

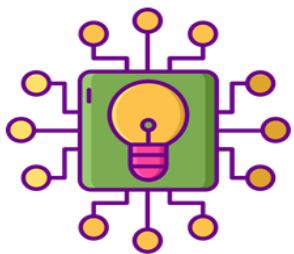
# **지도학습 & 비지도학습**

# 머신러닝이란?

## 머신러닝

### Machine Learning

컴퓨터가 스스로 학습하여  
인공지능의 성능을  
향상 시키는 기술 방법



## 지도학습

### Supervised

Target 예측을 위해  
데이터를 학습시킴

## 비지도학습

### Unsupervised

Target이 없이  
데이터의 특성을 추출

- 머신러닝은 데이터의 훈련 샘플들을 학습해 데이터의 규칙과 패턴을 학습하여 결과를 예측하는 알고리즘과 시스템을 연구하는 분야이다.
- 통계학에서 유래된 머신러닝 기법들이 많다.
- 머신러닝에서는 데이터에 대한 규칙을 프로그래밍으로 구현하지 않아도 파이썬 API 기반의 라이브러리로 손쉽게 모델을 구축할 수 있다.

# 지도학습 & 비지도학습

## 지도학습

레이블된 데이터를 학습하여  
미래 데이터 예측

정답 학습 시키기



토끼



고양이



이 사진은 무엇인가요?



토끼입니다.

## 비지도학습

레이블 없이 학습  
데이터의 숨겨진 특징 발견



데이터 특성 스스로 파악



긴 꼬리를 가졌고 귀가 뽕족함



짧은 꼬리를 가졌고 귀가 김

# 지도학습 & 비지도학습

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	adult_male	embark_town
0	0	3	male	22.0	1	0	7.2500	S	Third	True	Southampton
1	1	1	female	38.0	1	0	71.2833	C	First	False	Cherbourg
2	1	3	female	26.0	0	0	7.9250	S	Third	False	Southampton
3	1	1	female	35.0	1	0	53.1000	S	First	False	Southampton
4	0	3	male	35.0	0	0	8.0500	S	Third	True	Southampton
...	...	...	...	...	...	...	...	...	...	...	...
886	0	2	male	27.0	0	0	13.0000	S	Second	True	Southampton
887	1	1	female	19.0	0	0	30.0000	S	First	False	Southampton
888	0	3	female	NaN	1	2	23.4500	S	Third	False	Southampton
889	1	1	male	26.0	0	0	30.0000	C	First	True	Cherbourg
890	0	3	male	32.0	0	0	7.7500	Q	Third	True	Queenstown

891 rows × 11 columns

# 지도학습 & 비지도학습

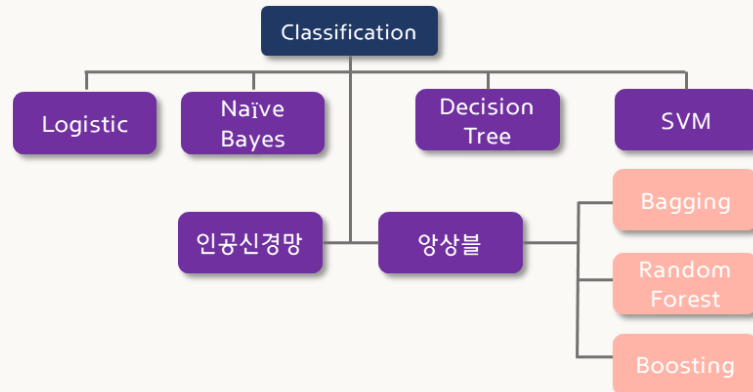
	sepal length	sepal width	petal length	petal width	gmm_cluster
0	5.1	3.5	1.4	0.2	0
1	4.9	3.0	1.4	0.2	0
2	4.7	3.2	1.3	0.2	0
3	4.6	3.1	1.5	0.2	0
4	5.0	3.6	1.4	0.2	0

	sepal length	sepal width	petal length	petal width
gmm_cluster				
0	5.006000	3.418000	1.464000	0.244000
1	6.554545	2.950909	5.489091	1.989091
2	5.904444	2.775556	4.193333	1.293333

# 지도학습 & 비지도학습

## 지도학습

분류 : Predict between several categories

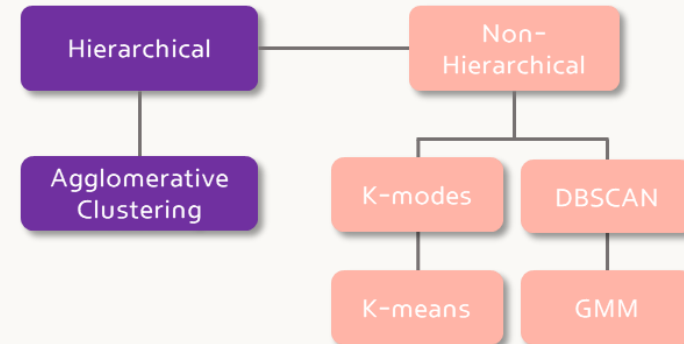


회귀 : Predict Value

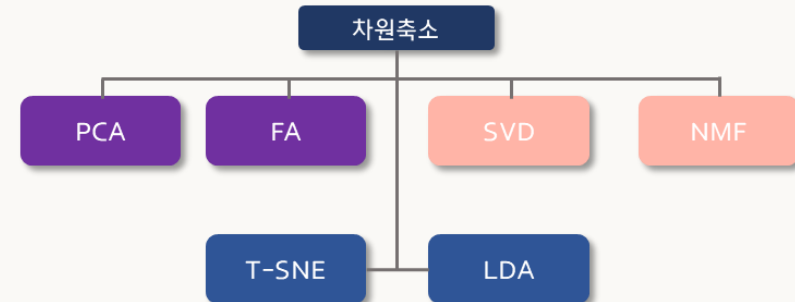


## 비지도학습

군집분석 : Discover structure



차원축소 : Dimension reduction



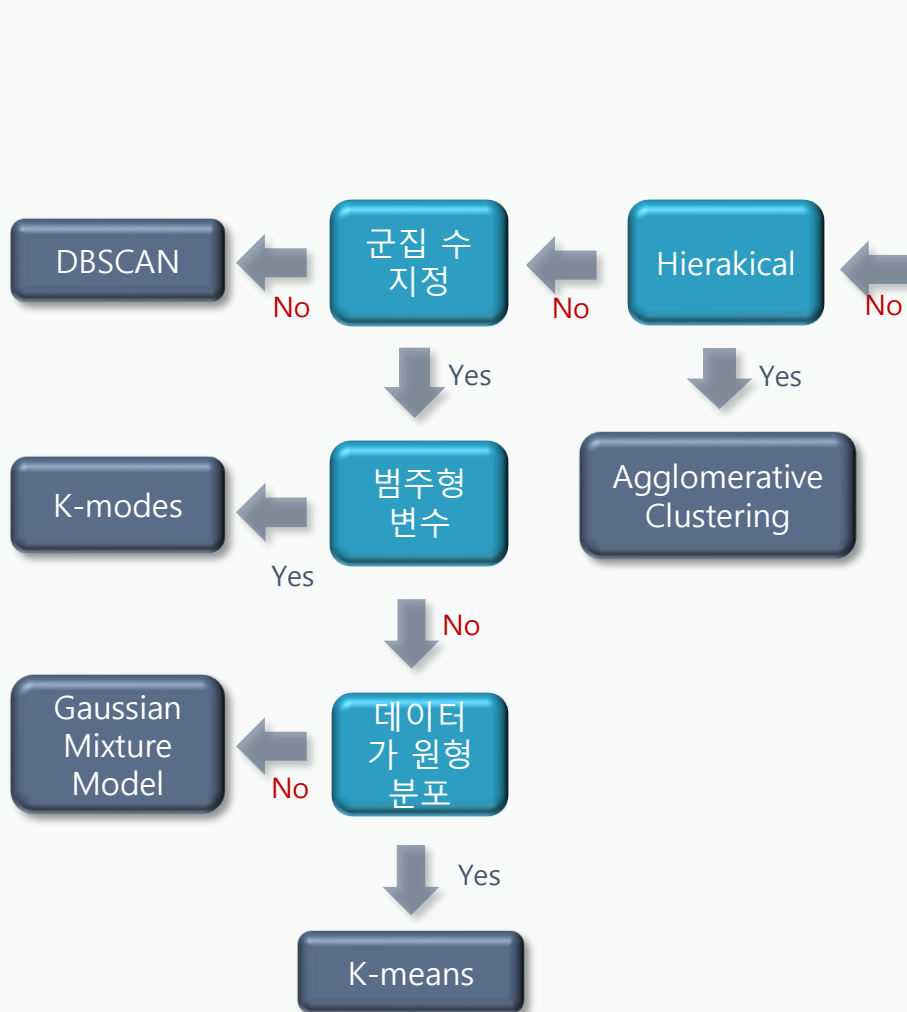
# 지도학습 & 비지도학습

---

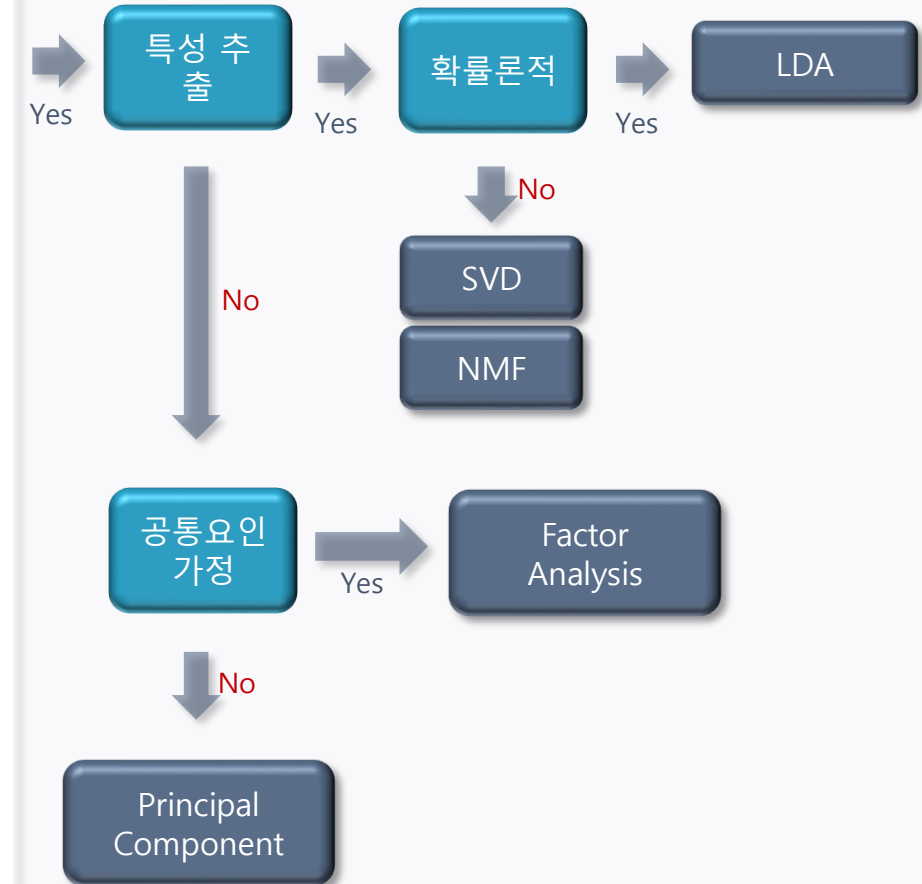
- 분류분석 : 학습된 데이터로 종속변수의 Class 예측
- 회귀분석 : 종속변수의 값을 예측 / 종속변수와 독립변수의  
관계 정량화
- 군집분석 : 데이터의 특성에 따라 묶음으로 유형별 특징을 분석
- 차원축소 : 다차원의 데이터 셋을 축소하여 새로운 차원의 데이터  
셋으로 변환

# 머신러닝 로드맵

## 군집분석 : Clustering



## 차원 축소 : Dimension Reduction





# 머신러닝 로드맵

