

**머신러닝**

**Machine Learning**

# 머신러닝이란?

---

- 머신러닝은 데이터의 훈련 샘플들을 학습해 데이터의 규칙과 패턴을 학습하여 결과를 예측하는 알고리즘과 시스템을 연구하는 분야이다.
- 통계학에서 유래된 머신러닝 기법들이 많다.
- 머신러닝에서는 데이터에 대한 규칙을 프로그래밍으로 구현하지 않아도 파이썬 API 기반의 라이브러리로 손쉽게 모델을 구축할 수 있다.

## 머신러닝

### Machine Learning

컴퓨터가 스스로 학습하여  
인공지능의 성능을  
향상 시키는 기술 방법



## 지도학습

### Supervised

Target 예측을 위해  
데이터를 학습시킴

## 비지도학습

### Unsupervised

Target이 없이  
데이터의 특성을 추출

# 머신러닝이란?

## Step 1 데이터 확인

연속형/범주형  
/회귀/분류/비지도

### Step 1 실전팁

- 독립변수, 종속변수 확인
- 연속형 vs 범주형 확인
- 적용가능한 분석모델 확인  
(회귀·분류·비지도 학습)

## Step 2 전처리

정규화/표준화  
/결측치·이상치 처리

### Step 2 실전팁

- 표준화(평균0, 표준편차1)  
또는 MinMax 정규화
- 결측치 확인 후 처리
- 이상치 확인 후 처리

## Step 3 데이터 분할

학습·검증·평가셋  
/교차검증 방법

### Step 3 실전팁

- 학습셋 : 60~80%
- 검증셋: 10~20%
- 평가셋: 10~20%
- 교차검증방법 적용 가능

## Step 4 모델학습

회귀/분류/비지도  
하이퍼파라미터 조절

### Step 4 실전팁

- 머신러닝 알고리즘 적용
- 회귀/분류/비지도 학습
- 최적모델을 결정하기 위해  
하이퍼파라미터 탐색·조절

## Step 5 성능평가

분석 정확도 확인  
/알고리즘 성능 제시

### Step 5 실전팁

- 평가셋에 최종모델을 적용
- 평가셋에 대한 정확도를  
머신러닝 분석에 대한  
성능으로 제시