

인공지능(Artificial Intelligence)

Orientation

참고하면 좋을 책

Authors



Aston Zhang

Amazon Senior Scientist



Zack C. Lipton

Amazon Scientist
CMU Assistant Professor



Mu Li

Amazon Senior Principal Scientist



Alex J. Smola

Amazon VP/Distinguished Scientist

Vol.2 Chapter Authors



Brent Werness

Amazon Senior Scientist
Mathematics for Deep Learning



Rachel Hu

Amazon Scientist
Mathematics for Deep Learning



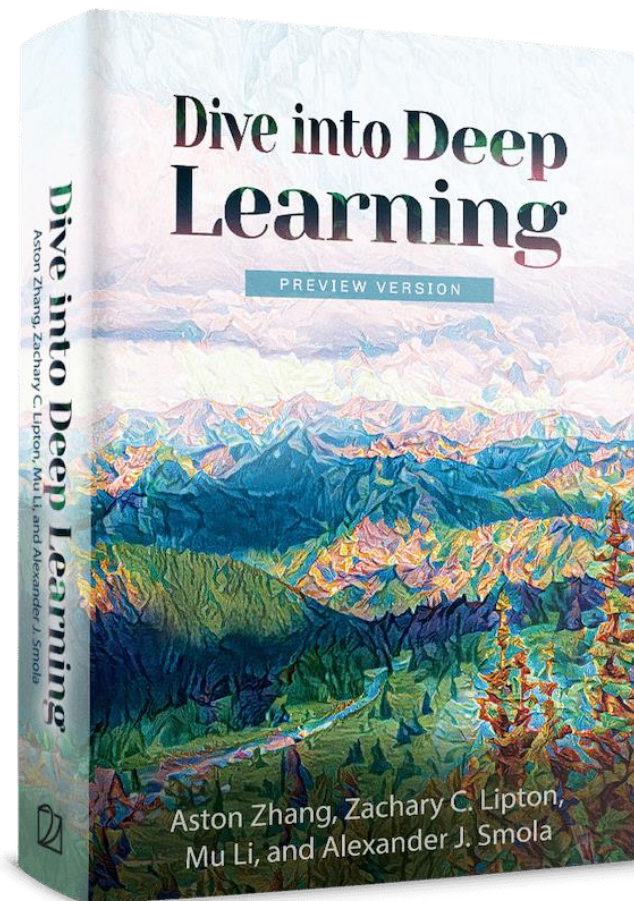
Shuai Zhang

ETH Zürich Postdoctoral Researcher
Recommender Systems



Yi Tay

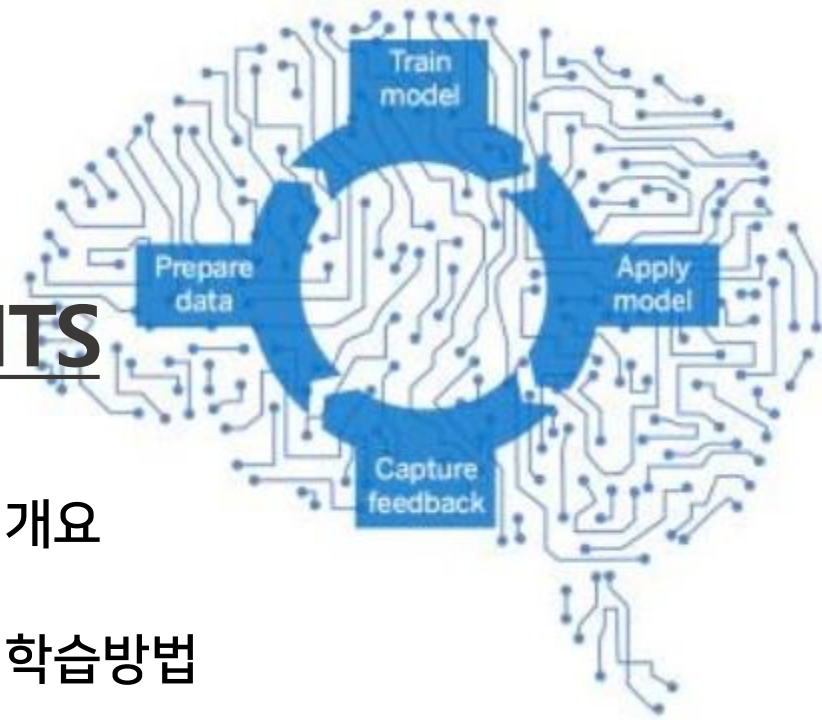
Google Scientist
Recommender Systems



Contents

CONTENTS

- 1 머신러닝의 개요
- 2 머신러닝의 학습방법
 - ✓ 지도학습
 - ✓ 비지도학습
 - ✓ 강화학습



머신러닝의 개요

머신러닝의 개요

머신러닝(Machine Learning)

- 규칙을 일일이 프로그래밍하지 않아도 자동으로 데이터에서 규칙을 학습하는 알고리즘을 연구하는 분야
- 머신러닝 알고리즘은 대규모 데이터 세트에서 패턴과 상관관계를 찾고 분석을 토대로 최적의 의사결정과 예측을 수행하도록 훈련됨

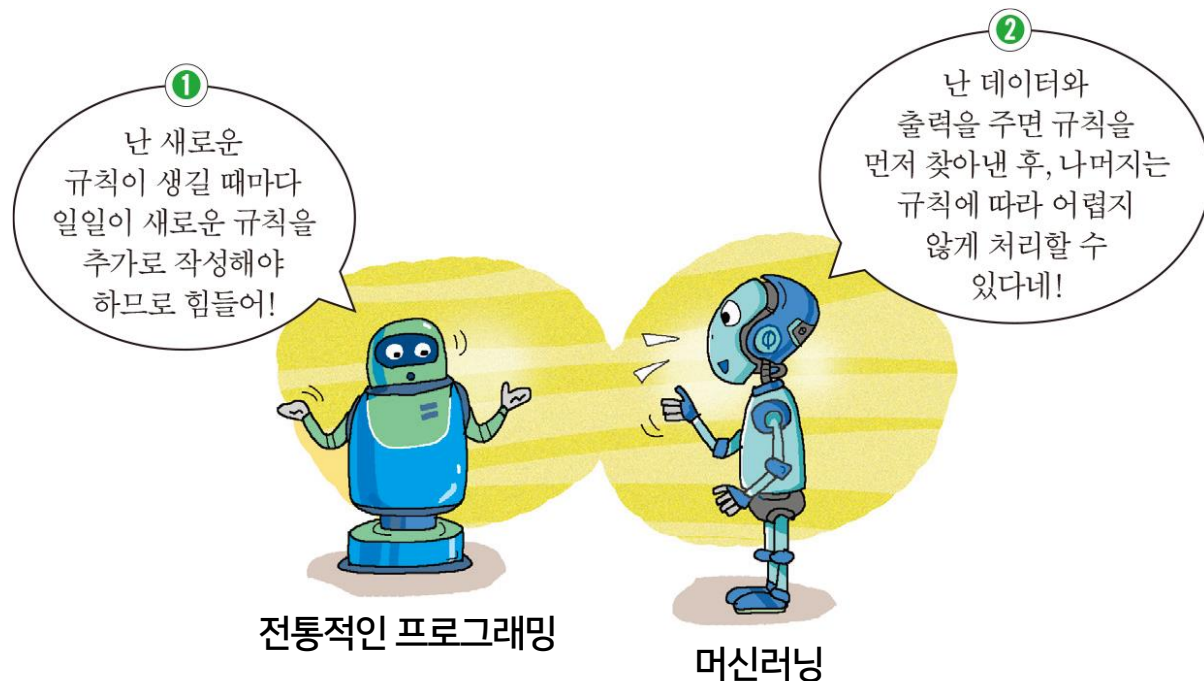
머신러닝을 통한 간단한 학습의 예

- 입력과 출력이 여러 개의 데이터 쌍으로 주어짐
ex. (1, 2), (2, 4), (4, 8), (7, 14), (5, 10) ...
- 학습 후, 출력이 입력의 2배임을 유추
- (3, ?), (8, ?) 등의 질문에 6, 16 등으로 답변

머신러닝의 개요

전통적인 프로그래밍과 머신러닝의 차이점

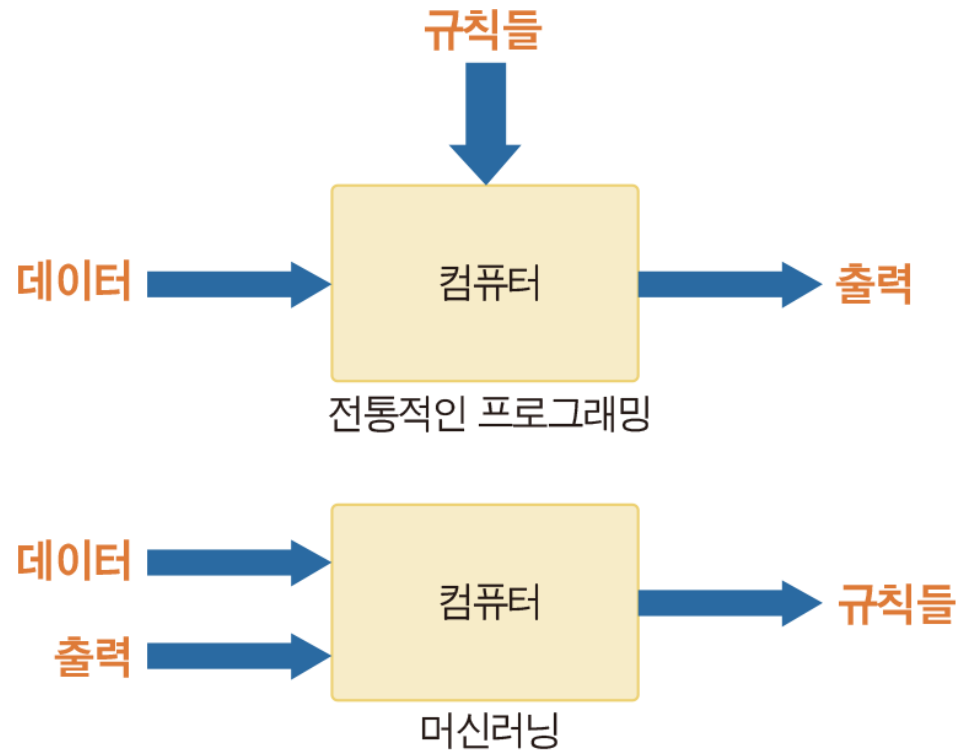
- 전통적인 프로그래밍에서는 모든 규칙들을 작성함
- 만약 규칙들이 추가로 작성될 경우, 유지 관리가 어려운 문제점
- 그러나 머신러닝은 시간에 따라 점차 효율이 향상됨
- 입출력 데이터의 관계를 학습하여 규칙을 생성



머신러닝의 개요

전통적인 프로그래밍과 머신러닝의 차이점

- 전통적인 프로그래밍에서는 데이터와 규칙이 결합하여 출력(Output)을 생성
- 머신러닝에서는 데이터와 출력(Output)이 함께 들어가서 규칙 생성



머신러닝의 개요

머신러닝의 활용 분야

활용 분야	응용
영상인식	문자인식, 물체인식
얼굴인식	Facebook에서의 얼굴인식
음성인식	Bixby, Siri, Alexa 등
자연어 처리	자동 번역, 대화 분석
정보 검색	스팸 메일 필터링
검색 엔진	개인 맞춤식 추천 시스템
로보틱스	자율주행 자동차, 경로 탐색

머신러닝의 개요

머신러닝과 인공지능과의 관계

- 머신러닝은 인공지능에 속하는 부분 집합
- 인공지능은 머신러닝을 포함하는 상위 개념



머신러닝의 개요

머신러닝과 인공지능의 차이점

- 머신러닝 : 데이터로부터 학습하여 지식을 획득
- 인공지능 : 지식을 획득한 후, 지식을 활용함



추구하는 개념과 목표가 다소 다름!

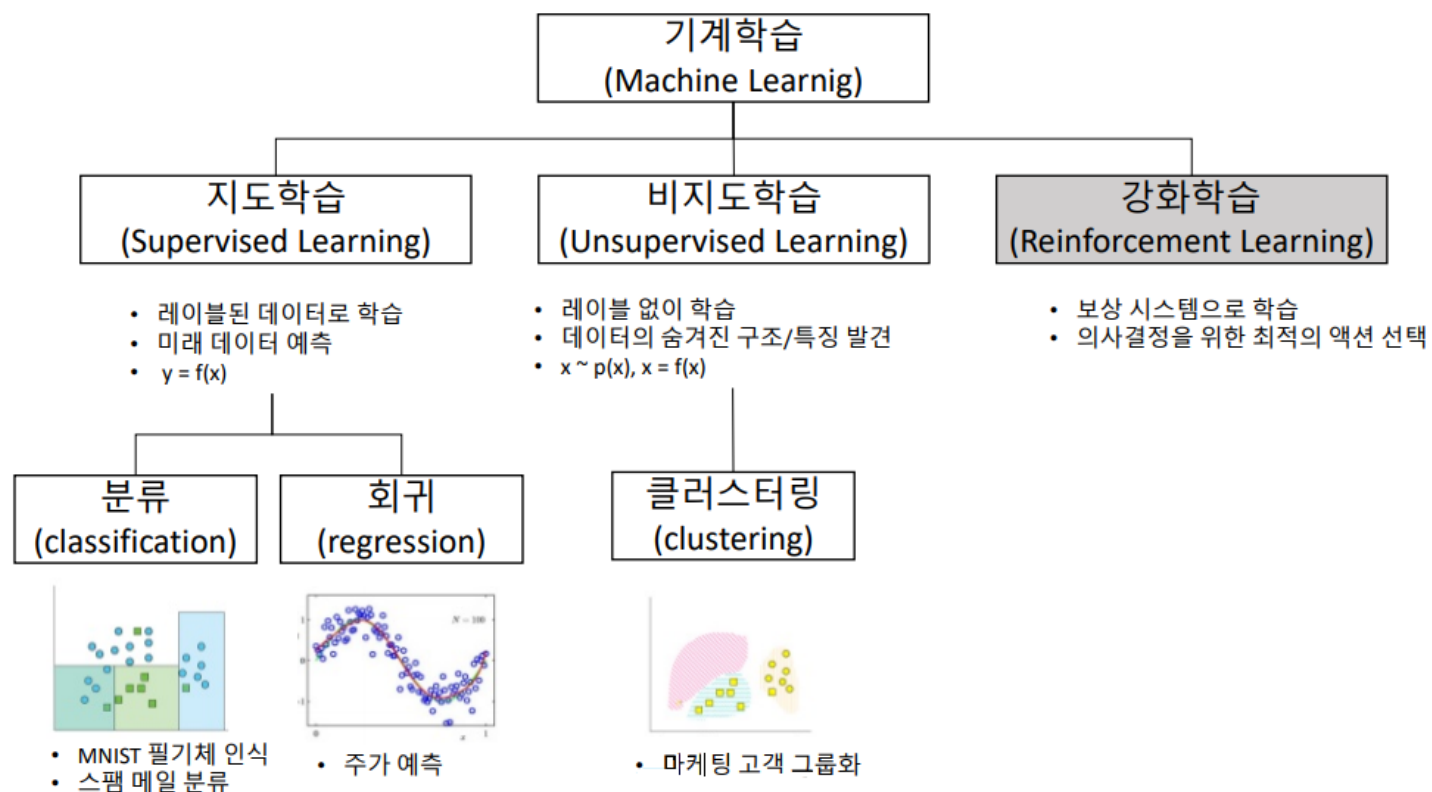
	머신러닝	인공지능
주요 활동	학습을 통한 지식의 획득	지식의 획득과 활용
구현과 실현	데이터로부터의 학습 실현	복잡한 문제 해결을 위한 지능의 구현
개발 목표	스스로 학습하는 알고리즘 개발	인간을 닮은 지능적인 시스템의 개발

머신러닝의 학습방법

머신러닝의 학습 방법

머신러닝에서의 학습 방법

- 학습의 형태에 따라 지도 학습, 비지도 학습, 강화 학습으로 구분
- 지도학습은 분류와 회귀, 비지도 학습은 클러스터링과 차원 축소로 다시 나누어짐



머신러닝의 지도학습

머신러닝의 학습 방법

지도 학습

- 알고리즘에 주입하는 훈련데이터에 레이블(label)이라는 원하는 답이 포함됨
- 주어진 입력과 정해진 출력을 연관시키는 관계를 학습



ex. 여러 가지 메일과 발송 기관 등을 샘플로 훈련하여 스팸메일인지 아닌지를 분류할 수 있도록 훈련

머신러닝의 학습 방법

지도 학습의 장단점

① 장점

- 이전의 경험으로부터 데이터 출력을 생성
- 경험을 사용하여 성능 기준을 최적화
- 다양한 유형의 문제 해결에 도움이 됨
- 비지도 학습에 비하여 성능이 좋음

② 단점

- 출력에 반드시 레이블이 있는 데이터들을 사용해야 함
- 일반적으로 레이블링 하는데 많은 시간이 걸림
- 빅데이터의 경우 엄청난 시간이 걸릴 수도 있음

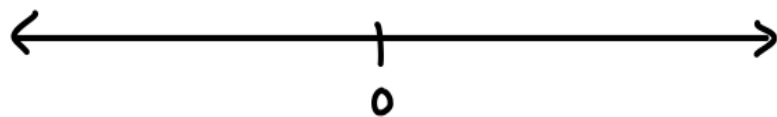
머신러닝의 학습 방법

지도 학습에서의 분류와 회귀

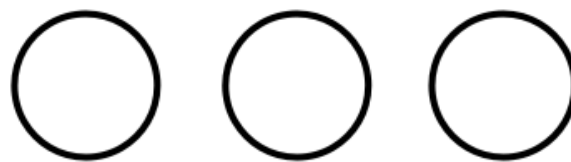
- 지도 학습은 **분류**와 **회귀**로 나누어짐
- 예측하려는 값이 연속값이면 회귀 문제, 이산값이면 분류 문제

cf. 연속값 : 연속하는 값 (ex. 0.31, 0.301, 0.3001처럼 끝없이 연속되어 나갈 수 있는 경우)

이산값 : 한정된 수로 끊어진 값 (ex. 매우 그렇다 / 그렇다 / 보통이다 / 그렇지 않다 / 매우 그렇지 않다)



연속값



이산값

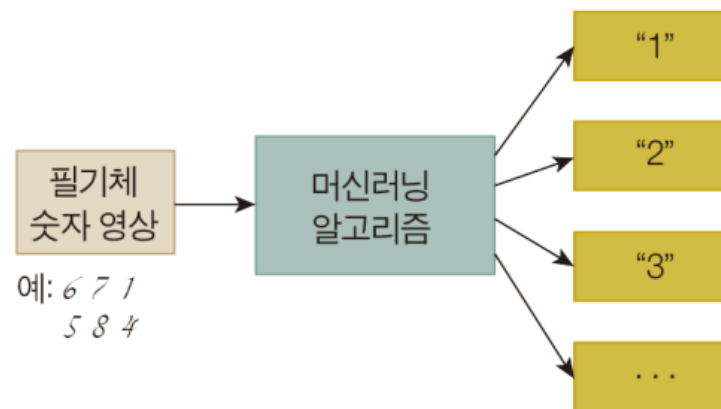
머신러닝의 학습 방법

① 분류(classification)

- 유사한 특성을 가진 데이터들끼리 묶어서 나누는 것
- 이항 분류 : 2개의 Label을 갖는 데이터가 들어왔을 때, 0 또는 1로 분류
ex. 스팸 메일/정상 메일, 합격/불합격 등
- 다항 분류 : 3개 이상의 Label을 갖는 데이터에 대한 분류 작업
ex. 0에서 9까지의 아라비아 숫자 인식



이항 분류

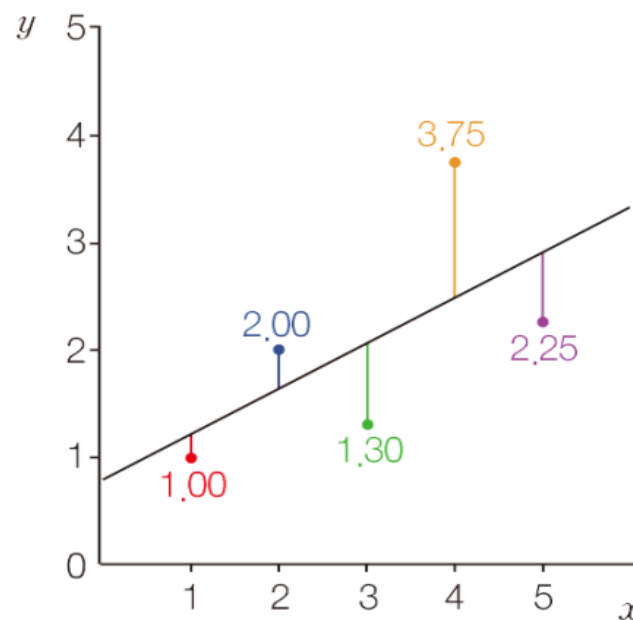
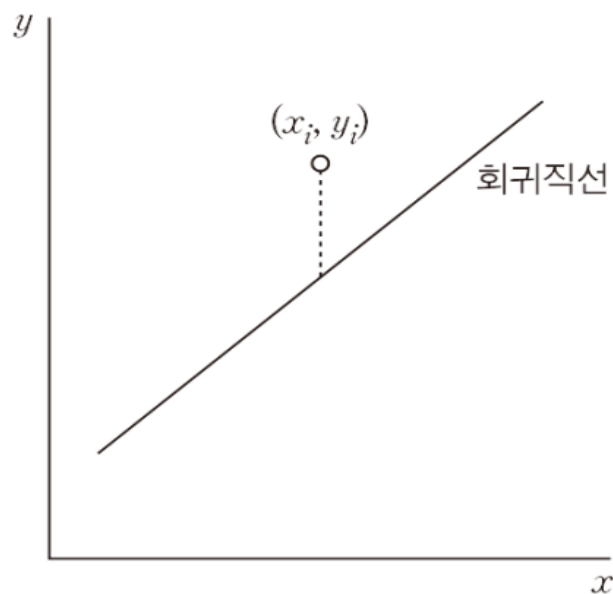


다항 분류

머신러닝의 학습 방법

② 회귀(regression)

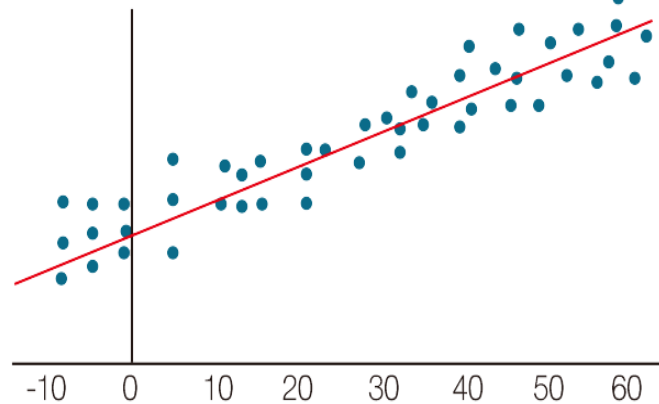
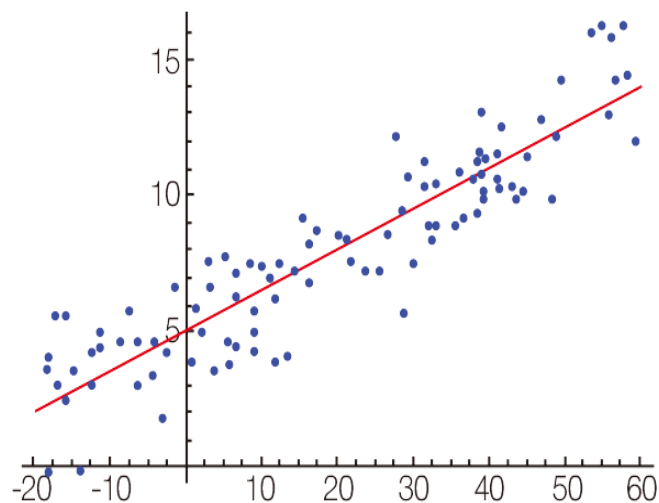
- 변수들 사이의 관계를 결정하는 통계적 측정
- 하나의 독립 변수를 사용하는 직선 형태의 '선형 회귀'
- 선형 회귀는 각 점에서 회귀 직선까지의 y축 방향의 거리 제곱의 총합을 최소로 해서 얻어지는 직선
- 직선 $y = a + bx$ 를 x 에 대한 y 의 회귀 직선이라 함



머신러닝의 학습 방법

회귀 분석(regression analysis)

- 독립변수가 종속변수에 영향을 미치는지 알아보고자 할 때 실시하는 분석 방법
- 변수 사이의 회귀에 대해 검정이나 추정을 하는 것
- 회귀 분석은 학습 데이터를 사용하여 출력값 예측
- 산출물은 항상 확률론적 의미를 내포



머신러닝의 학습 방법

회귀와 회귀 분석의 예측에의 활용

- 날씨에 대한 예측
- 월별 판매액을 보고 다음 달 판매액 예측
- 금융, 투자, 비즈니스적 가격 판단
- 금값이나 원유 가격 예측
- 주택 가격의 예측
- 장단기 주가 예측
- 원유 가격 추정 등

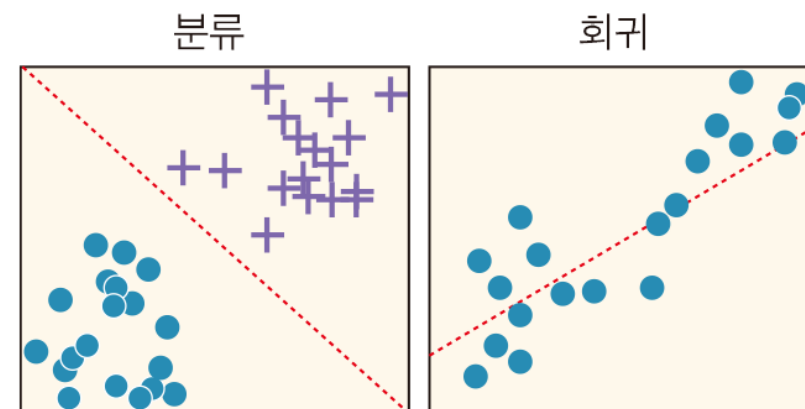
머신러닝의 학습 방법

분류와 회귀의 차이 구분

- 분류는 일정한 기준에 따라 명백하게 구분 짓는 것
- 회귀의 출력은 연속값으로 나타냄
- 회귀는 오차 제곱의 합을 최소화하는 직선을 긋는 작업이므로 명확히 직선으로 구별되는 것이 아님

분류와 회귀의 차이점 예시

- 분류의 출력은 남자/여자 등과 같은 선택식 출력
- "내일 날씨는 더울 것이다."와 같은 이분법적 선택
- "내일 기온?"에 대해 "18.3도로 추정된다." 등의 형태

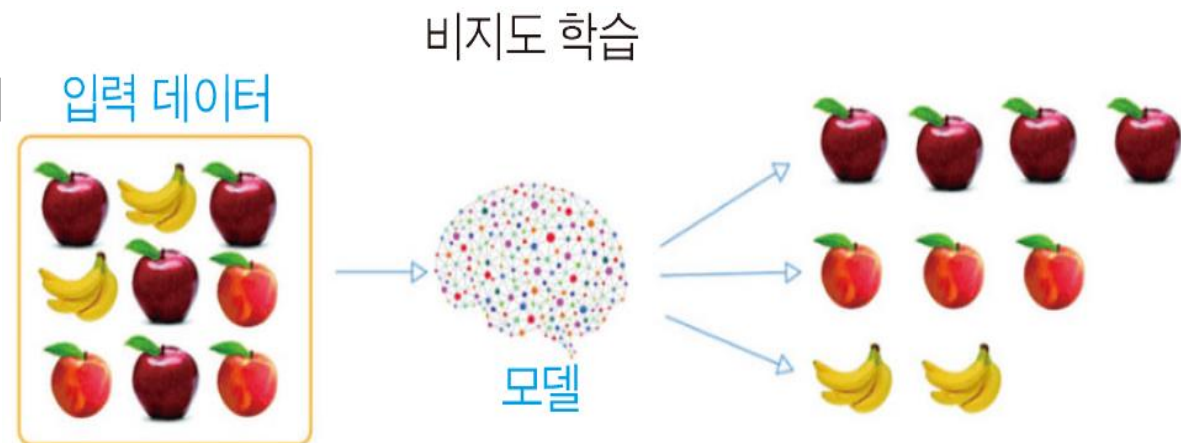
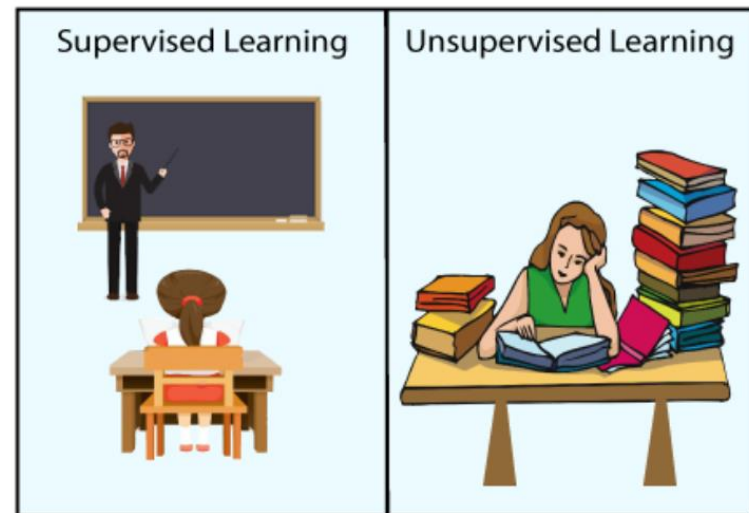


머신러닝의 비지도학습

머신러닝의 학습 방법

비지도 학습(unsupervised learning)

- 출력값을 알려주지 않고 스스로 모델을 구축하여 학습
- 비지도 학습은 입력만 있고 출력 즉, 레이블(label)이 없음
- 규칙성을 스스로 찾아내는 것이 학습의 주요 목표
- 또는 전문가에 의해 해석되어 다른 용도로 활용됨
- 데이터 마이닝(data mining) 기법은 비지도 학습의 예
- 주어진 입력에 대응하는 출력 정보 없이 학습
- 데이터 분류에 대한 정보가 전혀 없이 패턴을 찾거나 데이터를 분류하려고 할 때 사용하는 학습 방법
- 데이터에 레이블을 전혀 사용하지 않음
- 관계를 스스로 학습한 후, 과일들을 각 그룹으로 알아서 묶기



머신러닝의 비지도 학습

비지도 학습을 통한 클러스터링과 추천 시스템, 머신러닝의 동영상에의 활용

- 머신러닝에서의 주요 비지도 학습 방법
- 유튜브(YouTube)에 '내 관련 재생 목록'으로 응용됨
- 즐겨 찾던 동영상 리스트가 추천 리스트로 올라옴
- 온라인 광고에 많이 활용됨



[바둑랩 강좌] 두칸벌림 2선침입
바둑랩 강좌
조회수 44,926회



프로 9단들도 헛갈리는 함정수!
[무서운 함정수 5강]
바둑 도서관
맞춤 동영상 새 동영상

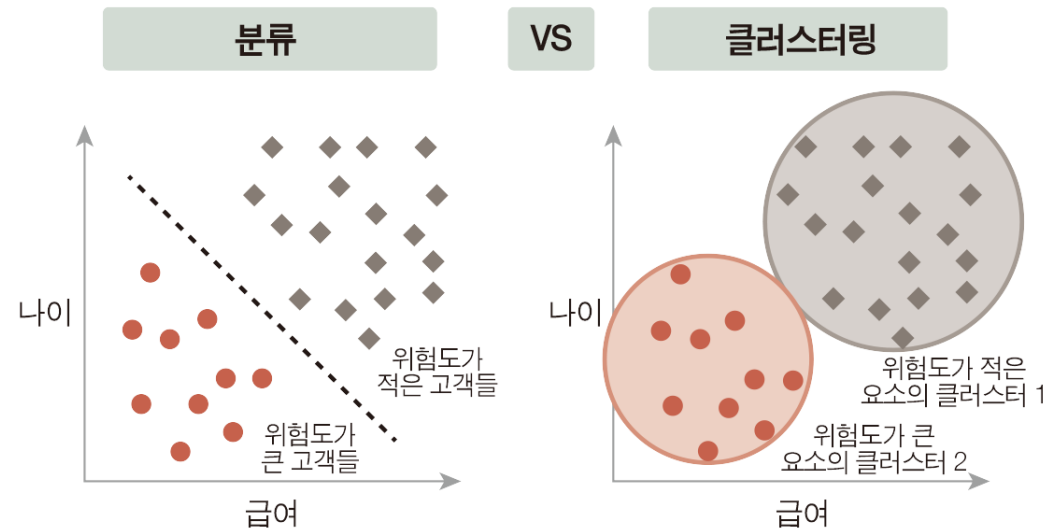


[바둑랩 강좌] 죽보에도 없는 꿈수를 당하신다면~
바둑랩 강좌
조회수 18,382회

머신러닝의 비지도 학습

분류와 클러스터링의 차이점

- 분류는 지도 학습 영역, 클러스터링은 비지도 학습 영역
- 분류는 데이터를 기준에 따라 직선으로 분류하는 것
- 클러스터링은 유사성에 따라 몇 개의 클러스터들로 묶는 것
- 급여, 나이, 위험도 상관관계의 예에서의 차이점의 예



머신러닝의 비지도 학습







비지도 학습의 주요 응용 분야

- 비슷한 성향의 고객을 그룹으로 묶기
- 블로그에서 주제별로 구분하기
- 유사한 꽃이나 동물들끼리 묶기
- 네트워크상에서의 비정상적인 접근의 탐지

머신러닝의 비지도 학습

추천 시스템의 활용과 기타 비지도 학습 방법

- 가령 교보문고에서 책을 검색하면 그 사람이 이전에 검색했던 도서나 관련 도서를 알려줌
- 사용자의 검색 경험 정보 파악, 적절한 광고 내보내기

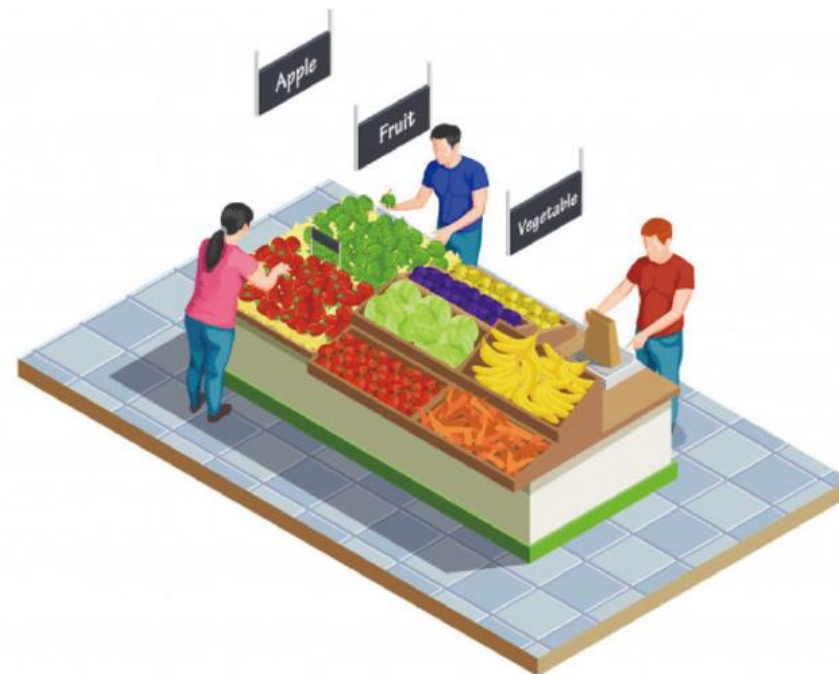
이 책을 구입하신 분들이 함께 산 책			이 책을 구입하신 분들이 많이 산 책		
					
두근두근 언플러그드 컴퓨팅 이재우 저 생능출판사 18,000원	소프트웨어와 컴퓨팅 사고 김대수 저 생능출판사 ★★★★★ 1건 20,000원	언플러그드 놀이 2 신갑천, 홍지연 저자 ... ★★★★★ 3건 13,500원 (10% 할인)	코딩을 위한 컴퓨팅 사고력 채성수, 오동환 공저 ... ★★★★★ 18건 10,800원 (10% 할인)	code.org를 활용한 컴퓨팅 사고력과 코딩 교육 제인 크라우스, 키키 프... 19,000원	놀이와 게임으로 만나는 코딩의 세계 짐 크리스티안 저/전성... ★★★★★ 32건 13,320원 (10% 할인)

- 그 외 비지도 학습 방법에는 가우스 혼합 모델, 계층적 클러스터링, PCA/T-SNE 등이 있음

머신러닝의 비지도 학습

(1) 클러스터(cluster)와 클러스터링(clustering)

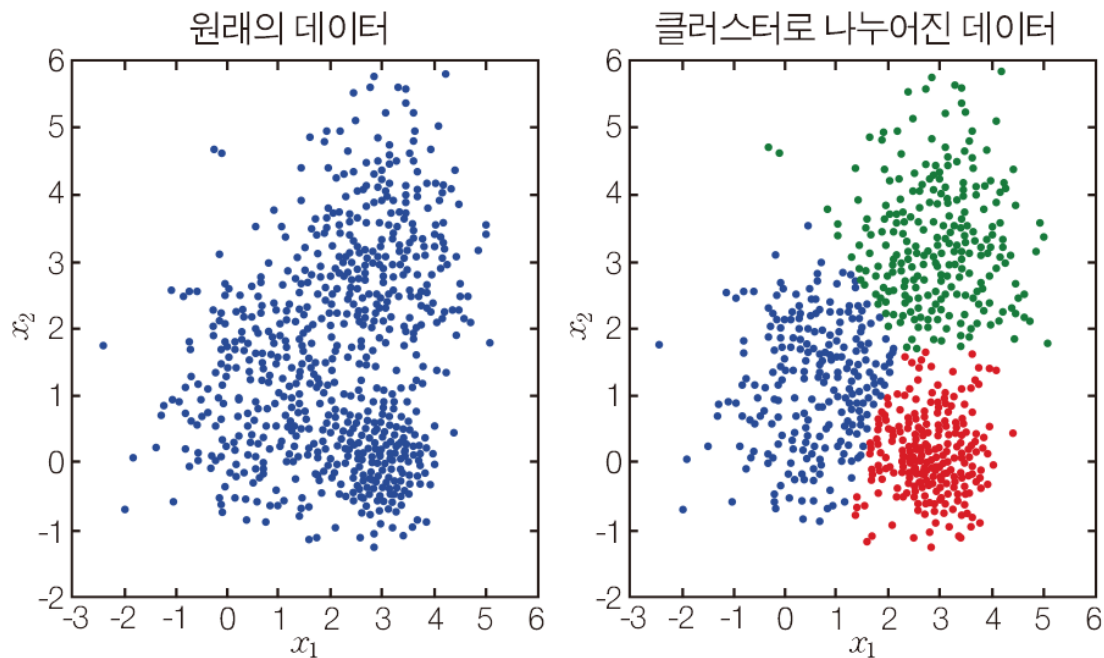
- 클러스터는 유사한 여러 개의 클래스로 나뉘어진 데이터
- 클러스터링은 유사한 특성을 가진 그룹들로 묶는 작업
- 같은 클러스터의 것은 다른 클러스터의 것보다 더 유사
- 이와 같은 유사한 것들끼리의 집합이 바로 클러스터



머신러닝의 비지도 학습

클러스터의 분류 예

- 원래 데이터를 3개의 클러스터로 분류해 놓은 예
- 빨간색, 파란색, 녹색



머신러닝의 비지도 학습

지도 학습과 비지도 학습의 특징 비교

기반	지도 학습	비지도 학습
입력 데이터	입력과 출력(값 또는 레이블)이 지정된 데이터를 사용하여 학습함	출력값이나 레이블이 전혀 없는 데이터를 사용하여 학습함
주요 기능	분류, 회귀	클러스터링, 추천 시스템
계산의 복잡성	비교적 간단함	상당히 복잡함
정확성	매우 정확함	다소 덜 정확함

머신러닝의 강화학습

머신러닝의 학습방법

강화 학습(reinforcement learning)

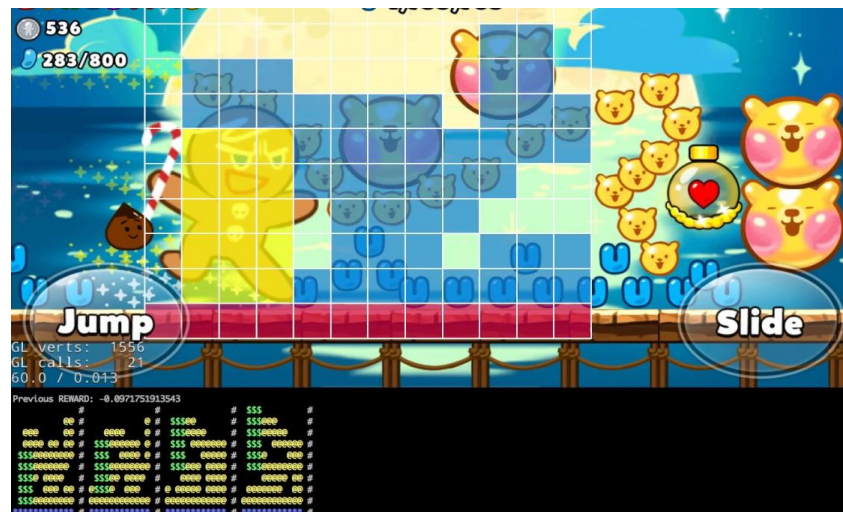
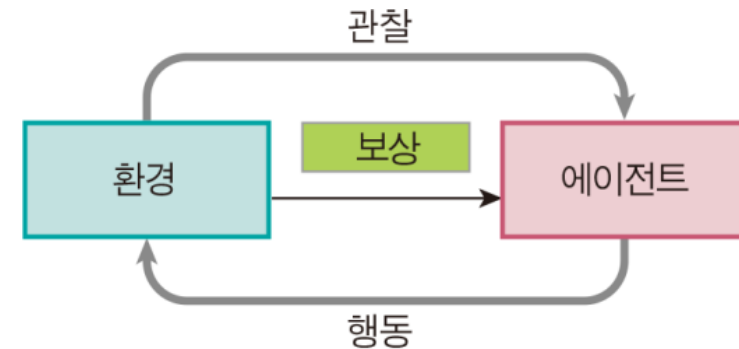
- 주어진 입력에 대응하는 행동에 대해 **보상**(reward)이나 **처벌**(punishment) 등이 작용하여 학습하는 방법
- 주어진 입력에 대한 출력, 즉 정답 행동이 주어지지 않음
- 주요 응용 분야로는 로봇, 게임, 내비게이션 등



머신러닝의 강화 학습

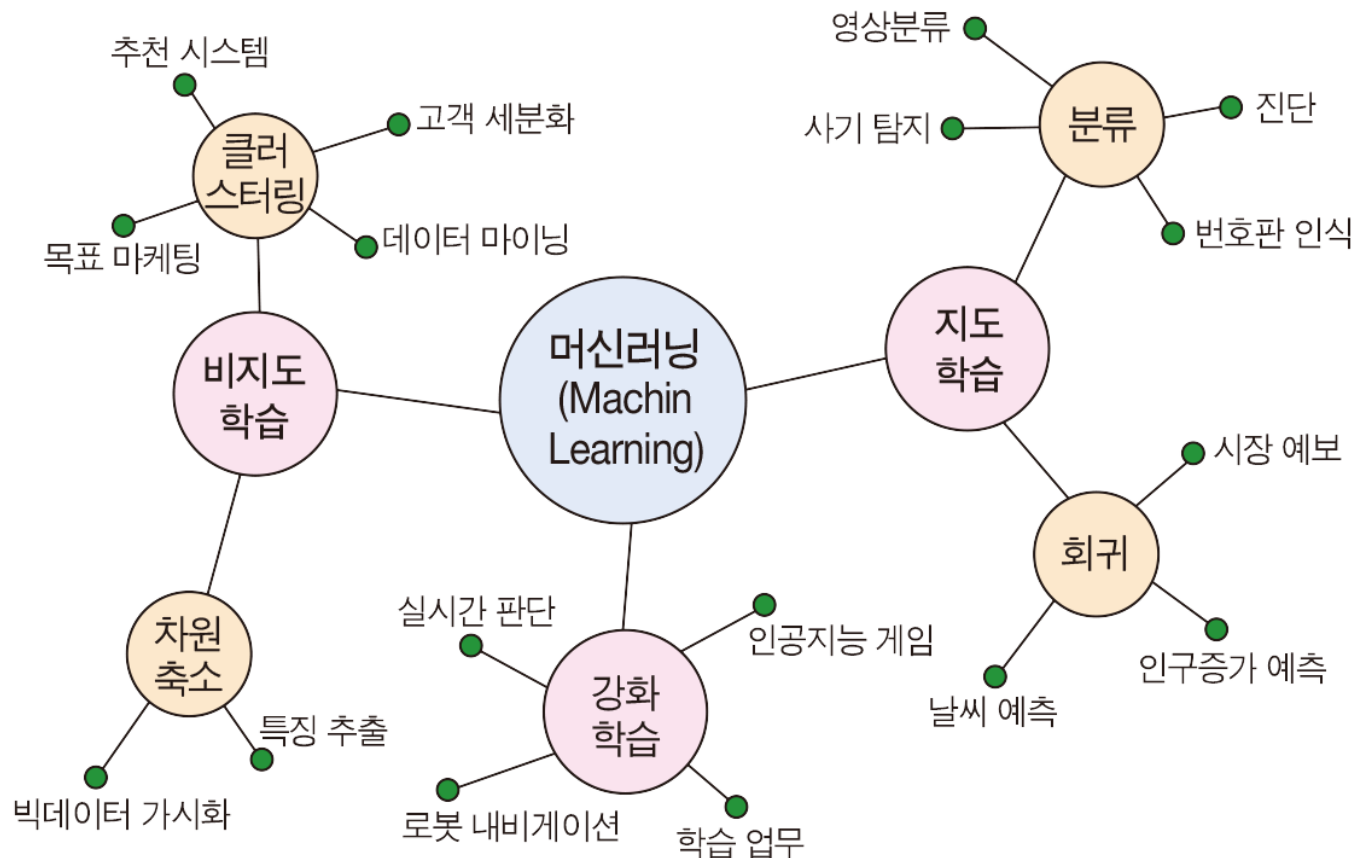
강화 학습의 응용 분야

- 보상(reward)이 주어지는 문제 해결에 매우 효과적
- 통신망, 로봇 제어, 엘리베이터 제어, 그리고 체스와 바둑 같은 게임에 주로 응용됨
- 알파고도 강화 학습을 통해 실력 향상
- 최근 게임에서는 거의 필수적으로 강화 학습이 사용됨



머신러닝의 학습 방법

머신러닝의 학습 방법과 활용 분야 체계



Thank you

