Week1

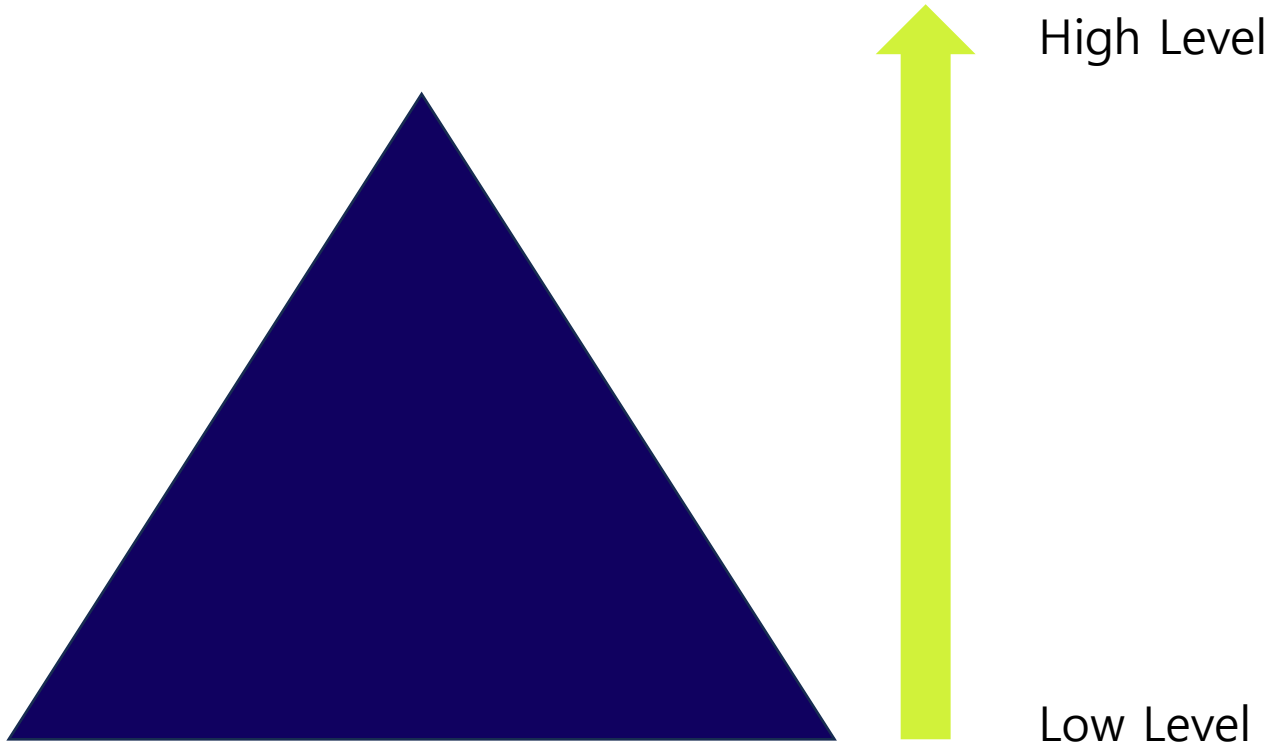
DataBiz팀 조용걸

Python

Python

https://www.codestates.com/blog/content/%EA%B0%9D%EC%B2%B4-%EC%A7%80%ED%96%A5-%ED%94%84%EB%A1%9C%EA%B7%B8%EB%9E%98%EB%B0%8D-%ED%8A%B9%EC%A7%95

Low Coupling, High Cohesion



AI/ML

AI/ML

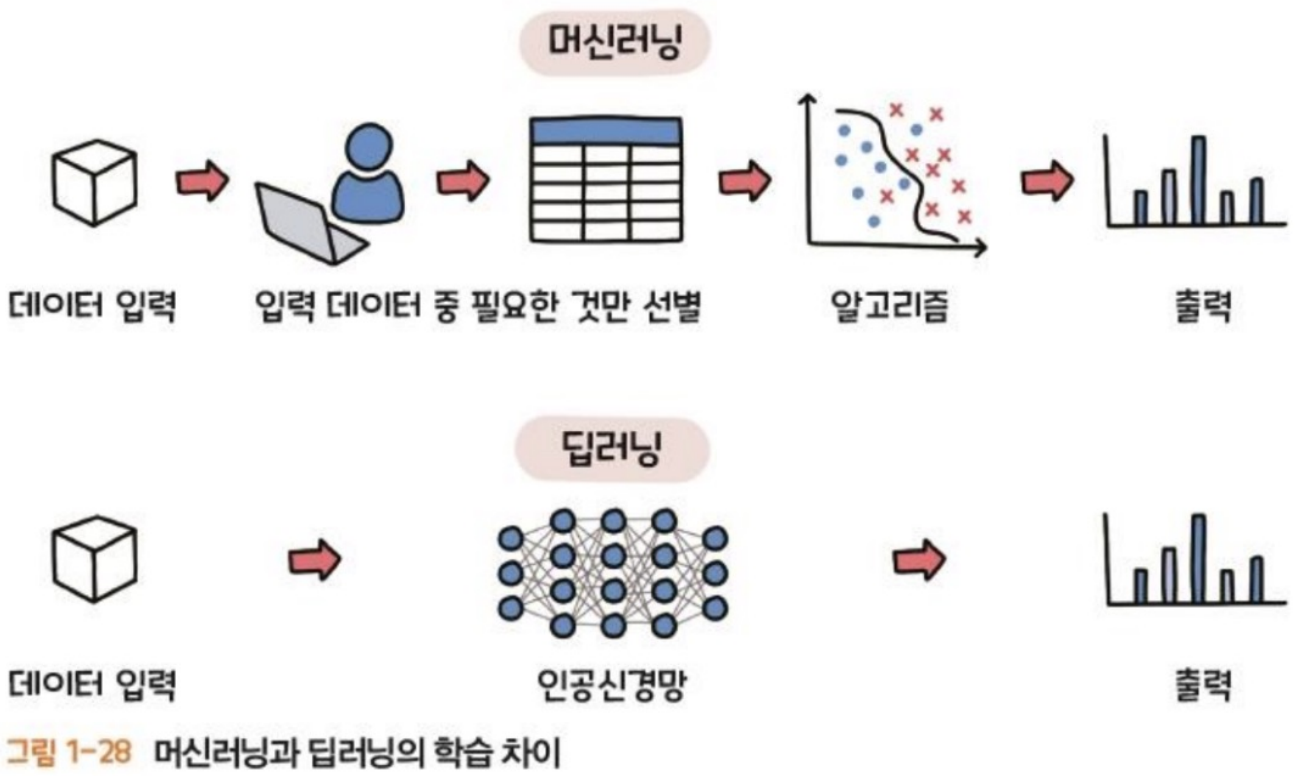


그림 1-28 머신러닝과 딥러닝의 학습 차이

https://wooono.tistory.com/206

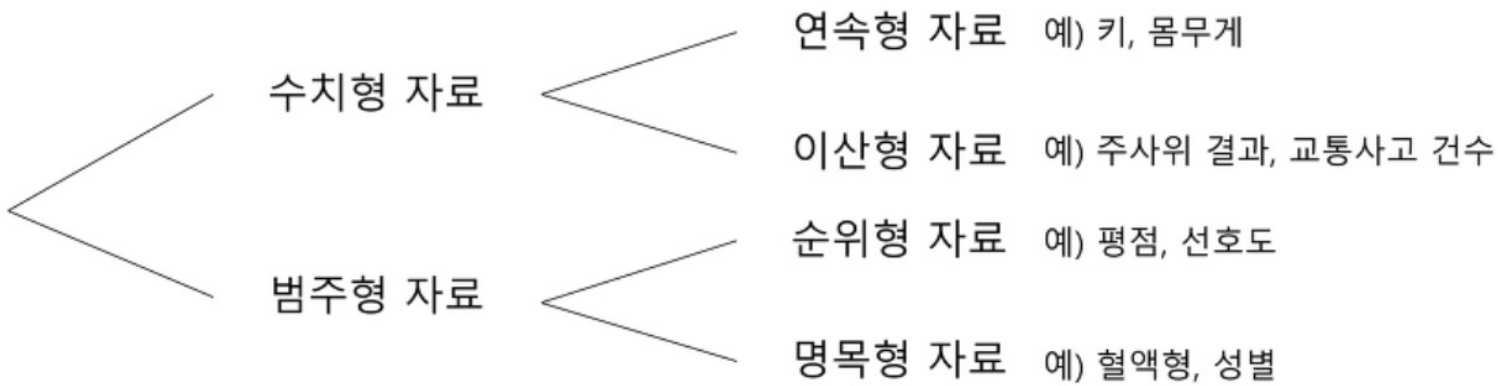
AI/ML

Features (x-Variable, Independent variable)
Target (label)

변수 Type

Categorical data
- 성별, 직업

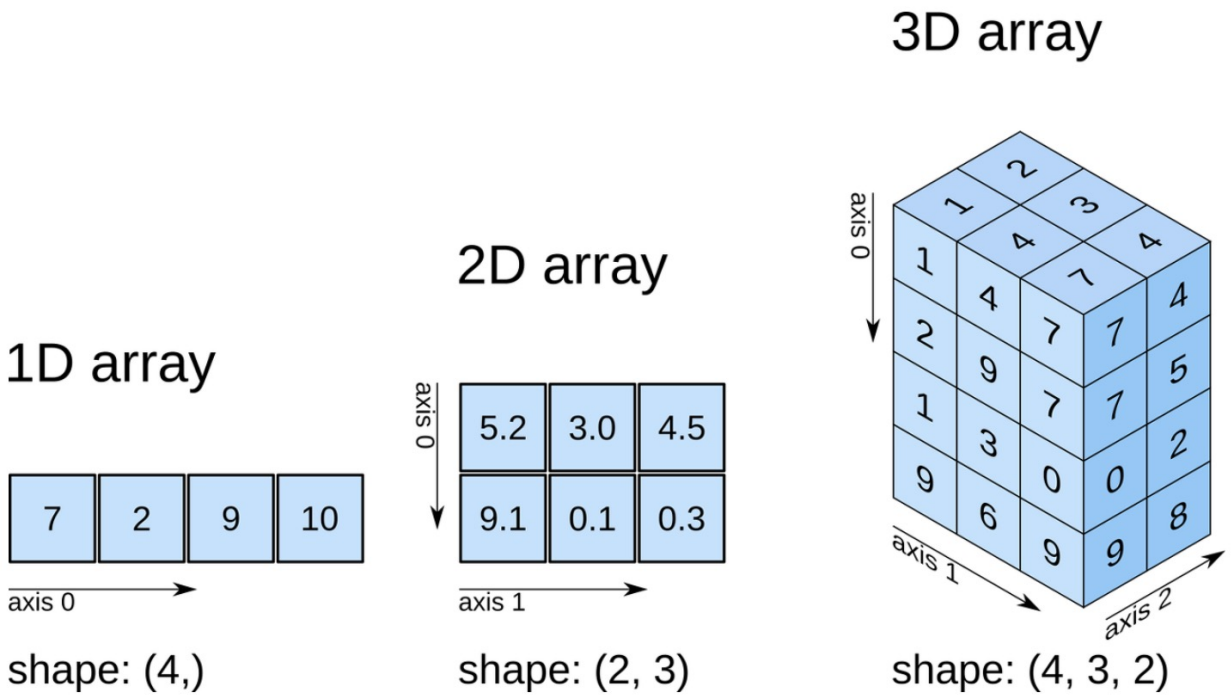
Numerical data
- 정수, 실수 등



https://leedakyeong.tistory.com/entry/%EC%88%98%EC%B9%98%ED%98%95-%EC%9E%90%EB%A3%8Cnumerical-data%EC%99%80-%EB%B2%94%EC%A3%BC%ED%98%95-%EC%9E%90%EB%A3%8Ccategorical-data

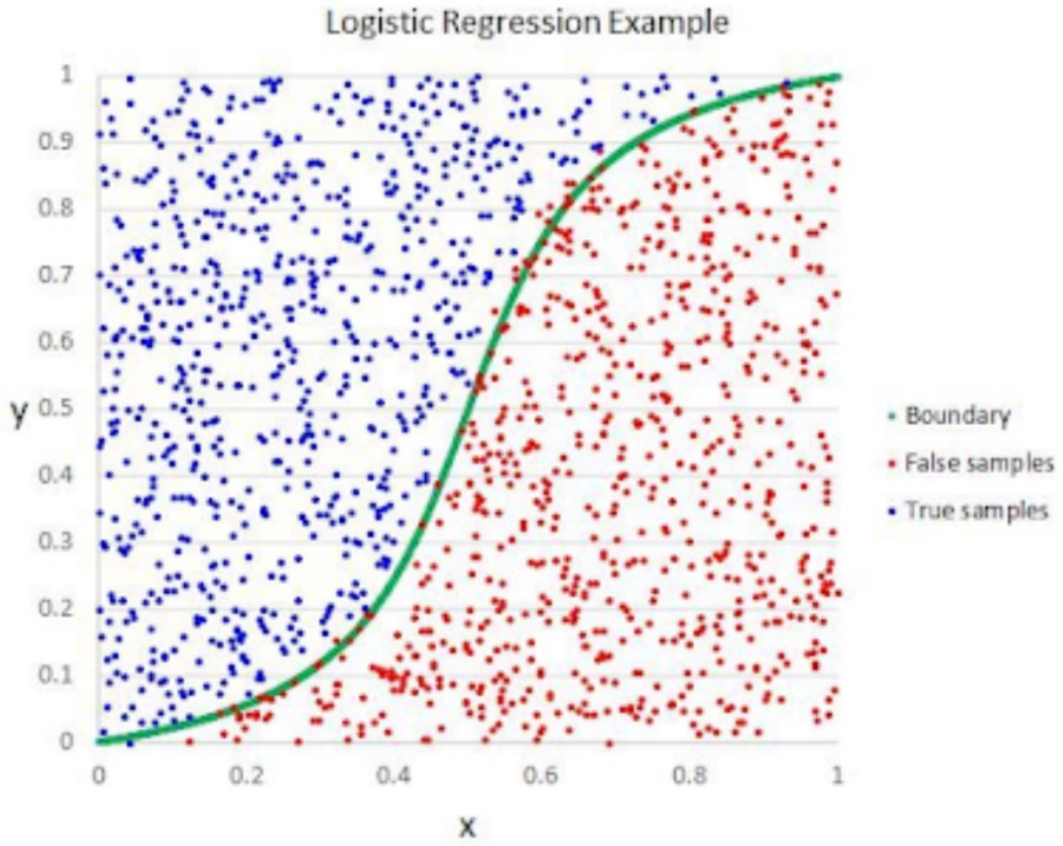
AI/ML

Vector
Matrix
Tabular Dataset



https://www.oreilly.com/library/view/elegant-scipy/9781491922927/ch01.html

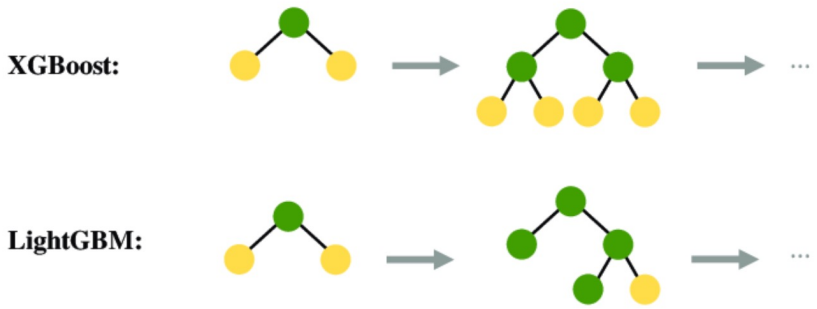
AI/ML


Logistic Regression Example
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
y
0
0.2
0.4
0.6
0.8
1
x
Boundary
False samples
True samples

AI/ML

ML

Explainable vs Interpretable
가용가능한 Resource
데이터 및 업무 활용 특징



scikits learn
machine learning in Python

Decision Tree(LGBM, XGBoost)
Support Vector Machine
Logistic Regression

AI

Pytorch
Tensorflow

공부 방향

과거

- 나만의 Dataset으로 모델 개발 및 훈련
- 적은 Resource
- PoC 작업 진행시 직접 데모 개발 및 시연
- Pre-trained model의 재학습
- 모델의 높은 자유도
- 약간 높은 진입장벽
- 가능성 확인 및 아카데믹한 분위기

현재

- ChatGPT ???
- 매우 큰 Resource 필요
- API 호출 및 개발 (Langchain, Streamlit 등)
- Fine tuning
- 모델의 아키텍쳐가 어느정도 수렴
- 낮아진 진입장벽
- Business Application 개발

감사합니다