

기초 이론부터 실무 실습까지
머신 러닝 익히기

Part 02. 머신 러닝 기초와 배경

정 정 민

Chapter 03. 머신 러닝 기초 개념

1. 머신 러닝이 뭔데?
2. 명시적 프로그램과 머신 러닝
3. 머신 러닝의 활용 사례
4. 머신 러닝 프로젝트의 흐름

머신 러닝이 뭔데?

머신 러닝이란..!

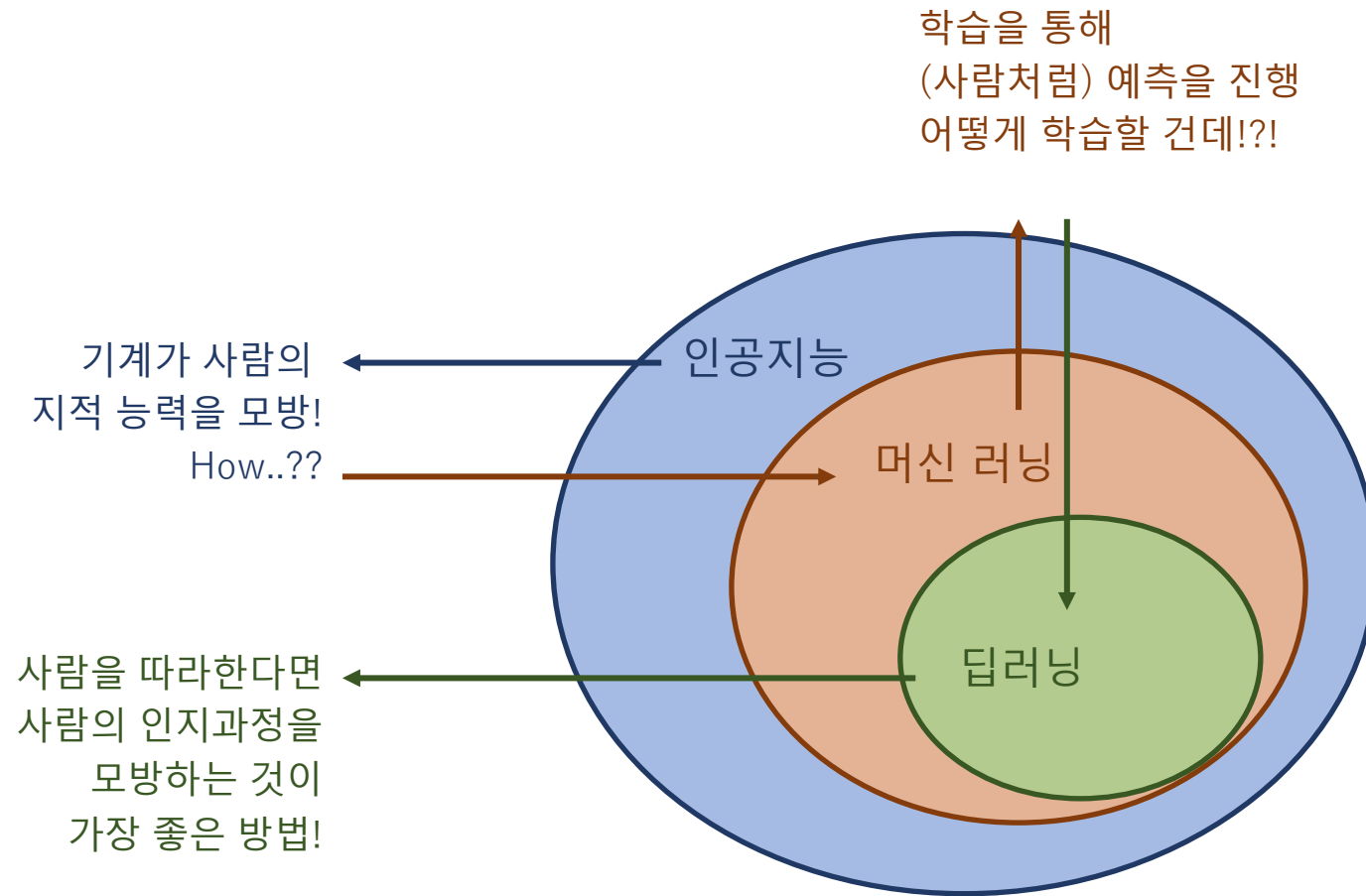
- 데이터에서 지식을 추출하는 작업
- 즉, 머신 스스로가 데이터를 바탕으로 그 안에 있는 특징과 패턴을 찾아냄
 - 패턴을 찾아내는 일 : 학습
- 특징과 패턴을 바탕으로 새로운 데이터에 대한 추론을 진행
- 과거 이런 방식으로 머신 러닝을 정의

“머신 러닝은 명시적인 프로그래밍 없이 컴퓨터가 학습하는 능력을 갖추게 하는 연구 분야”

Atrhur Samuel (1959)

*“어떤 작업 T 에 대한 컴퓨터 프로그램의 성능을 P 로 측정했을 때 경험 E 로 인해 성능이 향상됐다면,
이 컴퓨터 프로그램은 작업 T 와 성능 측정 P 에 대해 경험 E 로 학습한 것이다”*

Tom Mitchell (1997)



명시적 프로그램과 머신 러닝

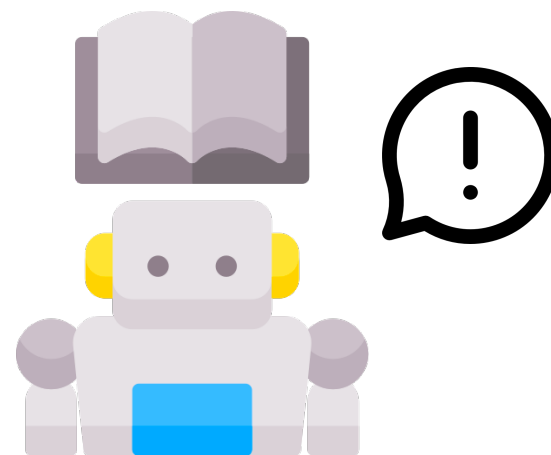
명시적 프로그램 : 규칙 기반 전문가 시스템 (Rule-Based Expert System)

- 머신 러닝 이전의 문제 해결 방법론
- 문제를 해결하기 위한 **규칙(rule)**을 **수동으로 사전에 정의**
 - 규칙 : 하드 코딩 된 if-else 명령어 집합
- 대표적으로 스팸 메일을 분류하는 과정이 이에 해당
 - 특정 단어 (Sale, Win, Free 등)의 유무로 스팸을 분류
- 장점
 - 처리 과정을 사람이 이해하기 쉬움
 - 작은 데이터에서 효과적
- 단점
 - 특정 규칙은 한 분야나 작업에 국한
 - 변경에 대응이 어려움
 - 규칙 설계 시 해당 분야의 전문가가 필요



머신 러닝

- 명시적 프로그램의 한계를 극복할 수 있는 기법
- 데이터 내부에서 자주 발생하는 특징과 패턴을 감지
 - 그러한 특징과 패턴을 feature라고 함
- 문제를 해결하기 위한 판단 기준을 시스템 스스로 찾아냄
- 구체적으로 명시할 수 없는 규칙을 찾을 수 없다면 머신 스스로가 그것을 찾아내게 하는 방법이 좋은 솔루션
- 장점
 - 예상치 못한 상관 관계를 파악하는데 탁월함
 - 특정 도메인에서 전문가가 필수로 필요하지 않음
- 단점
 - 머신이 패턴을 파악할 수 있도록 다양한 데이터가 필요
 - 결과 분석 과정에서 사람이 이해할 수 없는 포인트가 존재할 수 있음



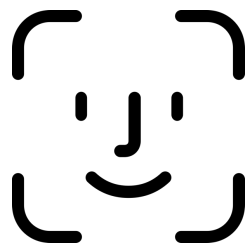
머신 러닝의 활용 사례

머신 러닝과 딥러닝의 다양한 사례

- 머신 러닝은 다양한 어플리케이션과 연구 분야에서 활용
- 최근 딥러닝의 발전으로 매우 다양한 분야에서 다양한 목적으로 사용



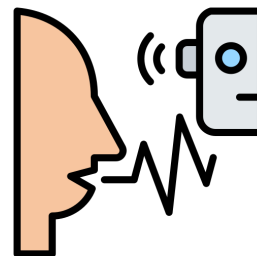
추천
영화, 음식, 쇼핑 등



FaceID
[출처](#)



의료 영상 처리
종양 진단, 병변 예측 등



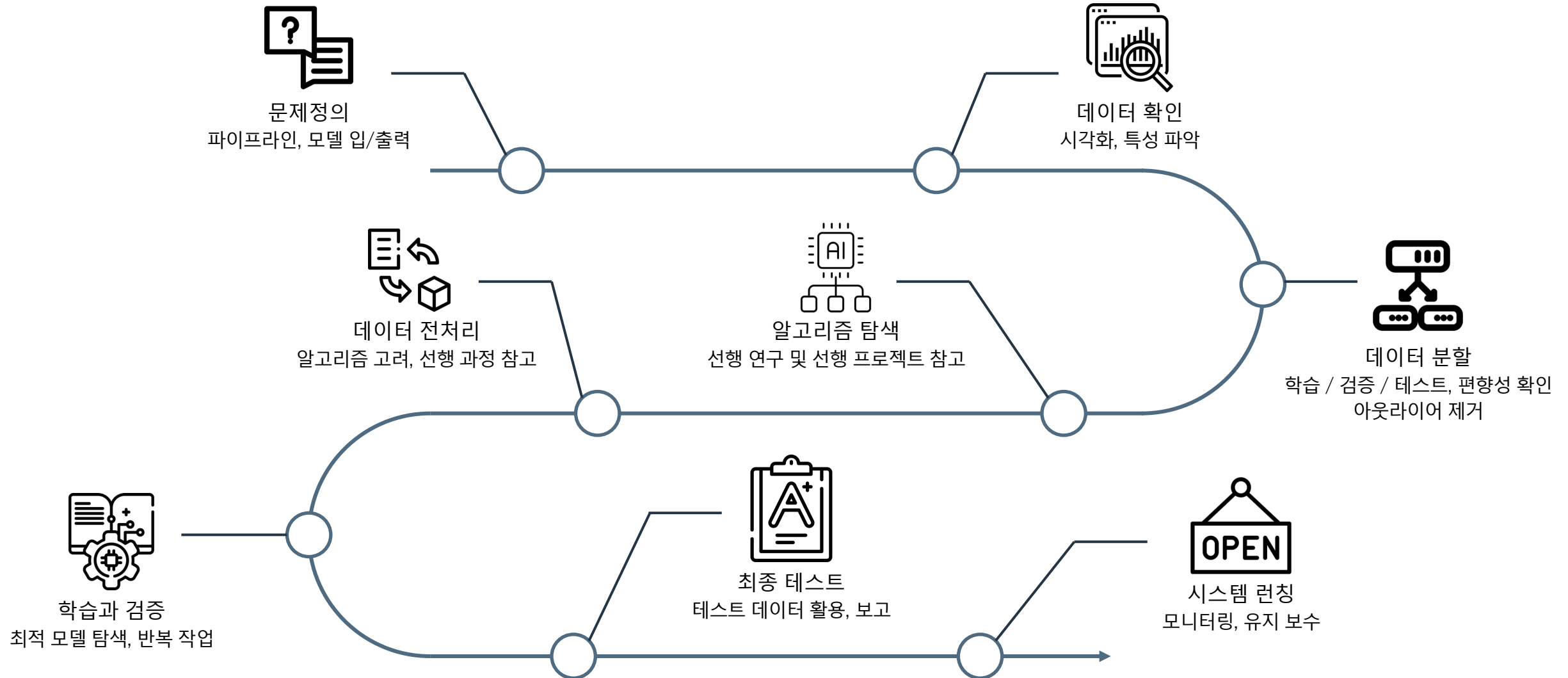
음성 처리
TTS, STT 등



금융 데이터 예측
수익 예측, 이상 거래 감지 등

머신 러닝 프로젝트의 흐름

머신 러닝 프로젝트 로드맵



E.O.D