

AI교과 교사연구회 연수

부산기계공업고등학교
2021.01



인공지능/머신러닝/딥러닝

인공지능(AI, Artificial Intelligence)

- 인공지능이란

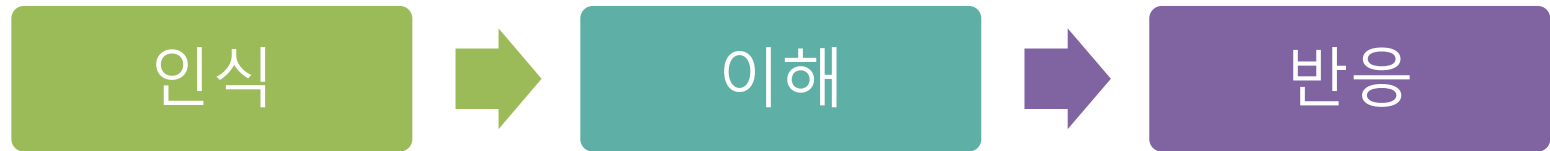
- 사고나 학습 등 인간이 가진 지적 능력을 컴퓨터를 통해 구현하는 기술
- 기계가 경험을 통해 학습하고 새로운 입력 내용에 따라 기존 지식을 조정하며 사람과 같은 방식으로 과제를 수행할 수 있도록 지원하는 기술
- 기계가 사람의 지능을 모방하여 사람과 같이 복잡한 일을 할 수 있도록 하는 다양한 기술

- 인공지능 분류

- 약인공지능 (Weak AI)
 - 유용한 도구로 사용하기 위해 설계된 인공지능
- 강인공지능 (Strong AI)
 - 인간을 완벽하게 모방한 인공지능



(약)인공지능



인공지능 발전

인공지능 (Artificial Intelligence)

사고방식이나 학습 등 인간이 가지는 지적 능력을 컴퓨터를 통해 구현하는 기술



머신러닝 (Machine Learning)

컴퓨터가 스스로 학습하여 인공지능의 성능을 향상시키는 기술 방법



딥러닝 (Deep Learning)

인간의 뉴런과 비슷한 인공신경망 방식으로 정보를 처리



인공지능 트렌트 분석



머신러닝

- 머신 러닝

- 사람과 동물에게는 자연스러운 일, 즉 경험을 통해 학습하는 것을 컴퓨터가 수행할 수 있도록 가르치는 데이터 분석 기법
- 머신 러닝 알고리즘은 미리 결정된 방정식을 모델로 의존하지 않고 계산 방법을 사용하여 데이터에서 직접 정보를 “학습”
- 학습할 수 있는 샘플 수가 증가함에 따라 알고리즘 성능이 향상

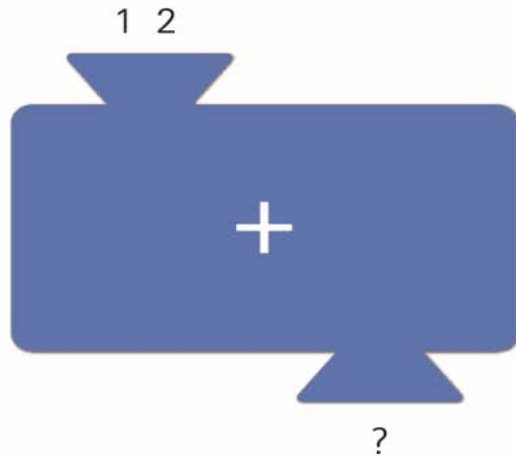
- 머신 러닝 처리에 적합 경우

- 대용량 데이터와 많은 변수가 관련되어 있지만 기존 공식이나 방정식이 없는 복잡한 작업이나 문제
 - 얼굴 인식, 음성 인식의 경우처럼 직접 작성하는 규칙과 방정식이 너무 복잡한 상황
 - 거래 기록에서 사기를 감지하는 경우처럼 작업 규칙이 지속적으로 바뀌는 상황
 - 자동화된 트레이딩, 에너지 수요 예측, 쇼핑 추세 예측의 경우처럼 데이터 특징이 계속 바뀌고 프로그램을 조정해야 하는 상황

머신러닝

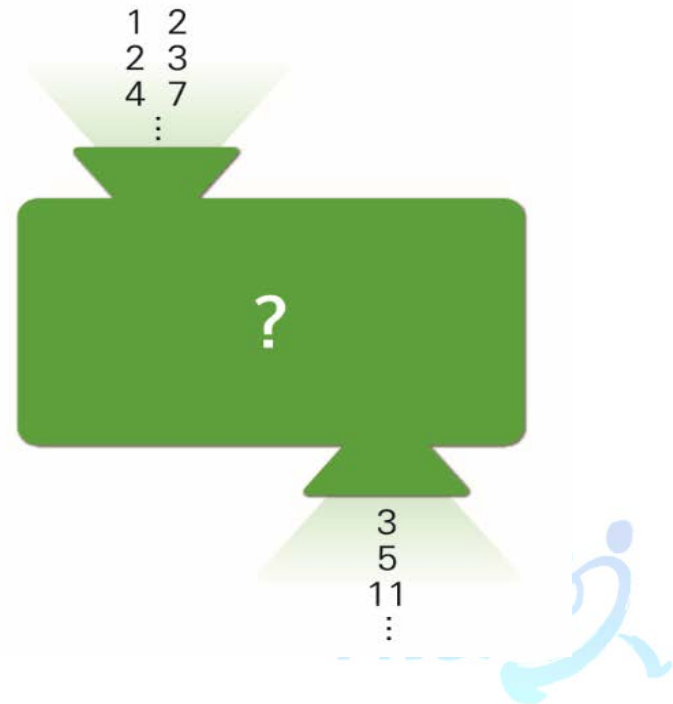
기존 프로그래밍

a+b의 결과값이 나오는 프로그래밍 설계

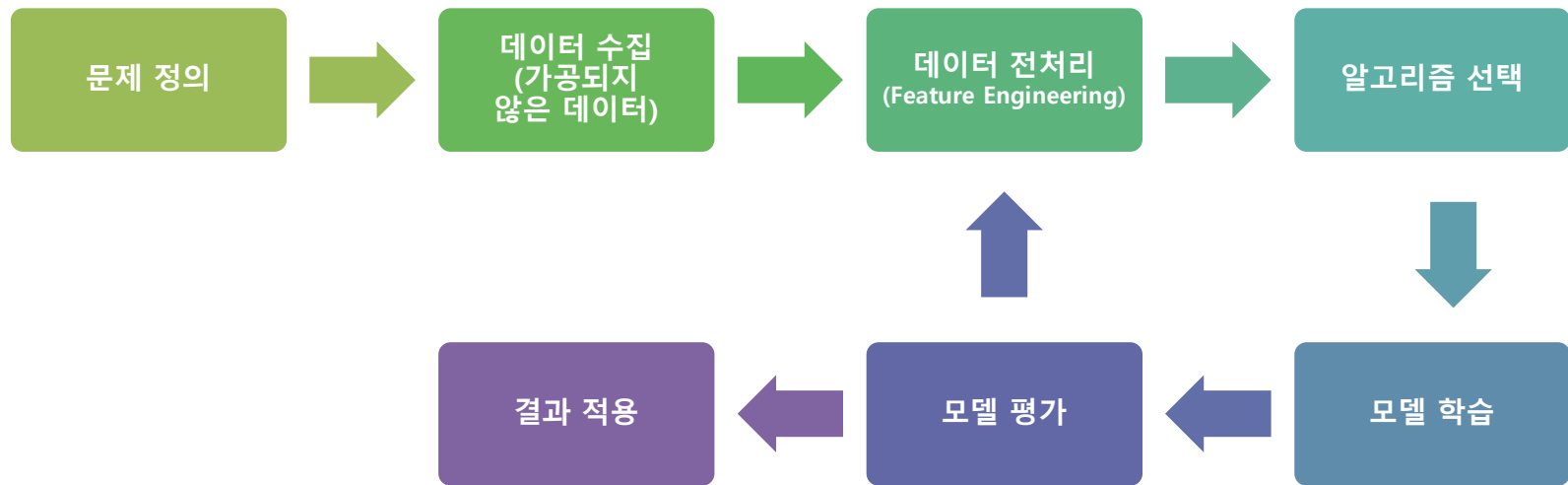


머신러닝

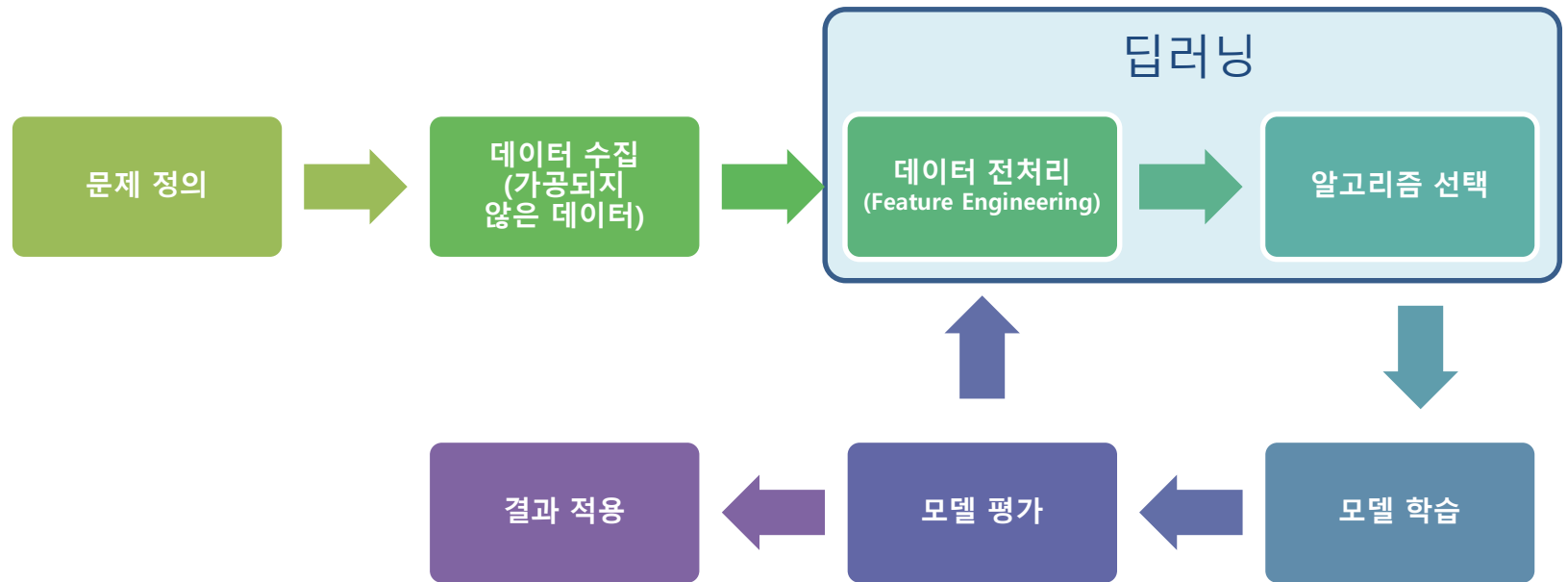
다양한 데이터를 통해 연산을 도출



머신러닝 개발과정



딥러닝 개발과정



머신러닝 분류



머신 러닝 분류

- 지도(supervised) 학습

- 알려진 입력 데이터셋 및 해당 데이터에 대한 알려진 응답(출력)을 사용하고 새 데이터에 대한 응답을 위해 합리적인 예측을 생성하도록 모델을 학습
- 출력에 대한 알려진 데이터가 있는 경우 지도 (supervised) 학습을 사용

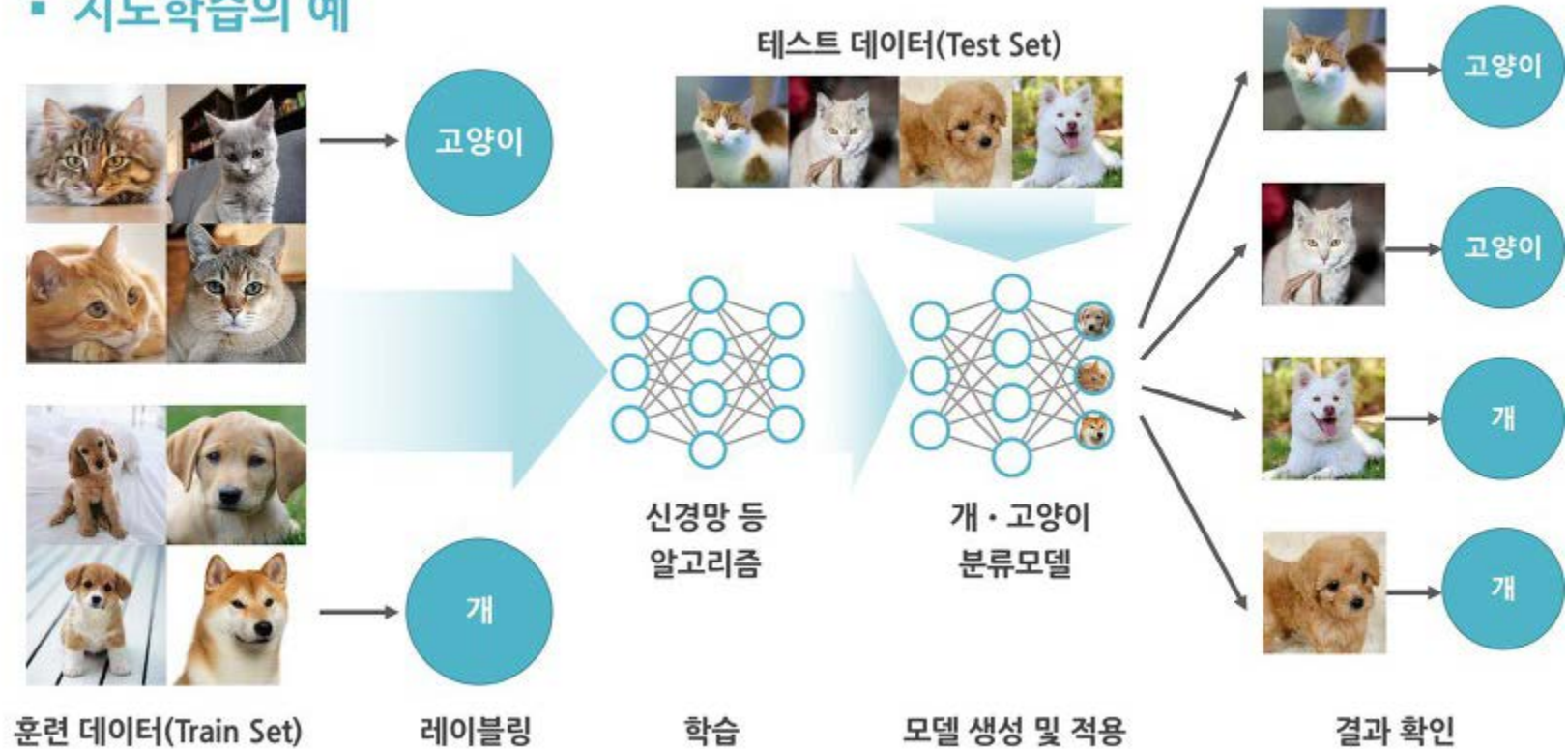
- 비지도(unsupervised) 학습

- 데이터에서 숨겨진 패턴이나 고유 구조체를 찾음
- 패턴이나 구조체는 분류된 응답 없이 입력 데이터로 구성된 데이터셋에서 추론



지도 학습

지도 학습의 예

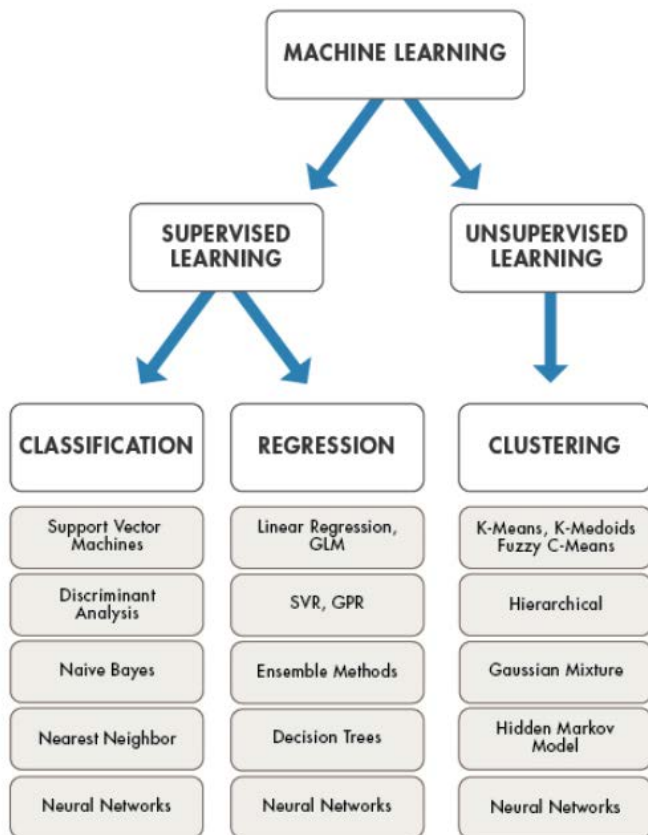


비지도 학습

▪ 비지도 학습의 예



머신 러닝 알고리즘



- **분류 기법**

- 데이터에 태그 지정하거나 범주화하거나 특정 그룹 또는 클래스로 구분할 수 있는 경우 분류를 사용
- 이메일이 진짜 또는 스팸인지 여부, 종양이 악성 또는 양성인지 여부 등의 개별 응답을 예측

- **회귀 기법**

- 데이터 범위로 작업하는 경우 또는 응답의 특성이 실제 숫자인 경우
- 온도 변화 또는 전력 수요 변동 등의 연속 응답을 예측

- **클러스터링**

- 탐색적 데이터 분석을 통해 데이터에서 숨겨진 패턴이나 그룹을 찾는 데 사용
- 휴대폰 기지국을 서울 위치를 기지국에 의존하는 사람들의 클러스터 수를 예측

머신러닝 알고리즘

Supervised Learning 지도학습

- SVM
- Logistic Regression
- Linear Regression
- Decision Tree
- Random Forest
- kNN

Unsupervised Learning 비 지도학습

- K-means
- PCA
- Apriori

Reinforcement Learning 강화학습

- Monte-Carlo
- Q-learning
- Markov Decision

머신러닝 알고리즘

Machine Learning Overview MACHINE LEARNING IN EMOJI Becoming Human.AI

SUPERVISED

human builds model based
on input / output


UNSUPERVISED

human input, machine output
human utilizes if satisfactory

REINFORCEMENT

human input, machine output
human reward/punish, cycle continues

BASIC REGRESSION

 **LINEAR**
`linear_model.LinearRegression()`
Lots of numerical data



 **LOGISTIC**
`linear_model.LogisticRegression()`
Target variable is categorical



CLUSTER ANALYSIS

 **K-MEANS**
`cluster.KMeans()`
Similar datum into groups based
on centroids



 **ANOMALY
DETECTION**
`covariance.EllipticalEnvelope()`
Finding outliers through grouping




CLASSIFICATION

 **NEURAL NET**
`neural_network.MLPClassifier()`
Complex relationships. Prone to overfitting
Basically magic.



 **K-NN**
`neighbors.KNeighborsClassifier()`
Group membership based on proximity



 **DECISION TREE**
`tree.DecisionTreeClassifier()`
If/then/else. Non-contiguous data.
Can also be regression.




 **RANDOM FOREST**
`ensemble.RandomForestClassifier()`
Find best split randomly
Can also be regression




 **SVM**
`svm.SVC()` `svm.LinearSVC()`
Maximum margin classifier. Fundamental
Data Science algorithm



 **NAIVE BAYES**
`GaussianNB()` `MultinomialNB()` `BernoulliNB()`
Updating knowledge step by step
with new info



FEATURE REDUCTION

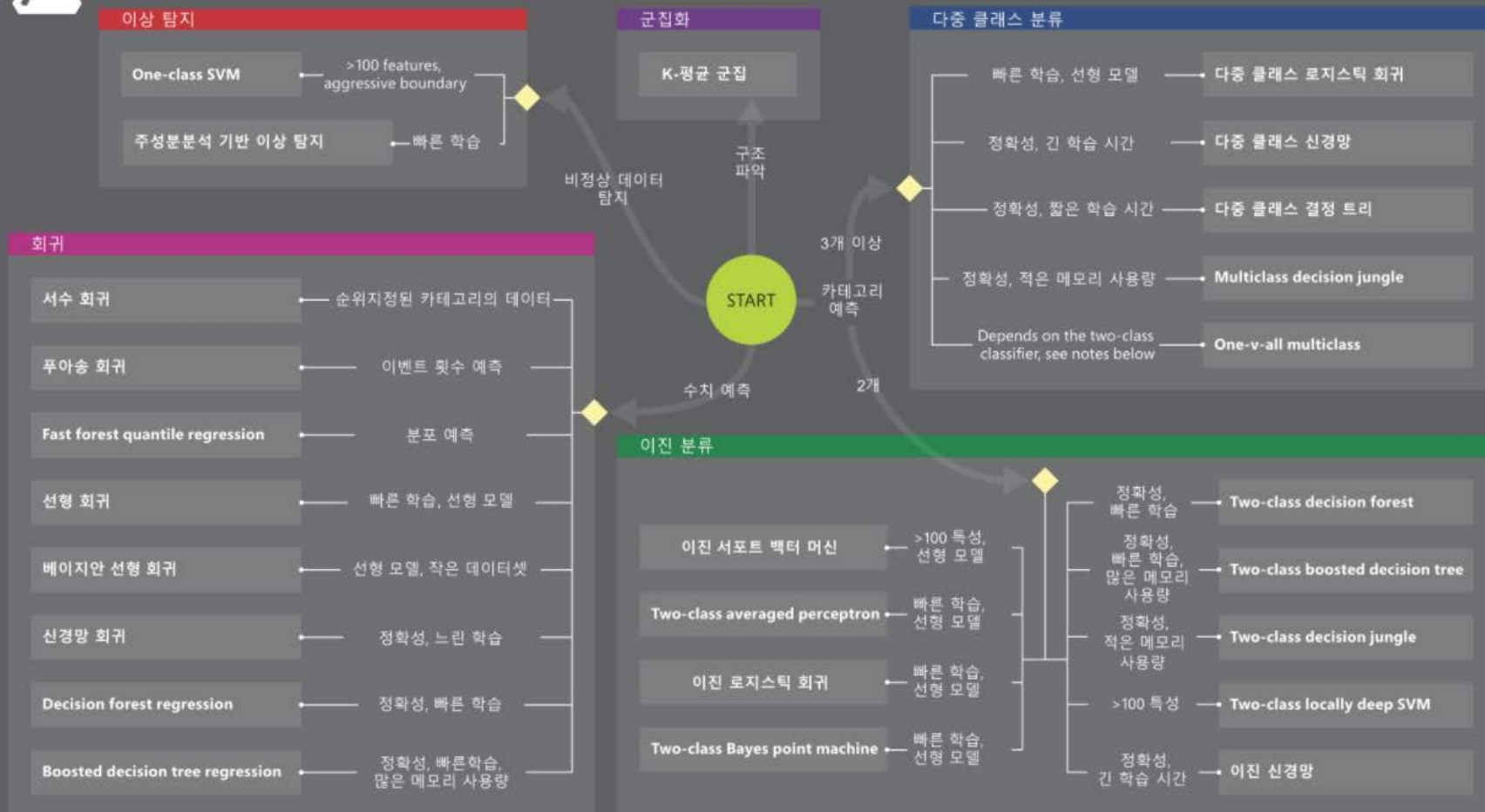
 **T-DISTRIB STOCHASTIC
NEIB EMBEDDING**
`manifold.TSNE()`
Visual high dimensional data. Convert
similarity to joint probabilities

머신러닝 알고리즘



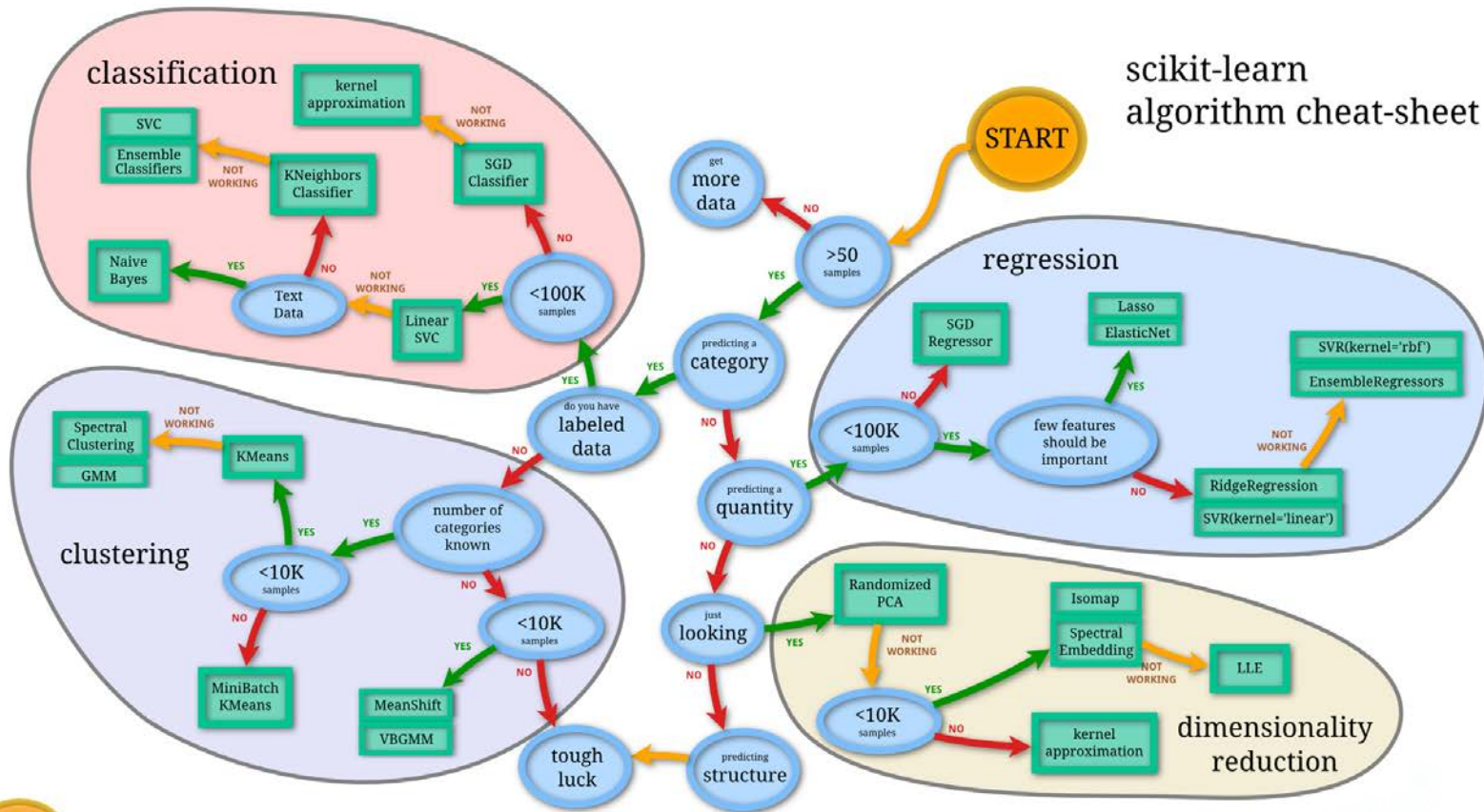
Microsoft Azure Machine Learning: Algorithm Cheat Sheet

This cheat sheet helps you choose the best Azure Machine Learning Studio algorithm for your predictive analytics solution. Your decision is driven by both the nature of your data and the question you're trying to answer.



머신러닝 알고리즘

scikit-learn
algorithm cheat-sheet



https://scikit-learn.org/stable/tutorial/machine_learning_map/

오렌지3를 이용한 인공지능 활용

오렌지3 설치

- 오렌지3

- 코딩없이 데이터 분석을 할 수 있는 데이터마이닝 도구

- 설치

- 아나콘다(Anaconda) 설치 후 오렌지3(Orange 3) 다운로드 받아 설치
- 아나콘다
 - <https://www.anaconda.com/products/individual>
- 오렌지3
 - <https://orange.biolab.si/download>



실습

- 문제정의
 - 주어진 수치 값으로 결과(더하기) 예측
- 데이터수집 및 데이터전처리
 - 엑셀을 이용한 데이터 만들기
- 알고리즘선택
 - 목표변수가 수치임으로 지도학습 회귀 알고리즘 선택
- 모델학습
 - 오렌지3로 학습
- 모델평가
 - 오렌지3로 평가
- 배포



결과



File Edit View Widget Options Help

Data

Visualize

Model

Evaluate

Test and Score

Predict...

Confusion Matrix

ROC Analysis

Lift Curve

Calibration Plot

Unsupervised

Data Table

View the dataset in a spreadsheet.
[more...](#)

i

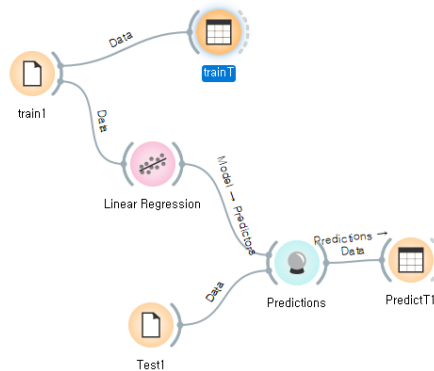
#

T

↵

||

?



trainT

Variables

☒ Show variable labels (if present)

☐ Visualize numeric values

☒ Color by instance classes

Selection

☒ Select full rows

Restore Original Order

☒ Send Automatically

?

📄

↶

20

📄

	결과	입력1	입력2
1	3	1	2
2	5	2	3
3	7	3	4
4	9	4	5
5	11	5	6
6	13	6	7
7	15	7	8
8	17	8	9
9	19	9	10
10	21	10	11
11	23	11	12
12	25	12	13
13	27	13	14
14	29	14	15
15	31	15	16
16	33	16	17
17	35	17	18
18	37	18	19
19	39	19	20
20	41	20	21

PredictT1

Variables

☒ Show variable labels (if present)

☐ Visualize numeric values

☒ Color by instance classes

Selection

☒ Select full rows

Restore Original Order

☒ Send Automatically

?

📄

↶

9

📄

해결문제

- 2020년까지 수능 응시자 데이터를 활용하여 2021년 졸업생 응시자를 예측

