Machine learning

1. 지도학습

회귀 : 입력 변수 x에 대해서 연속형 출력 변수 y를 예측

분류 : 입력 변수 x에 대해서 이산형 출력 변수 y(class)를 예측

1. 비지도학습

* 출력변수y가 존재하지 않고 입력변수 x간의 관계에 대해 모델링 하는 것
* 군집분석 : 유사한 데이터끼리 그룹화
* PCA : 독립변수들의 차원을 축소화

현실에는 라벨링된 데이터가 부족하기 때문에 비지도 학습이 더 중요하다고 하는 말도 있다

1. 강화학습

* Action과 reward를 통해 큰 보상을 취하는 방법을 배움

ML의 종류

* 선형회귀분석 : 독립변수와 종속변수가 선형적인 관계가 있다라는 가정하에 분석
  + Cf) 비선형적(이 모델로는 판단하기 어려운 단점), 선형적
  + 직선을 통해 종속변수를 예측하기 때문에 독립변수의 중요도를 파악하기 쉬움
* 의사결정나무 : 독립변수의 조건에 따라 종속변수를 분리 / 이해하기 쉬우나(간단하나) 과적합이 잘 일어남. 단일 모델로는 잘 사용안됨