

machine type  $\rightarrow$  분류 / 상관관계 파악 / 회귀

machine task  $\rightarrow$  감지/탐색 / 내감지/탐색 / 분류/탐색 / 객체/탐색

machine type  $\rightarrow$  선형성 / SUM / KLU / PCA 등

분류 : 데이터를 특정 카테고리 분류.

(식별) 분류 대상 카테고리 (클래스)는 사전에 지정함.

카테고리의 인스턴스가 속한 개체 / 개체일 때 속성을 의미

상관관계 분석.

· 선형성.

: 주어진 데이터 포인트 집합  $\{(x, y)\}$ 가 주어진 때

독립변수  $x$ 와 종속변수  $y$ 와의 관계를 나타내는 선형성을  
판단함.

- 단순 (다중) 선형성 (독립변수 개수)

- 양(음)성 / 미분 가능 선형성 (종속변수 개수)

· 단순 선형성.

일부 데이터  $X = \{x_k = (x, y) \mid k = 1, \dots, N\}$

$\hat{y} = f(x) = \beta_0 + \beta_1 x$ ,  $\beta_0, \beta_1$ 을 구함.

독립변수가  $a$ 개인 다중 선형성

$X = \{x_k = (x_1, y_1) \mid k = 1, \dots, N\}$

$\hat{y} = f(x) = \sum_{i=0}^a \beta_i x_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_a x_a, x_0 = 1$

여기  $\beta_1, \beta_2, \dots, \beta_a$ 를 구함

입력데이터:  $x_k = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_a \end{bmatrix}$

## 로지스틱 회귀

분류변수가 독립변 (각 클래스를 위하여 매핑해준다) Y 값  
class 1, 0

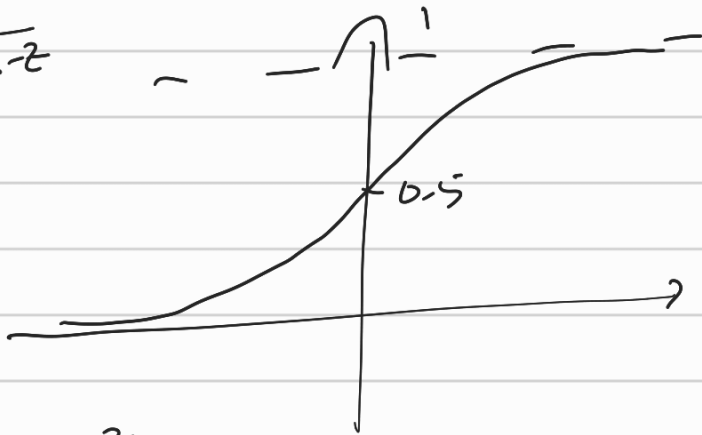
- 예/아니  $\begin{cases} \text{Yes} : 1 \\ \text{No} : 0 \end{cases}$

- 클래스 :  $C_1, C_2, \dots, C_n$

→ 각 클래스의 임계값을 찾아볼 수 있음.

• 시그모이드 함수를 매핑할 확률 (p) 값.

- 임의의 독립변 집합에 대해 선형 매핑을  $z = \sum_{i=1}^n \beta_i x_i$   
$$p(z) = \frac{1}{1 + e^{-z}}$$



$$\text{Odds} = \frac{p}{1-p} = e^z$$

$$\text{logit} : \ln\left(\frac{p}{1-p}\right) = z = -\sum_{i=1}^n \beta_i x_i$$

↳ logit는 {독립변의 선형조합}

Odds를 통해

Logit를 통해 매개변수

## 군집화 (Clustering)

∴ 여러 개의 클러스터가 여러 개의 패턴을 형성

## 감독학습. (Supervised Learning)

지도학습이라고도 함.

데이터  $X$ 와 답  $(f)$ 를 각각 학습데이터  $(x, y)$ 를 구성

→  $x$ 의 값을 통해  $y$ 의 값을 예측하는 예측 모델.

## 비감독학습.

: 미리 없이 데이터만으로 학습.

데이터에 나타나는 [특징]을 찾아, 공통점을 찾는

특징 학습의 방법들을 위해 비지도 학습이라고도 함

→ 유사도 기반 클러스터링, 연관분석

## 강화학습. (Reinforcement Learning)

: 어떤 상태 (State)의 환경의 보상을 최대화 하는 동작을  
매우 많은 시행을 통해 학습.

Model : (Markov, Decision, Process)

- 상태공간  $S$

작용공간  $A$

보상 함수 :  $R : S \times A \rightarrow \mathbb{R}$

상태-작용 공간.

비선형 기법도 있음,

