

ML 웹2

2017.6.28.

ML 나온 이유

- explicit programming 의 한계
- 예를들어 스팸필터라면 팝니다, 팝~니다,팝 니 다, 파압니다 등 너무 많은 선언이 있어야 스팸 가려낼 수 있음
- 개발자가 알려줘서 판단하는 것이 아닌 데이터를 기반으로 컴퓨터가 학습하여 판단하는 알고리즘 = ML

supervised / unsupervised

- 레이블 달린 데이터셋이 있으면 지도학습 아니면 비지도학습
- 결과에 따라 regression, binary classification, multi-label classification
- 스팸필터의 경우 지도학습 - binary classification

linear regression

- 이 많이 쓰이는데 그 이유는 세상 많은 부분이 선형적 비례관계가 있기 때문
- 예를 들면 공부 많이 하면 성적이 좋다.
- 스타트업경험이 많으면 대박날 확률이 높아진다. 등
- 가설과 실제 데이터의 제곱을 최소화하는 $Y=Wx+b$ 에서 W, b 를 찾는 것이 목표

linear regression

- w 와 $\text{cost}(w)$ 를 그래프로 표현하면 2차방정식 모양
- 어느점에서 시작하든 최저점으로 가기위해서는 기울기가 작아지는 쪽으로 이동
- 1차원적인 데이터 말고 행렬로 놓아도 적용가능
- $H(x) = Wx + b \quad \rightarrow \quad H = WX$