

# 빅데이터 분석 프로그래밍 Application

Introduction

서중원 2021-Fall

# 강사소개



서중원

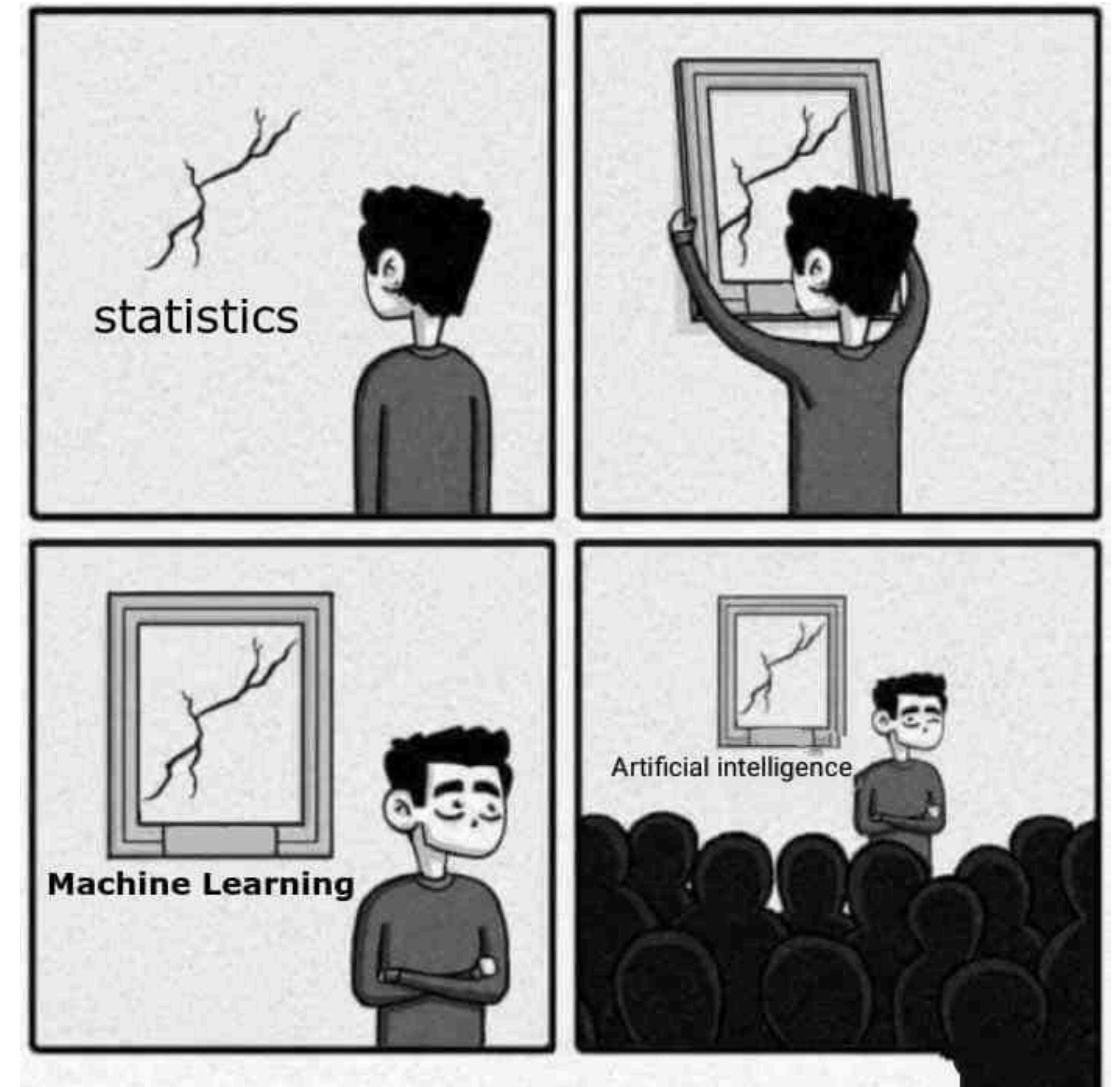
<https://github.com/thejungwon>

- 경력
  - 스타트업 BeBridge CTO
  - 연세대 정보대학원 강사 (2019-)
  - IT교육 및 컨설팅 업체 CodeVinci 대표
  - 스타트업 코스폴 Lead Developer
- 학력
  - University of Stavanger, 컴퓨터과학 석사
  - 연세대학교 컴퓨터과학 학사
- 기타
  - AWS (Seoul), Equinor (Norway) 인턴
  - NASA-Yonsei 큐브위성 프로젝트 개발자

# 머신러닝 이전의 의사 결정

머신러닝 기술 이전에는 어떻게 의사 결정을 하였을까?

- 규칙 기반 (Rule-based)
  - 1분 안에 10% 이상 떨어지면 매수!
  - 1분 안에 10% 이상 오르면 매도!
- 통계 기반
  - 단순 평균, 중위값, 최대 최소, 표준 편차, 분산을 이용
  - 상관관계 분석, 확률 분포
- 느낌
  - 내가 해봐서 아는데

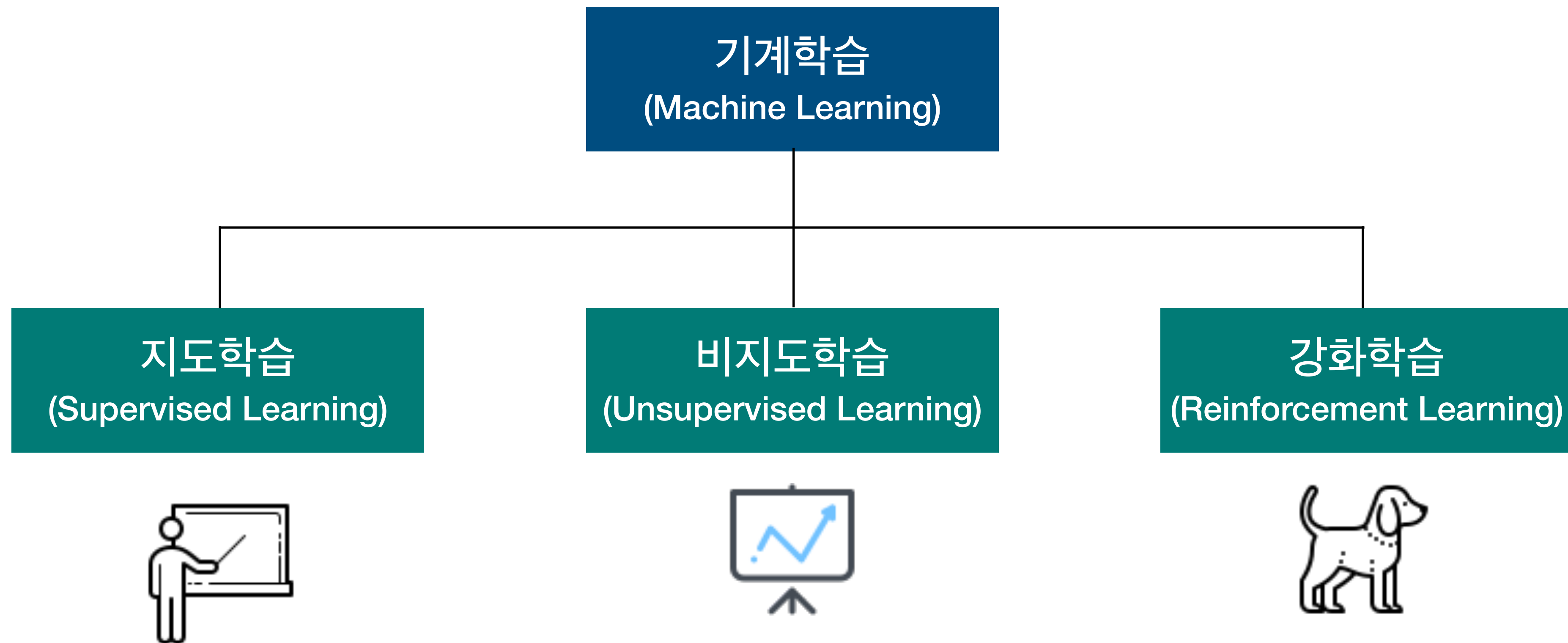


# 머신러닝

## 기계학습이란?

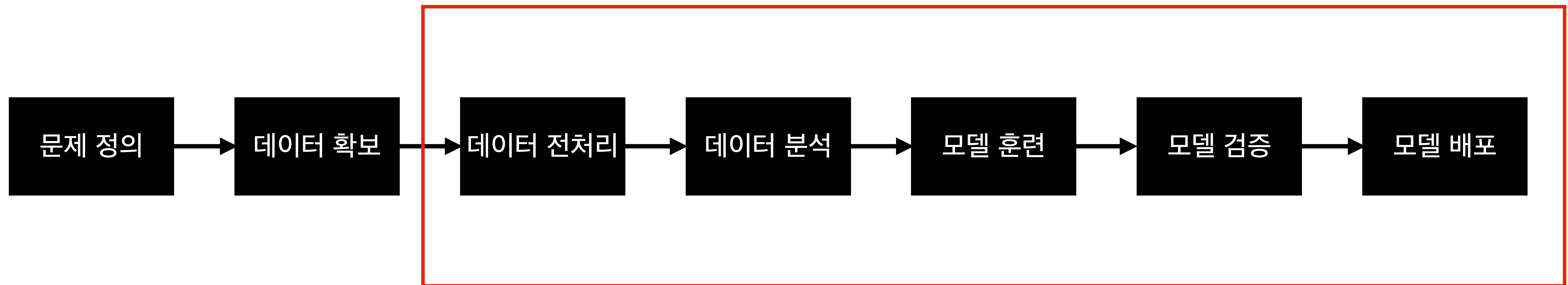
- 경험을 통해 자동으로 개선하는 컴퓨터 알고리즘 연구
  - 경험 = 데이터
  - 개선 = 더 많은 데이터, 더 좋은 성능
  - 알고리즘 = 모델 알고리즘

# 기계학습의 종류



# 머신러닝 사이클

구체적인 문제를 정의하고, 최적의 해답을 제공

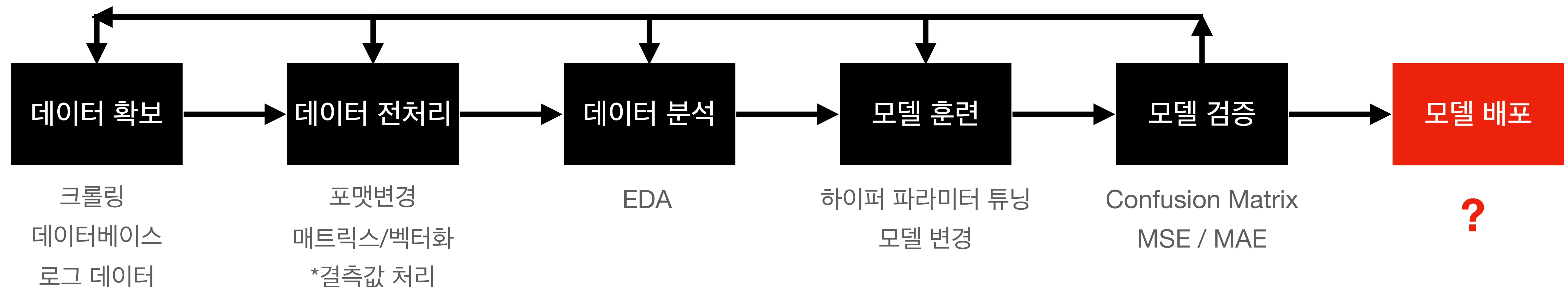




# 머신러닝 사이클

모델 검증 후 스토리는? 주피터 노트북을 깃허브에 올리면 되나?

데이터 확보, 전처리, 분석, 모델 훈련, 검증에 대해서는 많은 방법론이나 과정이 있지만, 실제 이렇게 생성한 모델은 어떻게 배포해야 할까요?



# 배울 내용들

- 지식적인 측면
  - 머신러닝 알고리즘 이해
  - 목적에 맞는 머신러닝 문제 정의
- 실무적인 측면
  - 전체적인 머신러닝 프로세스 개발 능력



# 성적평가

- 1시간 이론강의, 2시간 실습강의
- 매 수업 퀴즈
  - 성적평가 미반영 (점검용)
- 참여도 (40%)
  - 출석 + Q&A 답변
- 과제 (60%)
  - 고민 중

**E.O.D**