Data Science

What is data science?

Data science, also known as data-driven science, is an interdisciplinary field of scientific Methods, process, algorithms and systems of extract knowledge or insights from data in various forms, either structured or unstructured, similar to data mining

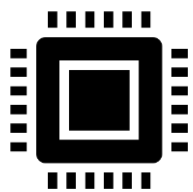
Machine Learning

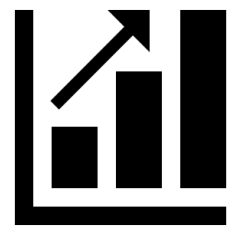
Why??

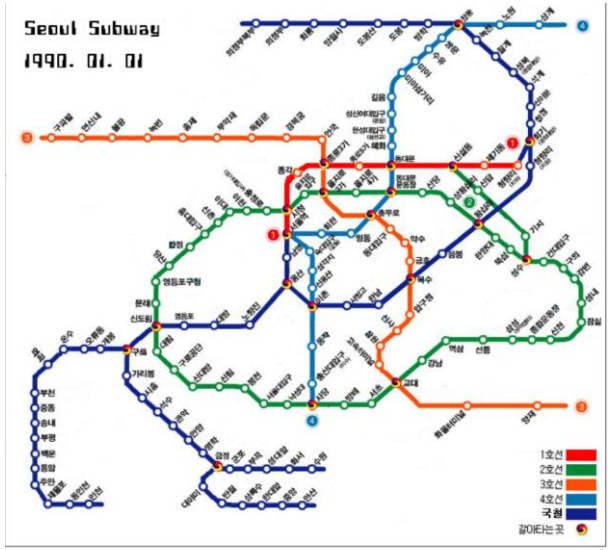
More simple

More Fast

More exactly

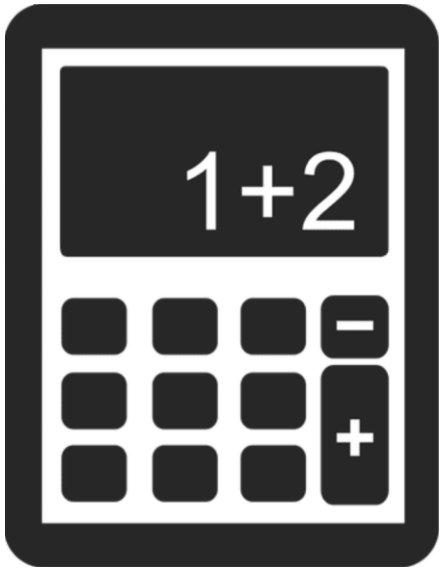




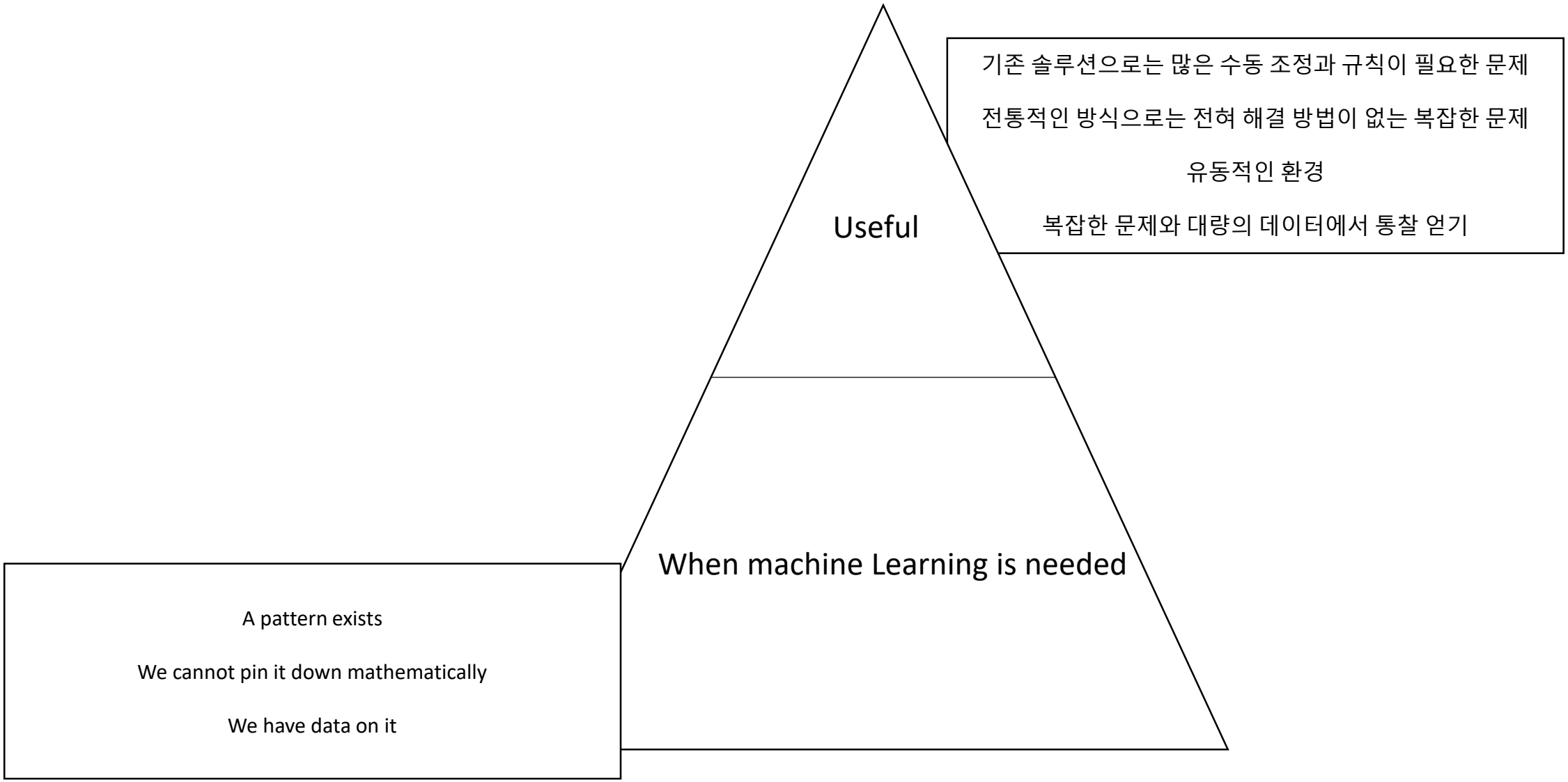


Shortest path problem

When?

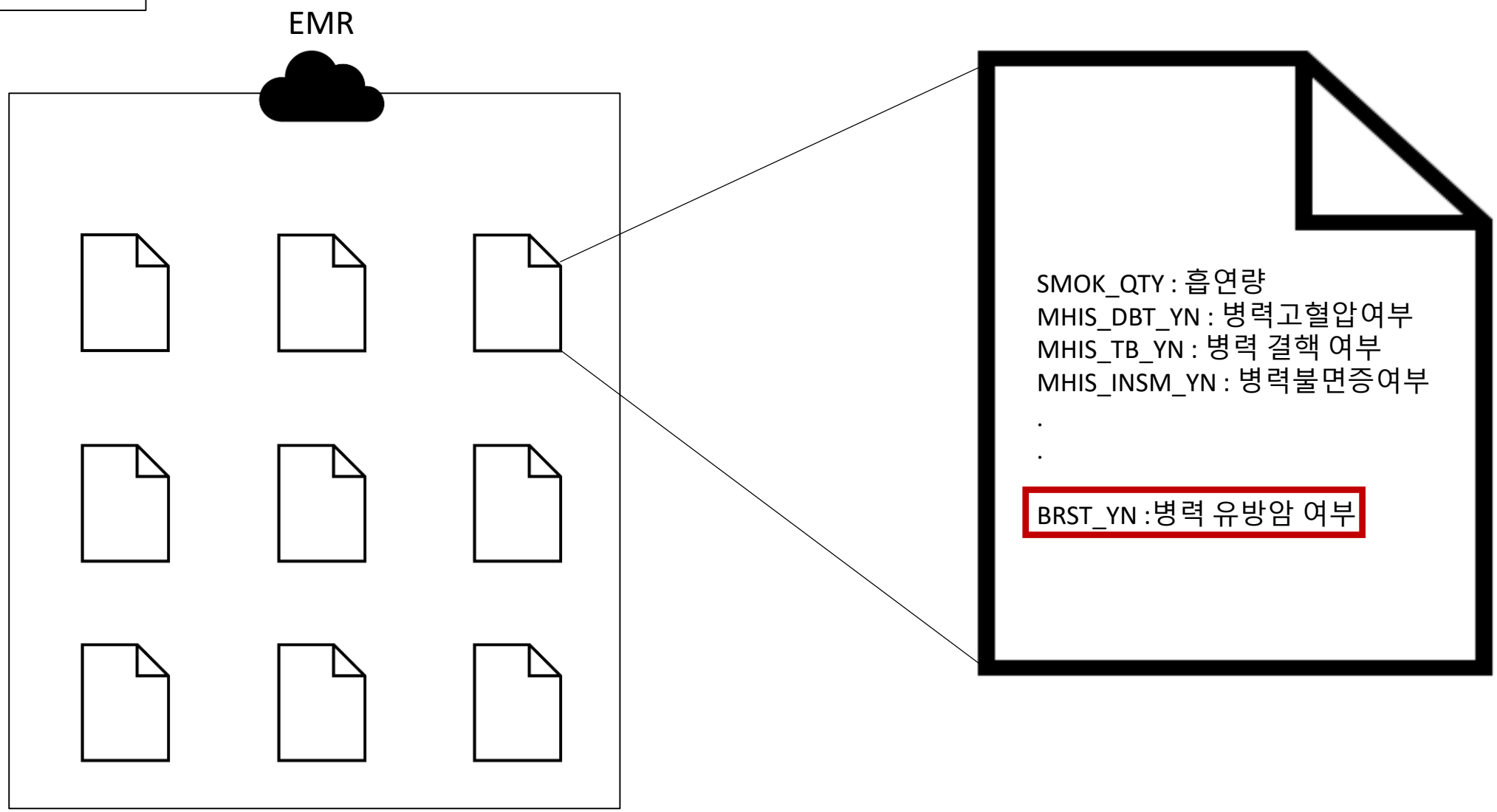


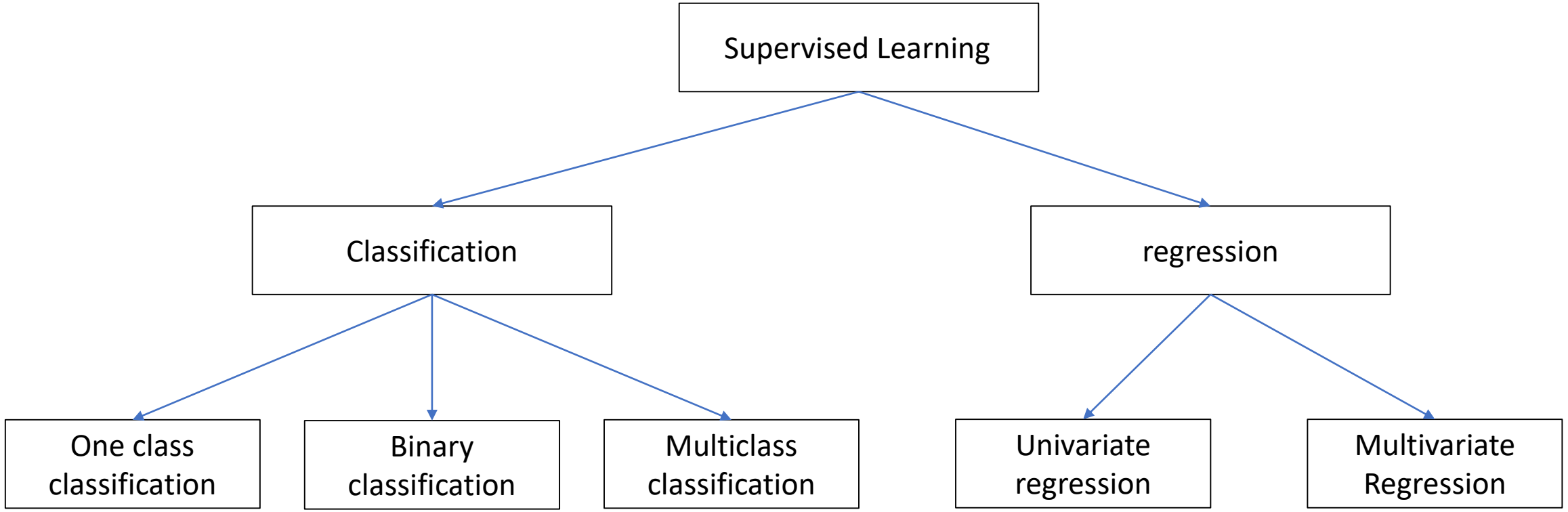
Arithmetic operation problem

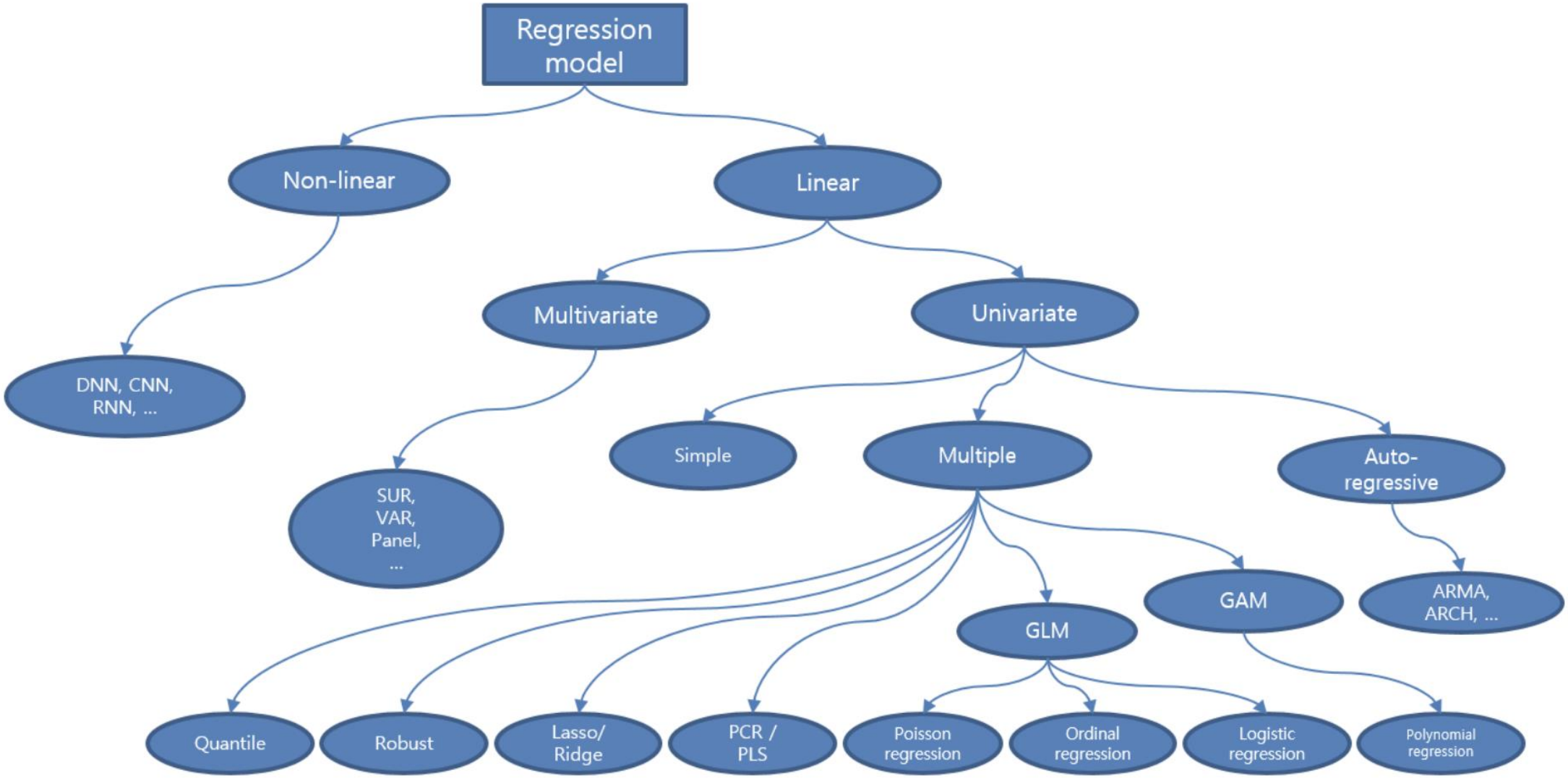

Useful
When machine Learning is needed
기존 솔루션으로는 많은 수동 조정과 규칙이 필요한 문제
전통적인 방식으로는 전혀 해결 방법이 없는 복잡한 문제
유동적인 환경
복잡한 문제와 대량의 데이터에서 통찰 얻기
A pattern exists
We cannot pin it down mathematically
We have data on it

Supervised Learning

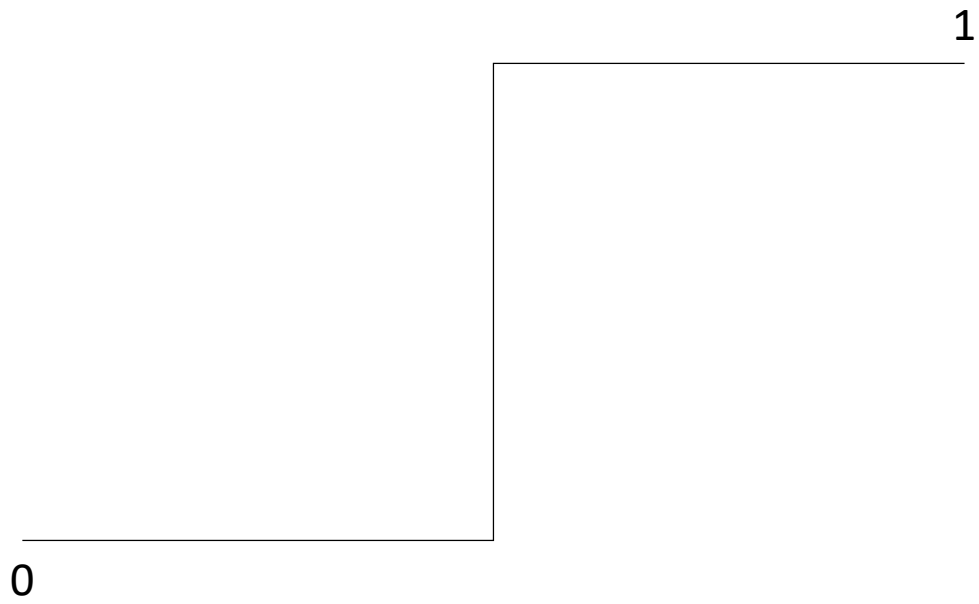
Supervised Learning



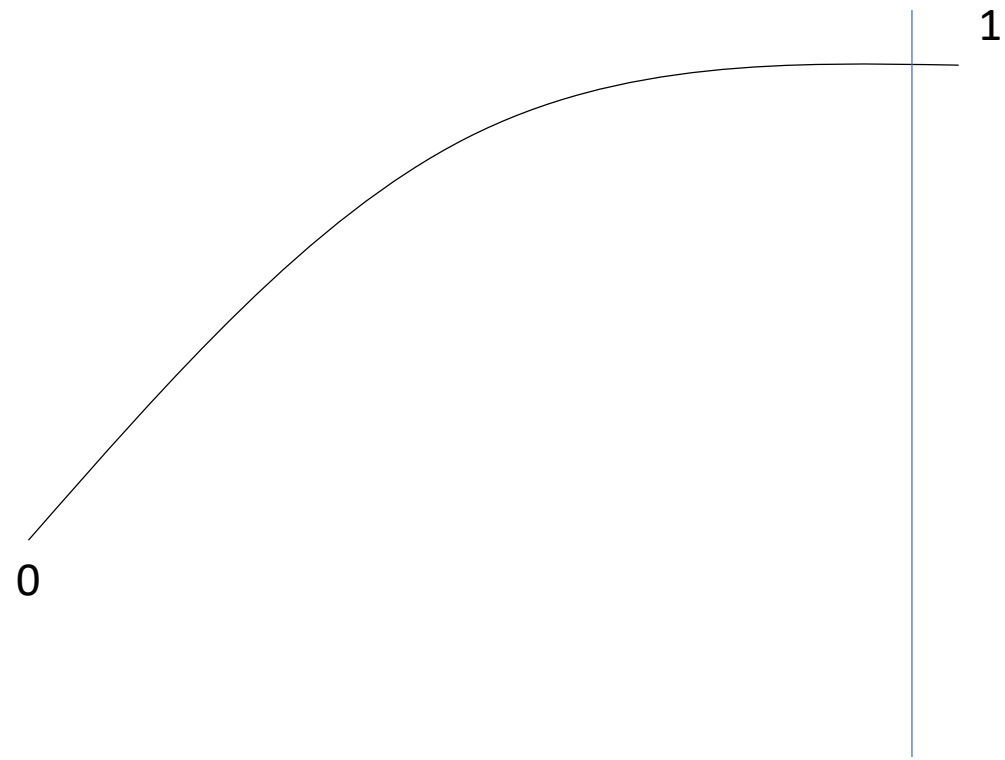

Supervised Learning
Classification
regression
One class classification
Binary classification
Multiclass classification
Univariate regression
Multivariate Regression


Regression model
Non-linear
Linear
DNN, CNN, RNN, ...
Multivariate
Univariate
SUR, VAR, Panel, ...
Simple
Multiple
Auto-regressive
GAM
ARMA, ARCH, ...
GLM
Quantile
Robust
Lasso/ Ridge
PCR / PLS
Poisson regression
Ordinal regression
Logistic regression
Polynomial regression

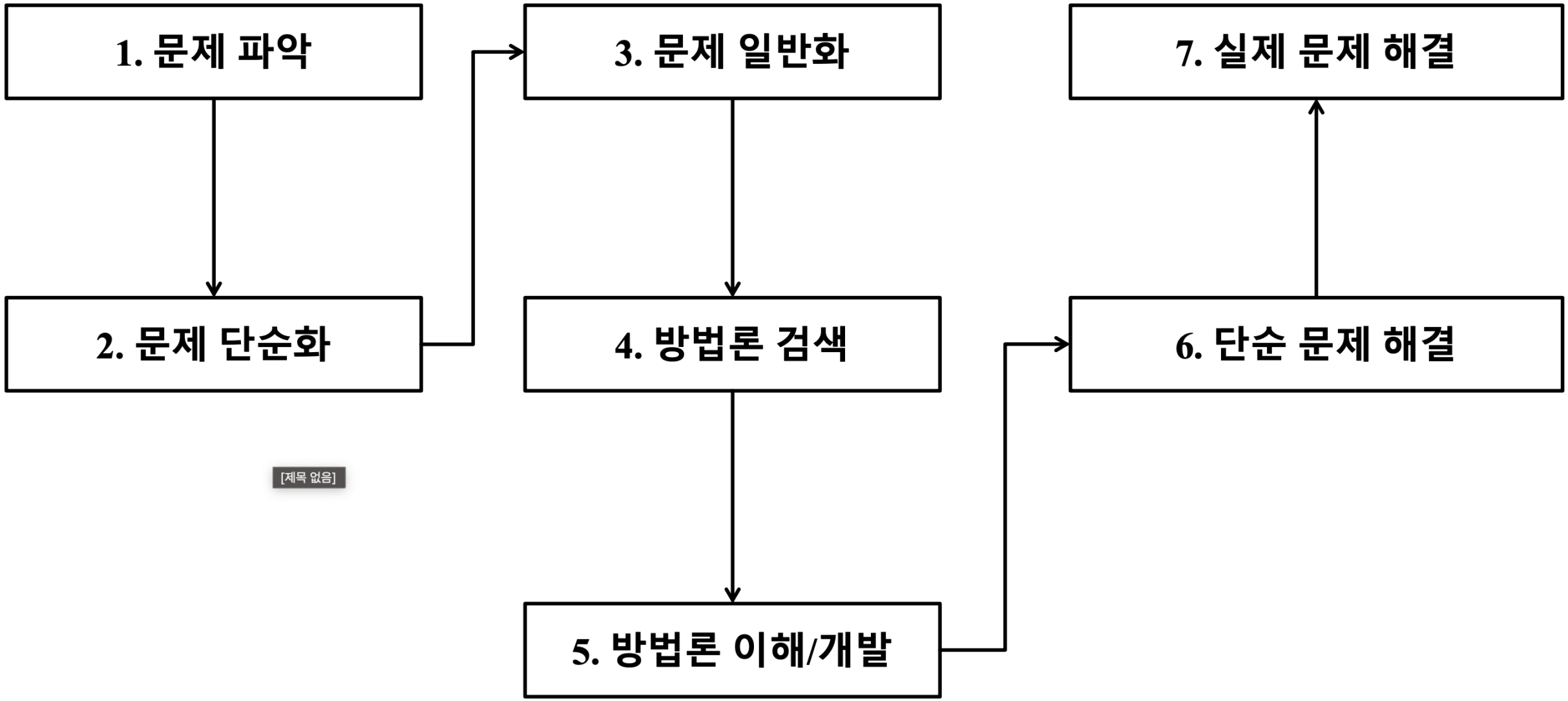
classification



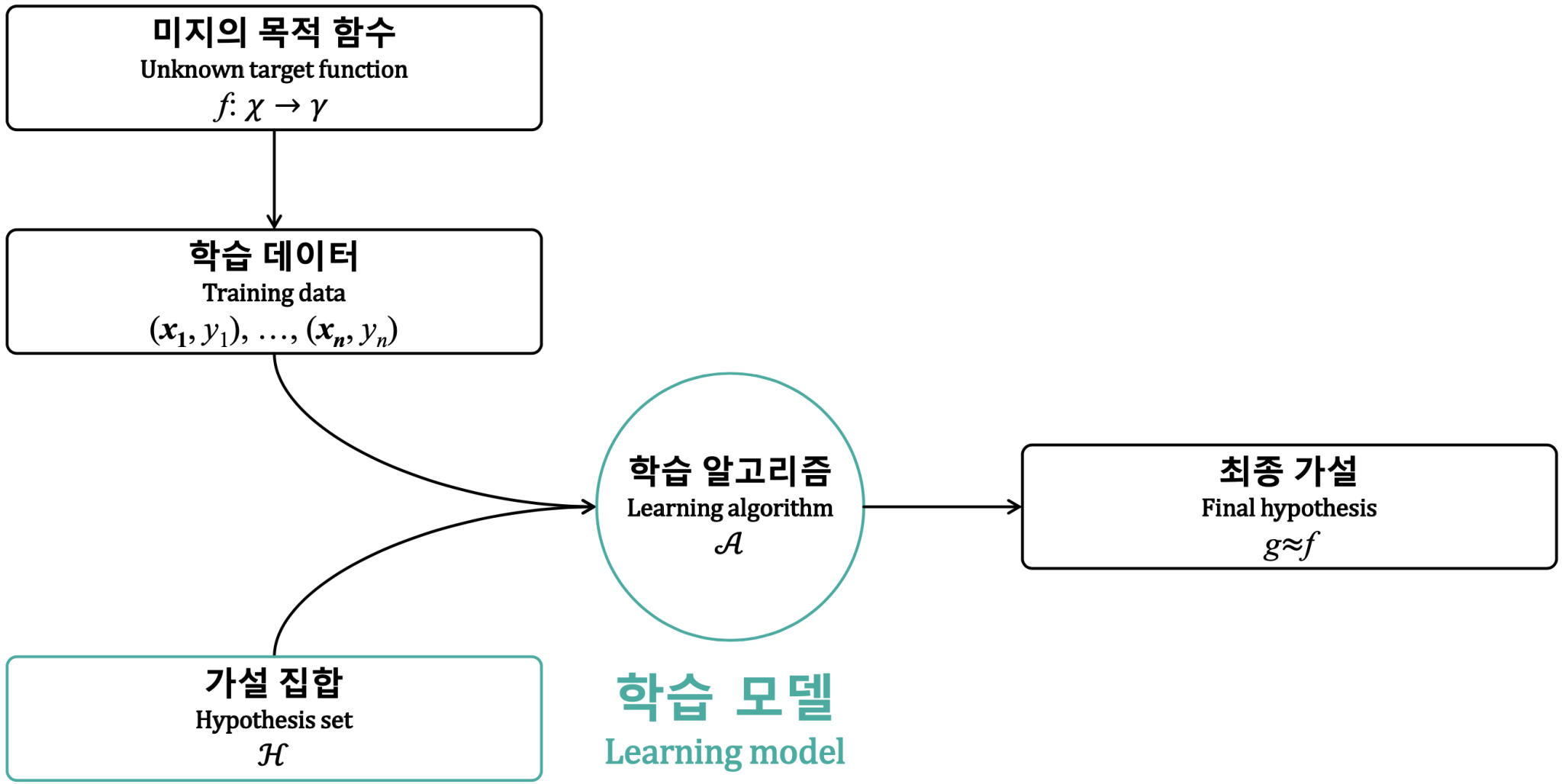
regression



How Data scientist solve the problem?



[제목 없음]


미지의 목적 함수
Unknown target function
$f: \chi \rightarrow \gamma$
학습 데이터
Training data
$(x_1, y_1), \ldots, (x_n, y_n)$
학습 알고리즘
Learning algorithm
$\mathcal{A}$
최종 가설
Final hypothesis
$g \approx f$
가설 집합
Hypothesis set
$\mathcal{H}$
학습 모델
Learning model