머신러닝

https://youtu.be/MzGvS5WsGrQ





머신러닝

• 기계가 학습 데이터로부터 규칙을 스스로 학습하여 새로운 데이터에 대한 적절한 작업을 수행하는 과정

• 지도학습 : 데이터가 레이블과 함께 주어짐

• 비지도 학습: 데이터가 레이블이 지정되어 있지 않음



지도 학습

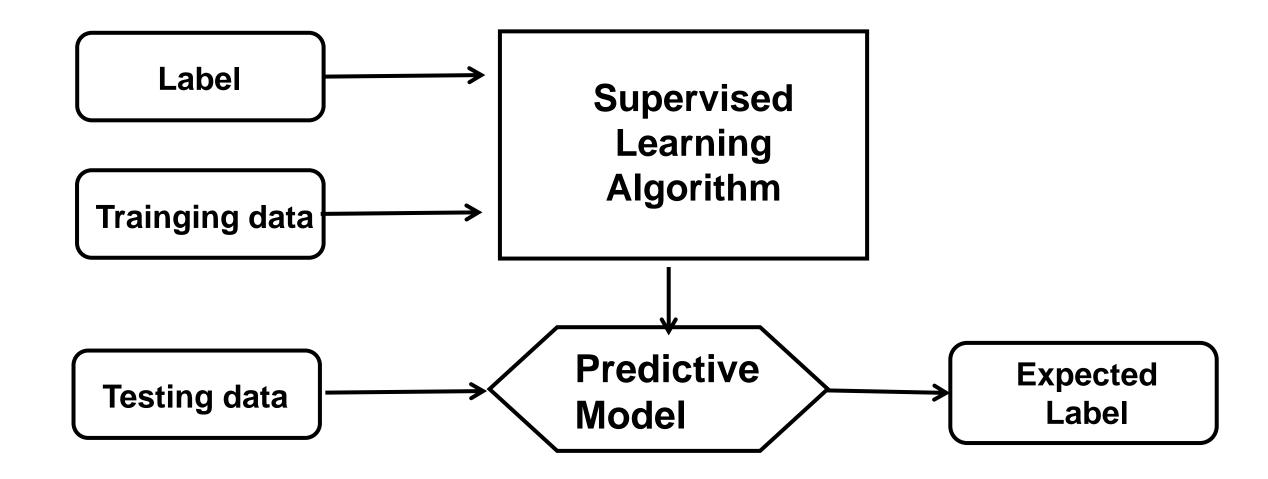
• 레이블이 지정된 학습 데이터를 추록하여 새로운 데이터의 출력 값을

예측하는 모델을 생성하는 머신러닝의 한 방법

• 신용카드의 사기성이나 보험 가입자의 보험금 청구 가능성 여부 예측에 효과적



지도 학습 프레임워크





지도 학습의 기법

구분	기법	설명
분류	이진 분류	두가지 중 하나로 분류
	다중 분류	여러 종류 중 하나로 분류
회귀	독립 변수 기반 분석	독립변수(입력값)의 개수에 따른 분석
	종속 변수 기반 분석	종속변수(독립변수에 의한 효과)의 개수에 따른 분석



지도 학습을 이용한 알고리즘

구분	알고리즘	설명
인공신경망 측면	CNN	– 이미지 기반 특징 추출, 차원 축소를 통한 인식, 예측
		– Convolution / Pooling Layer, Feature Map, Sub Sampling
	RNN	 현재와 과거 데이터 고려 순차 데이터 처리 순환 신경망
		– Input / Output / Hidden Layer, Time Unfolding, BPTT, LSTM
벡터 기반 측면	SVM	– 데이터를 두 클래스 분류 위해 Margin 최대 결정직선 탐색
		– Support Vector, Margin, 초평면, 결정 직선, 커널 함수
	회귀분석 (Regression)	– 변수 집합에서 독립/종속변수 간 상관관계를 함수로 표현
		– 독립/종속 변수, 회귀 계수, 최소 자승법, 회귀 방정식



비지도 학습

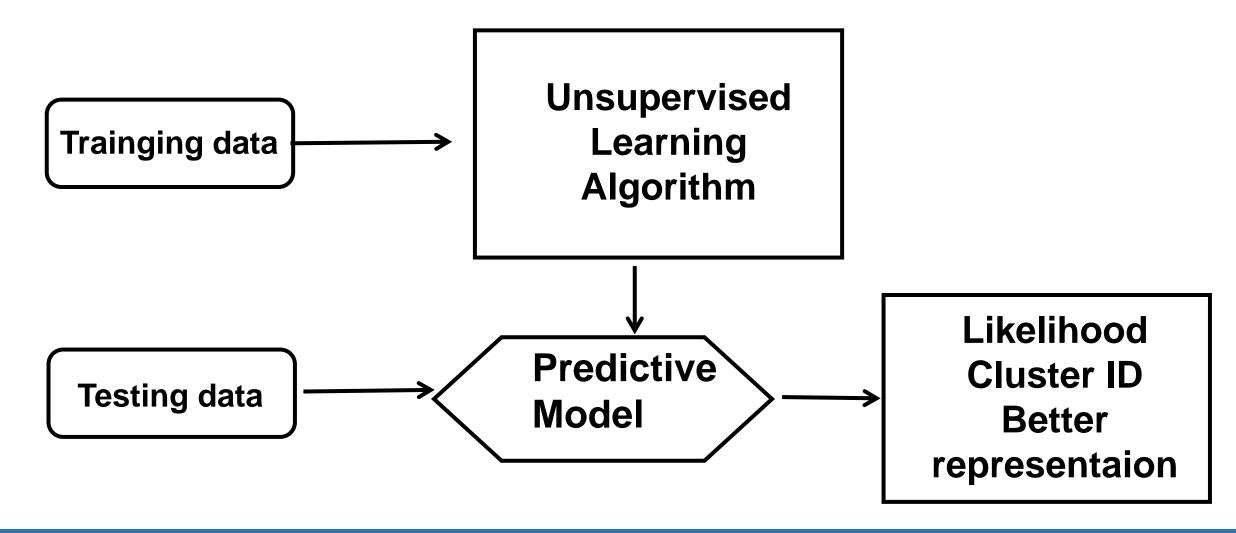
• 입력데이터에 대한 목표값 없이 데이터가 어떻게 구성되었는지를 알아내는 기법

• 트랜잭션 데이터를 탐색하여 내부 구조를 파악

• FDS 시스템과 군집화에 사용



비지도 학습 프레임워크





비지도 학습의 기법

구분	기법	설명
군집화	거리 기반 군집화	– 중심값과의 최소거리 기반 군집 형성 – 군집 수 선정 → 좌표 계산 → 중심값이동(반복)
	밀도 기반 군집화	– 군집을 이루는 벡터 밀도 기반 군집 형성 – 군집 벡터 수 선정 → 반경 내 군집 → 중심벡터 변 경(반복)
패턴인 식	전처리 / 특징 추출	– 표본화, 정규화, 노이즈 제거 – 주성분 분석, 데이터 마이닝
	모델 선택 / 인식	– Bagging/Boosting, 앙상블 학습



비지도학습을 이용한 알고리즘

구분	알고리즘	설명
데이터 관계 측면	K-Means	– 임의의 중심점 기준 최소 거리 기반 군집화 – Code-Vector, 유클리드 거리 계산, 노이즈에 민감
	DBSCAN	– 반경 내 데이터 벡터 밀도 기반 군집화 – ε, minPts, Core Point, 노이즈에 강함
	Mean-Shift	– 컴퓨터 비전, 머신 비전, 영상 분할 – 임의 영상을 몇 개 영역으로 분할, 군집화
특징 추출 측면 	주석분 분석	- 사물의 주요 특징 분석 및 추출 - 차원 축소, 축 상의 투영으로 표시



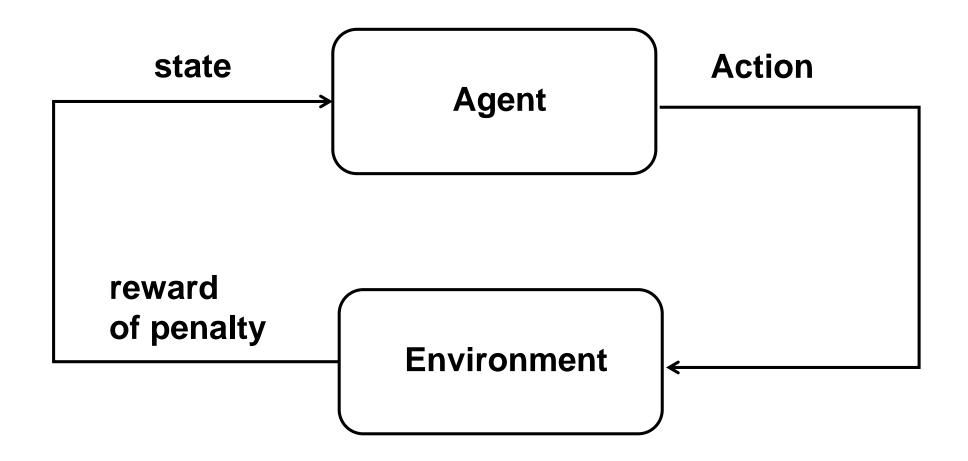
강화 학습 알고리즘

• 지도학습이나 비지도 학습 과는 다른 종류의 학습 알고리즘으로 구분

 에이전트가 현재 환경의 상태를 인식하여 행동하고 행동으로 인해 환경이 변화하는 일종의 동적인 상태에서 동작



강화학습 프레임워크





Q&A

