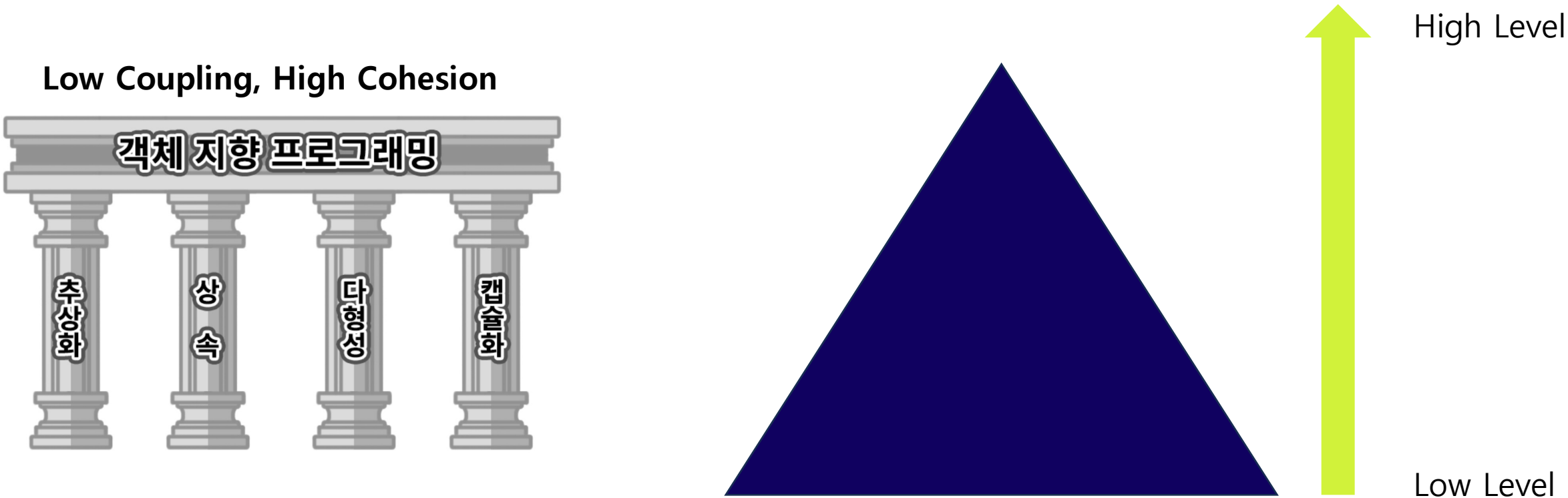


Week1

DataBiz팀 조용걸

Python



AI/ML

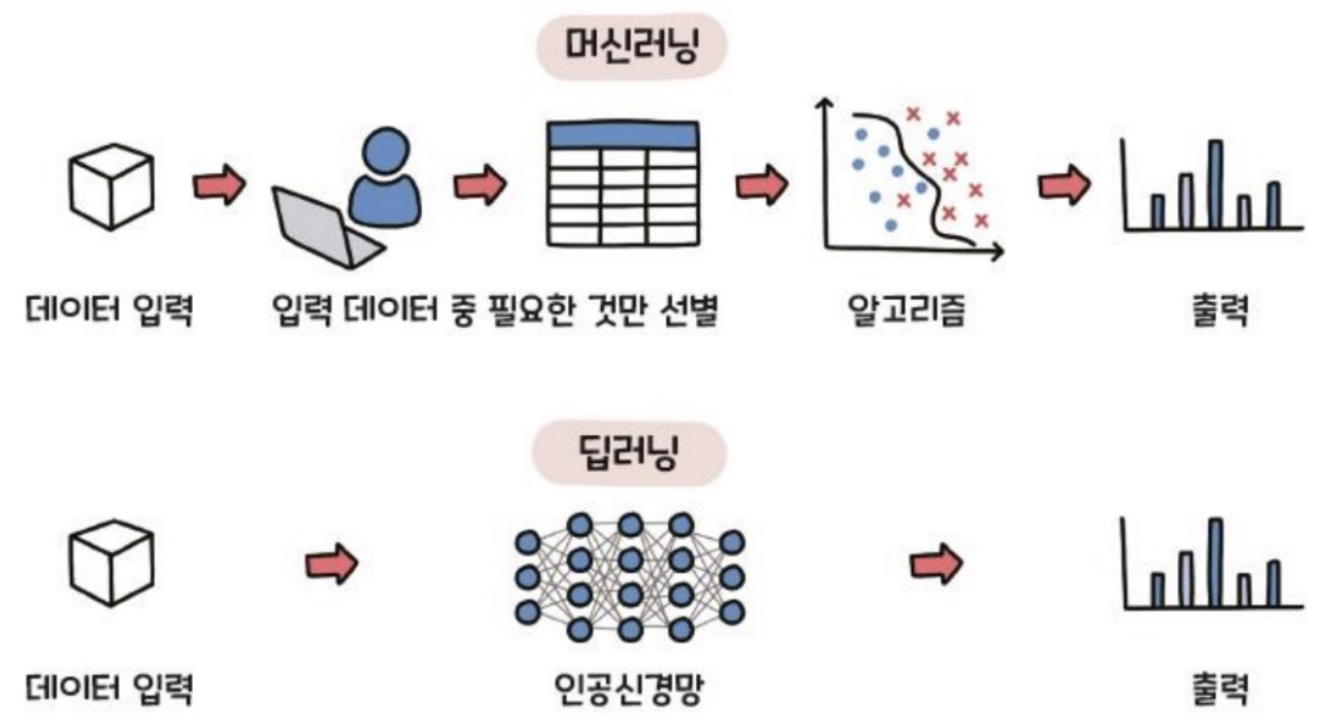


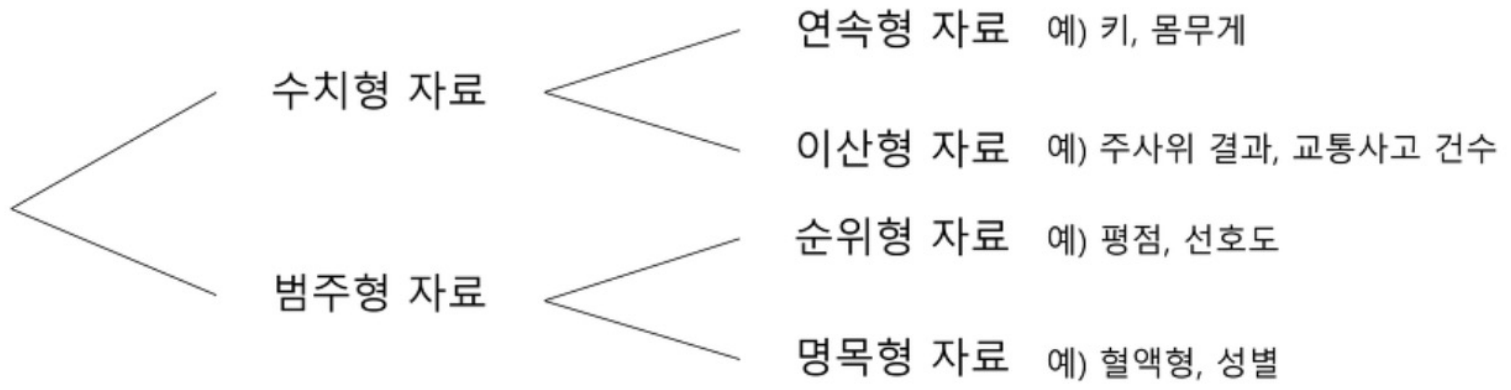
그림 1-28 머신러닝과 딥러닝의 학습 차이

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

변수 Type

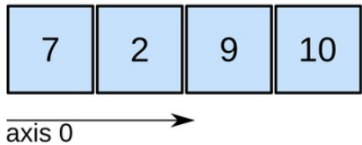
Categorical data
- 성별, 직업

Numerical data
- 정수, 실수 등



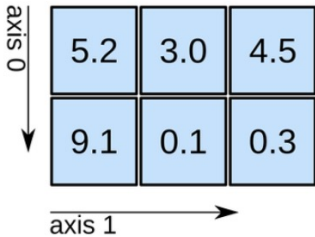
Vector
Matrix
Tabular Dataset

1D array



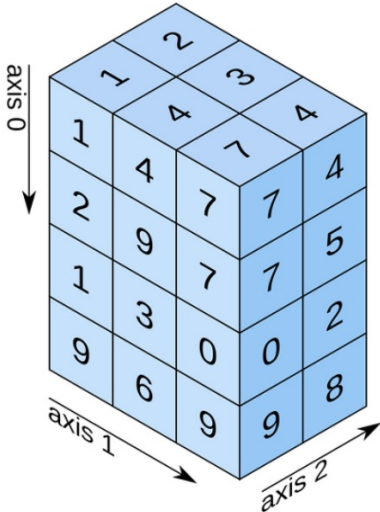
shape: (4,)

2D array

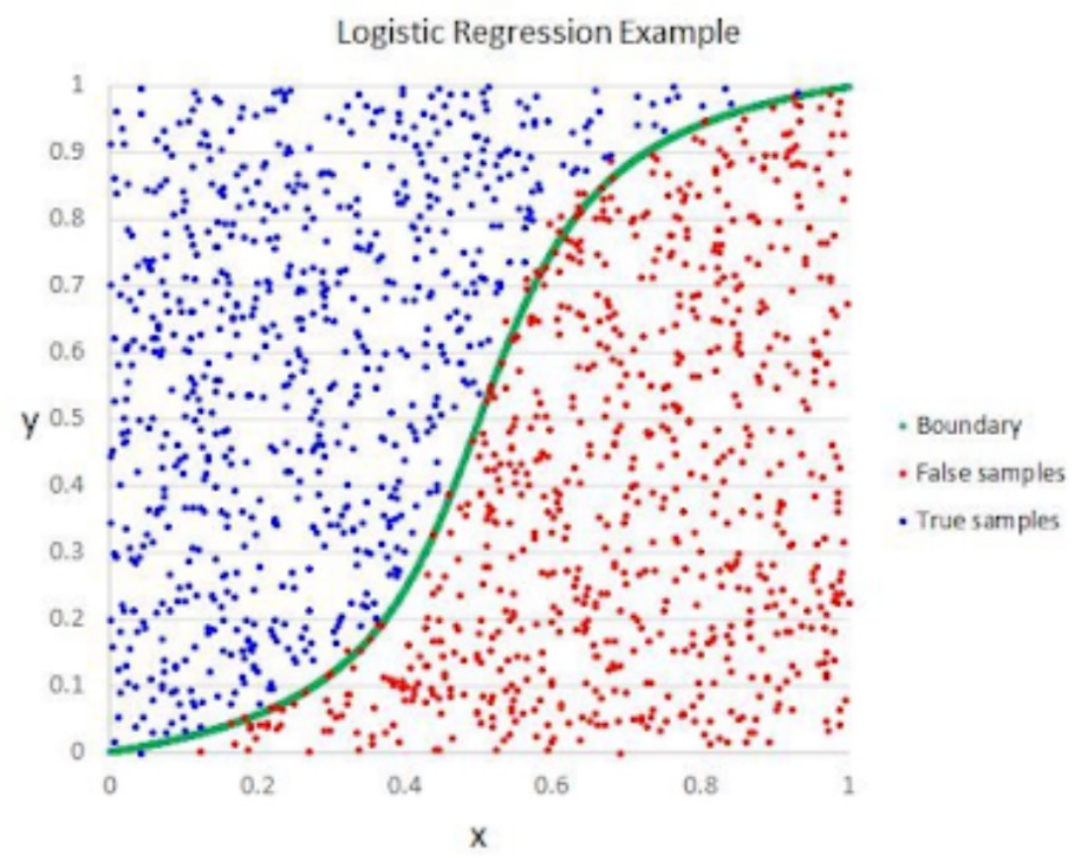


shape: (2, 3)

3D array



shape: (4, 3, 2)

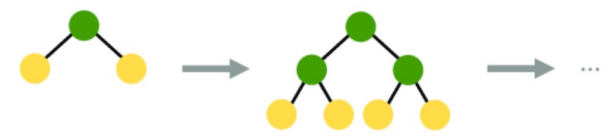


ML

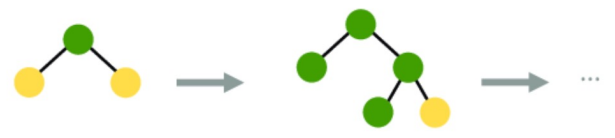
Explainable vs Interpretable
가용가능한 Resource
데이터 및 업무 활용 특징

AI

XGBoost:



LightGBM:



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

Pytorch
Tensorflow

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 아카데미한 분위기

현재

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

감사합니다