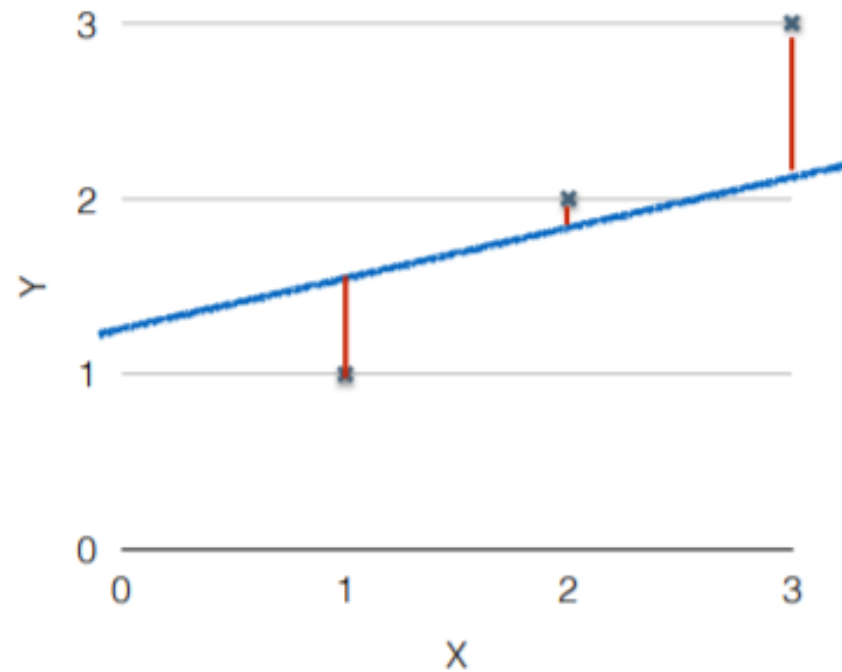


# (Linear) Hypothesis

$$H(x) = Wx + b$$



# Which hypothesis is better?



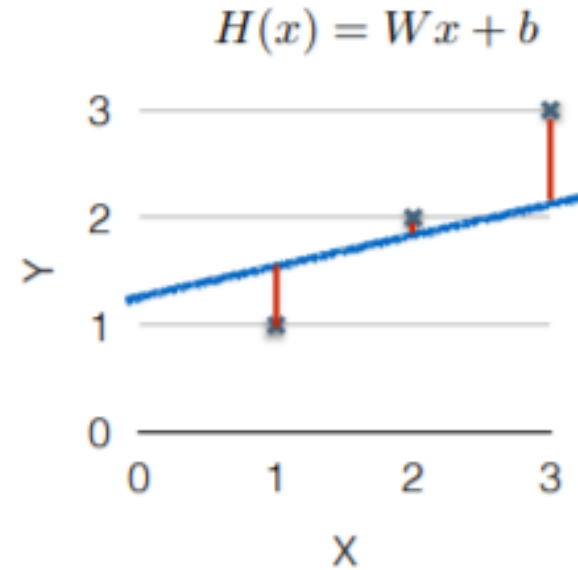
# 비용 함수(Cost Function)

→ (예측값-실제값)<sup>2</sup> 의 평균

- How fit the line to our (training) data

$$\frac{(H(x^{(1)}) - y^{(1)})^2 + (H(x^{(2)}) - y^{(2)})^2 + (H(x^{(3)}) - y^{(3)})^2}{3}$$

$$cost = \frac{1}{m} \sum_{i=1}^m (H(x^{(i)}) - y^{(i)})^2$$



# Cost function

$$cost = \frac{1}{m} \sum_{i=1}^m (H(x^{(i)}) - y^{(i)})^2$$

$$H(x) = Wx + b$$

$$cost(W, b) = \frac{1}{m} \sum_{i=1}^m (H(x^{(i)}) - y^{(i)})^2$$

Goal: Minimize cost

$$\underset{W, b}{\text{minimize}} \text{cost}(W, b)$$



# Simplified hypothesis

$$H(x) = Wx$$

$$\text{cost}(W) = \frac{1}{m} \sum_{i=1}^m (Wx^{(i)} - y^{(i)})^2$$

# What $\text{cost}(W)$ looks like?

$$\text{cost}(W) = \frac{1}{m} \sum_{i=1}^m (Wx^{(i)} - y^{(i)})^2$$

| x | Y |
|---|---|
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |

- $W=1, \text{cost}(W)=0$

$$\frac{1}{3}((1 * 1 - 1)^2 + (1 * 2 - 2)^2 + (1 * 3 - 3)^2)$$

- $W=0, \text{cost}(W)=4.67$

$$\frac{1}{3}((0 * 1 - 1)^2 + (0 * 2 - 2)^2 + (0 * 3 - 3)^2)$$

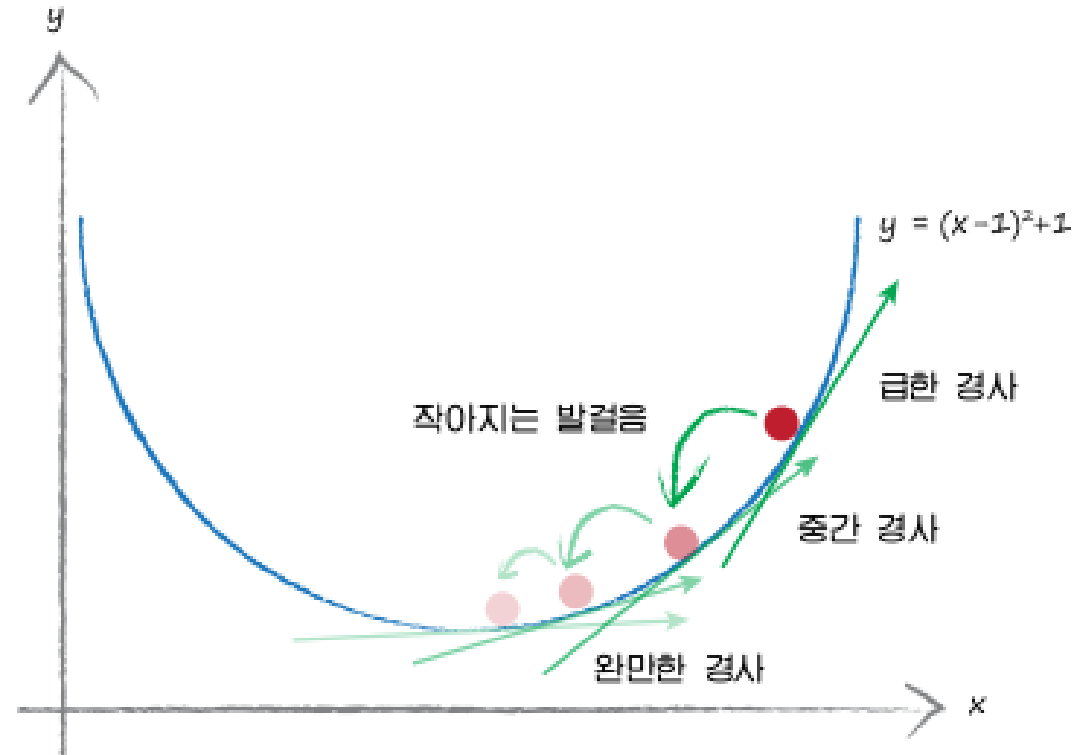
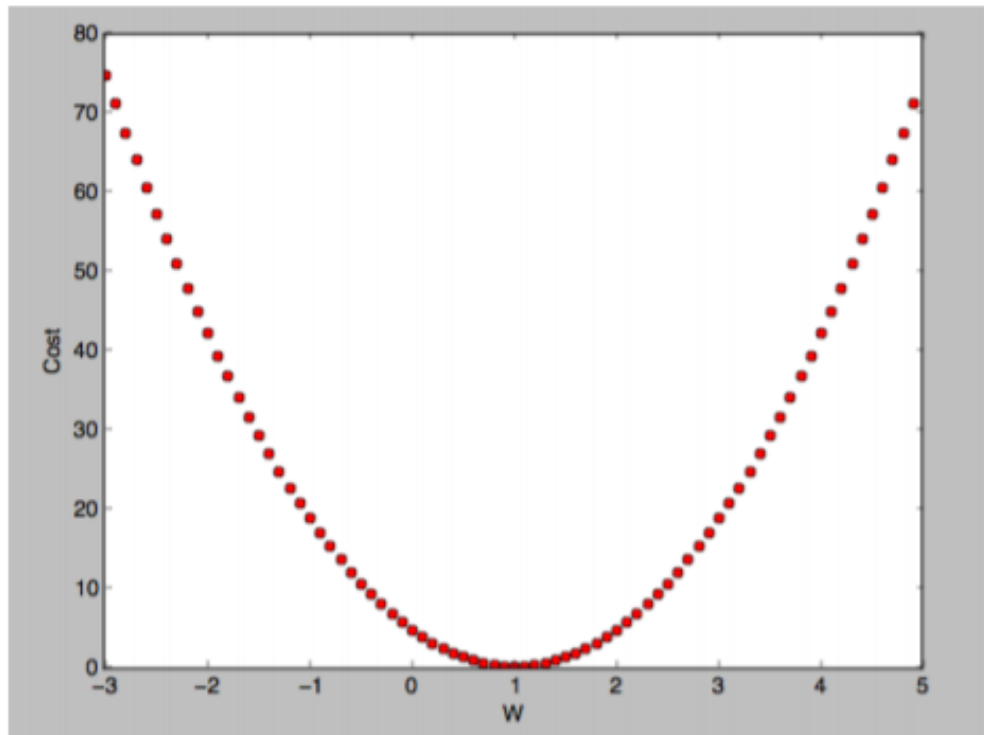
- $W=2, \text{cost}(W)=?$

- $W=1, \text{cost}(W)=0$
- $W=0, \text{cost}(W)=4.67$
- $W=2, \text{cost}(W)=4.67$

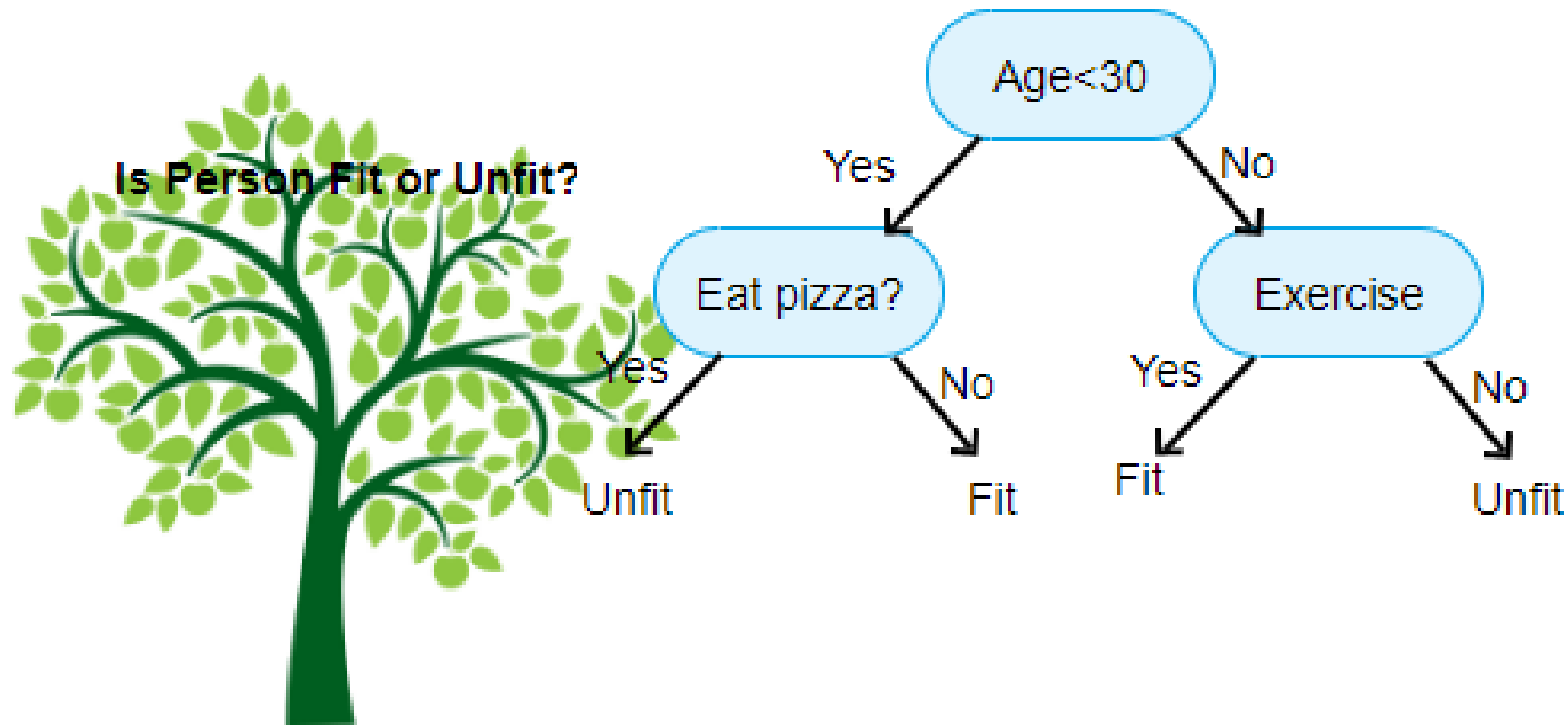
# 경사하강법(Gradient descent algorithm)

What  $\text{cost}(W)$  looks like?

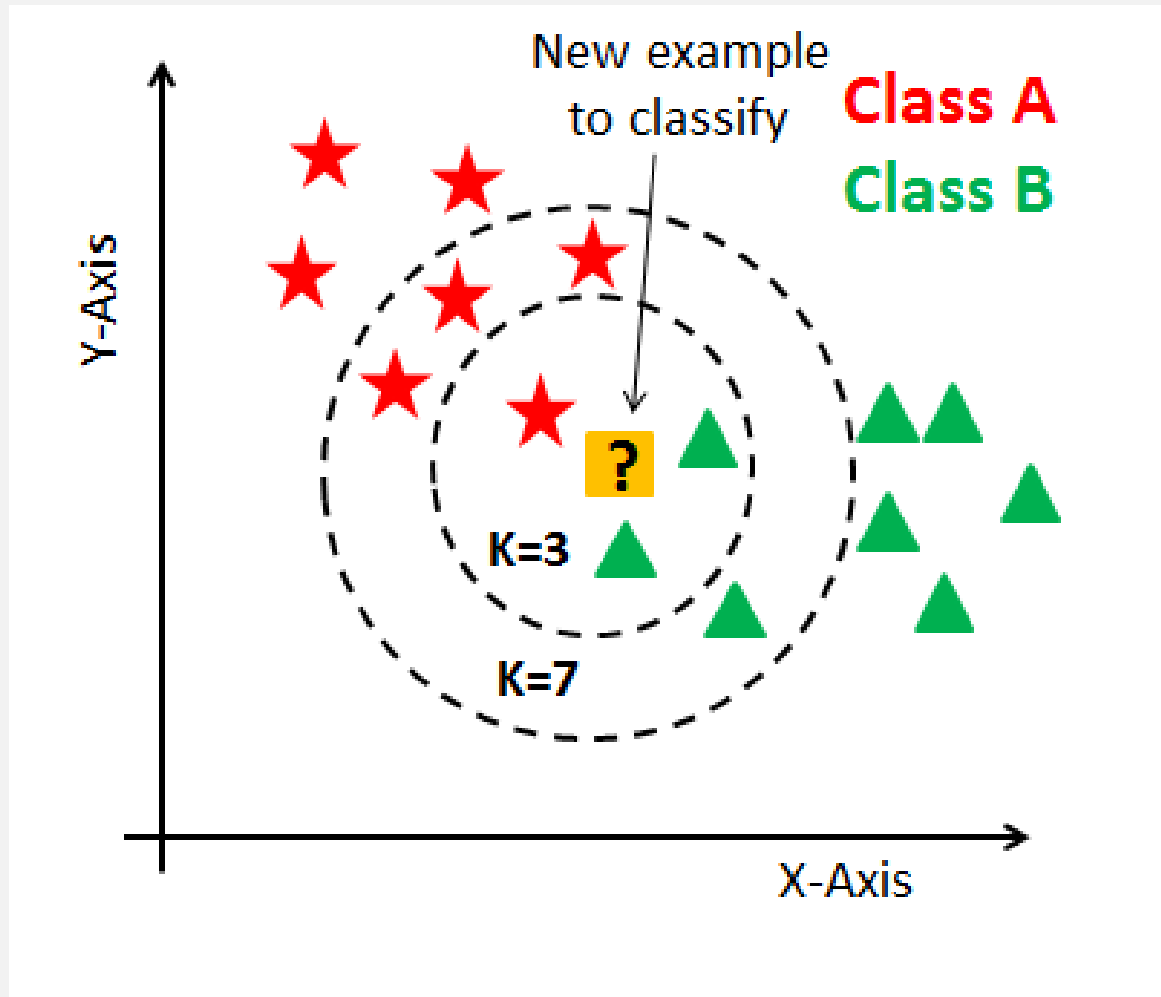
$$\text{cost}(W) = \frac{1}{m} \sum_{i=1}^m (Wx^{(i)} - y^{(i)})^2$$



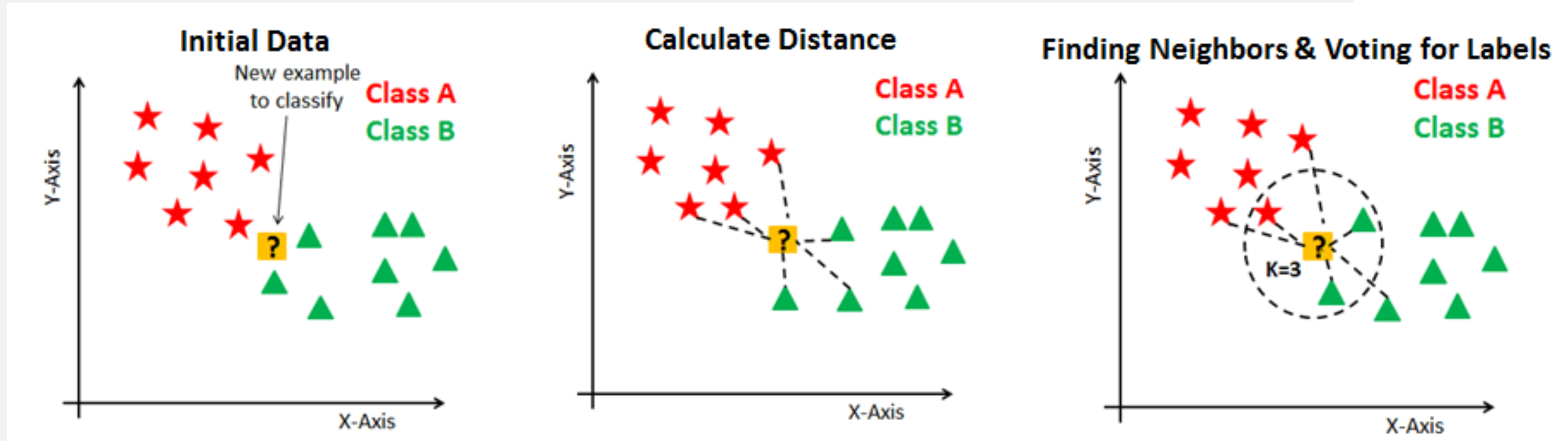
# 지도학습(분류) - 의사결정 트리(Decision Tree)



# 지도학습(분류) - K-nearest neighbor classifier(k-최근접 이웃 분류기)



# 지도학습(분류) - K-nearest neighbor classifier(k-최근접 이웃 분류기)





# 비지도학습(군집) - K-means Clustering(평균 군집)

KMeans 클러스터링 알고리즘은  $n$ 개의 중심점을 찍은 후에, 이 중심점에서 각 점간의 거리의 합이 가장 최소화 되는 중심점  $n$ 의 위치를 찾고, 이 중심점에서 가까운 점들을 중심점을 기준으로 묶는 클러스터링 알고리즘이다.

