

머신러닝 (ML)

1. 머신러닝이란

01 머신러닝이란?

- 알고리즘(algorithms) : 어떠한 문제를 해결하기 위한 일련의 **절차나 방법**
- 머신러닝(machine learning) : 기계가 **패턴을 학습**하여 자동화하는 알고리즘
- 유튜브는 개인이 유튜브 영상을 보는 패턴에 대해 학습하는 프로그램(머신러닝)을 만든 다음 그 패턴(알고리즘)에 맞게 다음 영상을 계속 추천

01 머신러닝이란?

2. 머신러닝의 실제 사용 사례

2.1 구매 추천

- 인터넷 쇼핑몰에서 장바구니에 추가한 제품과 비슷한 제품 구매를 추천

이 책을 구입하신 분들이 함께 산 책

이 책을 구입하신 분들이 많이 산 책

					
강원국의 어른답게 말 합니다	2000원으로 밥상 차리 기 (7월호) [202	손미나의 나의 첫 외국 어 수업	타인의 집	달리구르 품 벽화점	어떻게 말해줘야 할까
강원국 지 동진지식...	본집부 그리고작	손미나 지 후네이도	손원환 지 창비	이매예 지 박로리나연	오운영 지/차상미 그림
★★★★★ 39건	★★★★★ 21건	★★★★★ 23건	★★★★★ 5건	★★★★★ 418건	★★★★★ 199건
14,400원 (10% 할인)	2,850원 (5% 할인)	15,120원 (10% 할인)	12,600원 (10% 할인)	12,420원 (10% 할인)	15,750원 (10% 할인)

구입하려는 책과 관련된 또 다른 책을 추천하는 YES24의 사례

01 머신러닝이란?

2.2 번역

- 머신러닝에 전문 번역가의 번역을 학습시켜 새로운 문장을 번역하게 함



네이버 번역기 파파고(Papago)

01 머신러닝이란?

2.3 자율주행차

- 머신러닝 기술 중 하나인 이미지 처리 기술을 활용하여 도로상의 여러 이미지를 학습, 차량에 있는 컴퓨터가 스스로 판단하여 운행(최근에는 딥러닝 기술 활용)



머신러닝의 이미지 처리 기술을 활용한 자율주행차

01 머신러닝이란?

2.4 챗봇

- 머신러닝을 기반으로 사용자와 컴퓨터 간 대화를 지원
- 코로나19 사태로 인해 사람의 목소리를 흉내내어 응대하는 인공지능 콜센터 출현
 - 머신러닝이 인간의 대화 패턴과 목소리를 학습



챗봇 사용 사례

01 머신러닝이란?

3. 머신러닝의 키워드

3.1 인공지능, 머신러닝, 딥러닝

- 인공지능 ⊃ 머신러닝 ⊃ 딥러닝



01 머신러닝이란?

- **머신러닝**(machine learning) : 데이터를 컴퓨터에 학습시켜 그 **패턴과 규칙을 컴퓨터가 스스로 학습**하도록 만드는 기술
 - 이전에는 사람이 지식을 직접 데이터베이스화한 후 컴퓨터가 처리하도록 프로그램으로 만듦
 - 머신러닝은 데이터를 분류하는 **수학적 모델**을 프로그래밍하여, 데이터만 입력하면 **이미 만들어진 수학 모델**이 규칙으로 적용되어 여러 문제를 풀 수 있음
- **딥러닝**(deep learning) : 머신러닝 기법 중 신경망(neural network)을 기반으로 사물이나 데이터를 **군집화**하거나 **분류**하는 데 사용하는 기술

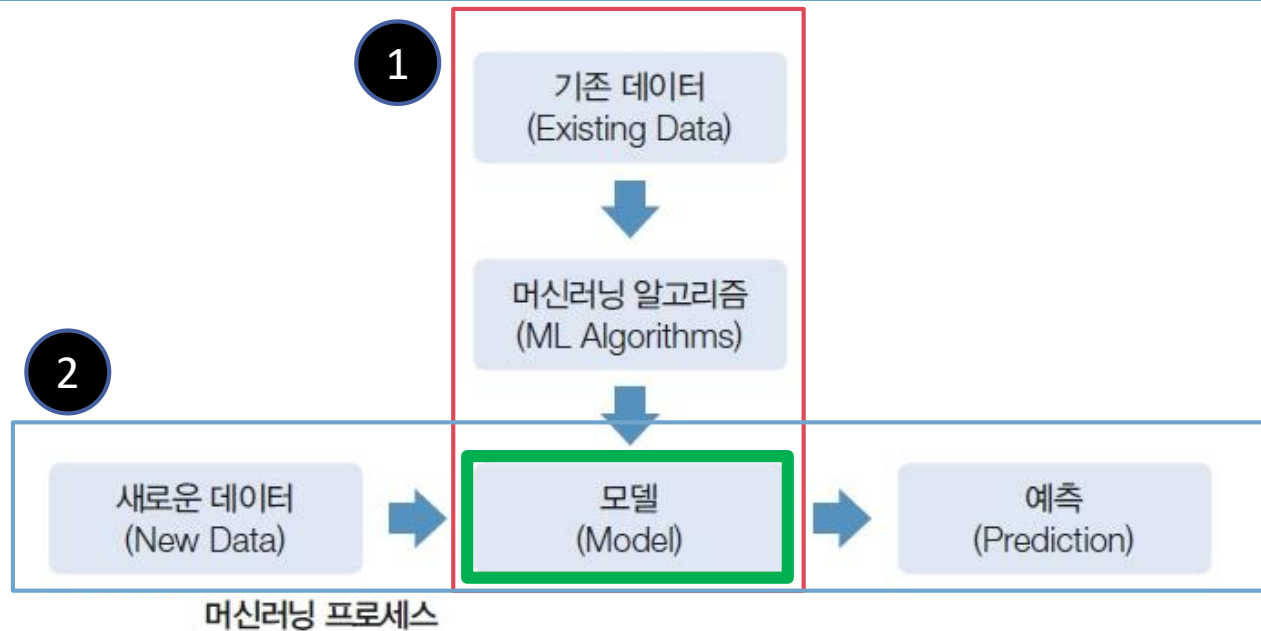
01 머신러닝이란?



02

**머신러닝의
학습 프로세스와 종류**

02 머신러닝의 학습 프로세스와 종류



[하나 더 알기] 모델과 알고리즘

- '모델'은 '수식'이나 '통계 분포', '알고리즘'은 모델을 산출하기 위해 규정화된 과정(훈련과정 = 학습)
- 보통 하나의 모델은 다양한 알고리즘으로 표현할 수 있다
- 때때로 '알고리즘'은 하나의 '수식'으로 표현 가능하다

02 머신러닝의 학습 프로세스와 종류

2. 머신러닝의 종류

- **지도학습**(supervised learning) : 문제와 답을 함께 학습
- **비지도학습**(unsupervised learning) : 조력자의 도움 없이 컴퓨터 스스로 학습. 컴퓨터가 훈련 데이터를 이용하여 데이터들 간의 규칙성을 찾아냄
- **실제 답**(ground truth) y 의 존재 여부에 따라 구분

전통적인 관점에서 머신러닝의 종류

머신러닝 대분류	머신러닝 종류	설명
지도학습	회귀	연속형 값인 y 의 특징을 찾아 데이터 x 를 사용하여 y 값을 예측하는 기법
	분류	이산형 값인 y 의 특징을 찾아 데이터 x 를 사용하여 y 값을 예측하는 기법
비지도학습	군집	y 값이 주어지지 않고, 데이터의 특징이 유사한 값들의 모임을 군집으로 표현하는 기법

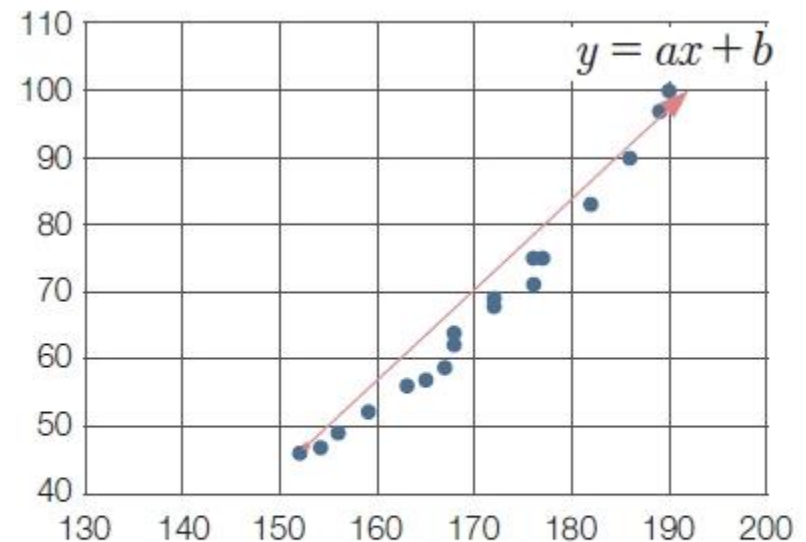
02 머신러닝의 학습 프로세스와 종류

2.1 회귀

- 회귀(regression) : 독립변수 x 와 종속변수 y 의 관계를 **함수식**으로 설명
- **추세선**을 표현하는 수학적 모델을 만드는 기법

키	몸무게	키	몸무게
152	46	172	69
154	47	172	68
156	49	176	71
159	58	176	75
163	56	177	75
165	57	182	83
167	59	186	90
168	64	189	97
168	62	190	100

(a) 키와 몸무게 데이터



(b) 회귀 표현

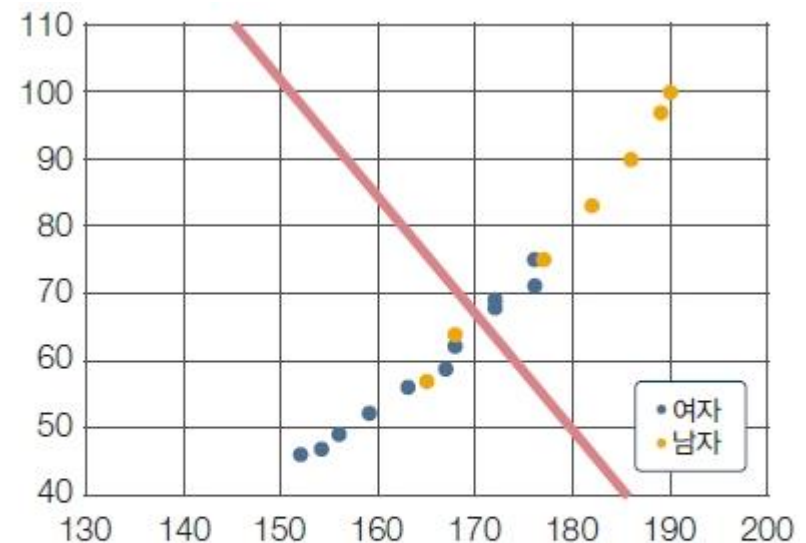
02 머신러닝의 학습 프로세스와 종류

2.2 분류

- 분류(classification) : 데이터를 어떤 기준(패턴)에 따라 나눔
- 이진분류(binary classification) : 2개의 값 중 1개를 분류
- 다중분류(multi-class classification) : 3개 이상 분류 실행

성별	키	몸무게	성별	키	몸무게
여자	152	46	남자	172	69
여자	154	47	여자	172	68
여자	156	49	여자	176	71
여자	159	58	여자	176	75
여자	163	56	남자	177	75
남자	165	57	남자	182	83
남자	167	59	남자	186	90
남자	168	64	남자	189	97
여자	168	62	남자	190	100

(a) 성별, 키, 몸무게 데이터



(b) 분류 표현

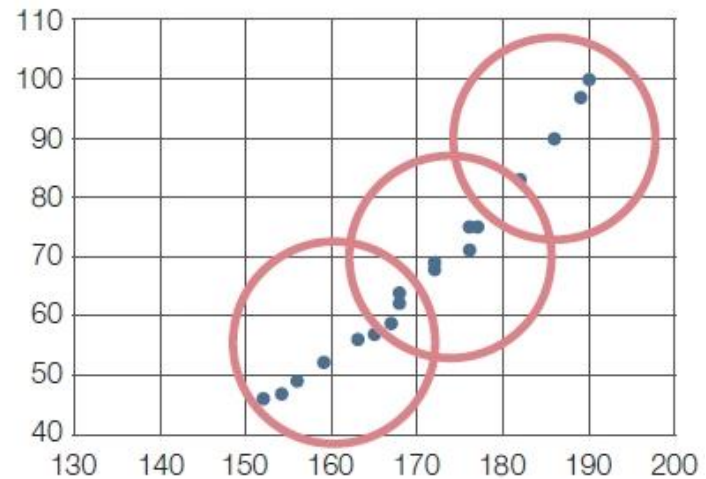
02 머신러닝의 학습 프로세스와 종류

2.3 군집

- 군집(clustering) : 기존에 모여 있던 데이터에 대해 따로 분류 기준을 주지 않고 모델이 스스로 분류 기준을 찾아 집단을 모으는 기법
- 비슷한 수준의 농구팀 3개 만들기

키	몸무게	키	몸무게
152	46	172	69
154	47	172	68
156	49	176	71
159	58	176	75
163	56	177	75
165	57	182	83
167	59	186	90
168	64	189	97
168	62	190	100

(a) 키와 몸무게 데이터



(b) 군집 표현

■ 군집 예제 : 키, 몸무게 데이터

03 머신러닝 환경 구축하기

2. 파이썬 머신러닝 환경 구축하기

파이썬 머신러닝의 도구들

도구 종류	설명	사용 도구
파이썬 인터프리터	파이썬 코드를 해석하고 실행시키기 위한 실행 프로그램	아나콘다(Anaconda)
코드 편집기	파이썬 코드를 수정할 때 사용하는 프로그램	주피터(Jupyter), VS코드(VSCode)
통계 분석 및 전처리 도구	데이터를 로드하고 전처리하기 위한 도구들	넘파이(NumPy), 판다스(Pandas), 사이파이(SciPy)
시각화 도구	데이터의 상태를 파악하기 위해 시각화를 지원하는 도구	맷플롯립(Matplotlib), 시본(Seaborn), 플롯리(Plotly)
머신러닝 프레임워크	실제 머신러닝 모델을 생성하고 데이터에 적용할 수 있도록 도와주는 도구	사이킷런(Scikit-learn)

03 머신러닝 환경 구축하기

01 실행된 터미널 환경에서 다음과 같은 명령어를 입력

```
(base) C:\...>conda create -n gj_env_02 python=3.9
```

02 해당 가상환경을 수행하는 명령어를 실행하면, 가상환경의 명칭이 (base)에서 gj_env_02로 변경됨

```
(base) C:\...>conda activate gj_env_02  
(edu_env) C:\...>
```

03 머신러닝 환경 구축하기

3.3 각 라이브러리 설치하기

- 머신러닝에 필요한 모듈을 설치
- 각 라이브러리들은 서로 의존성을 가지고 있기 때문에 필요에 따라 서로 맞는 버전을 설치

```
(gj_env_02) C:\...>pip install ...
```