

머신러닝이란?

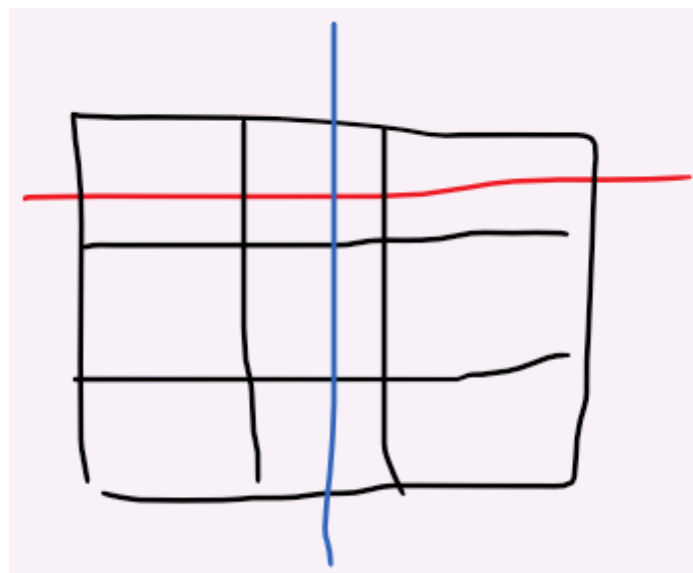
인간의 결정을 기계한테 맡긴다

판단력(모델)을 기계한테 부여한다

모델을 만드는 과정=학습

현실을 데이터화한다

Column(열)
=특성(feature)
속성(attribute)
변수(variable)
field



Row(행)
=개체(instance)
관측치(observed
value)
기록(record)
사례(example)
경우(case)

독립변수(원인)

종속변수(결과)

독립변수와 종속변수의 관계를 인과관계라고 한다.

모든 인과관계는 상관관계이다.

하지만, 모든 상관관계가 인과관계인 것은 아니다.

(인과관계는 상관관계에 포함된다.)

지도학습(supervised learning) ----배움

-분류(Classification): 예측하고 싶은 종속변수가 이름일 때 보통 분류라는 머신러닝의 방법을 사용한다.

-회귀(regression): 예측하고 싶은 종속변수가 숫자일 때 보통 회귀라는 머신러닝의 방법을 사용한다.

비지도학습(unsupervised learning)

-군집화(clustering): 비슷한 것들을 찾아서 그룹을 만드는 것

Cf) 그룹을 만들고 난 후에는 각각의 물건을 적당한 그룹에 위치시키겠죠? 이것이 분류입니다.

-연관 규칙 학습(Association rule learning): 서로 관련이 있는 특성(열)을 찾아주는 머신러닝의 기법

Ex)쇼핑 추천, 음악 추천, 영화 추천

Cf) 관측치(행)를 그룹핑 해주는 것 : 군집화

특성(열)을 그룹핑 해주는 것 : 연관규칙

강화학습(Reinforcement Learning)-----경험