

ML lec 01

기본적인 Machine Learning의 용어와 개념 설명

2021년 3월 12일 금요일 오후 10:55

Machine Learning

입력을 하게 되면, 입력을 기반으로 어떤 데이터를 읽어서 보여주는 것을 explicit programming 이라 한다.

특정 상황에서는 explicit programming하기 어려운 경우가 있다.

ex) 스팸 메시지 필터링, 자동운전과 같이 rule이 너무 많은 경우

아서 사무엘이 "우리가 일일이 프로그래밍하지 말고, 어떤 현상에서 자동적으로 배우게 하자"고 말하며 머신러닝에 대해 정의함

아서 사무엘(Arthur Lee Samuel)은 머신 러닝을 "컴퓨터가 명시적으로 프로그램되지 않고도 학습할 수 있도록 하는 연구 분야"라고 정의했다. (Andrew Ng 교수, Coursera Machine Learning 강의 Lecture 1, <https://www.coursera.org/learn/machine-learning/>)

Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed. (1959, by Arthur Lee Samuel)

일반적인 컴퓨터 프로그램은 "A 입력에 B 조건이 성립하면 X를 동작시킨다"를 인간이 작성하는 반면, 머신러닝에서는 A라는 정보를 입력할 때 대답이 X가 되는 조건 B를 찾도록 기계를 학습시킨다. ("구글에서 배우는 딥러닝" 닷케이 빅 데이터 저, 서재원 옮김, 영진닷컴, 28p)

Supervised/Unsupervised Learning

머신러닝을 할 때, 프로그램을 학습을 하기 위해서 어떤 데이터들이 미리 주어져야 하는데 학습하는 방법에 따라서 2가지 방법으로 나뉘어진다.

Supervised learning (지도학습)

Lable이 정해져 있는 데이터를 가지고 학습(Training Set)

ex) 고양이, 개, 머그컵, 모자 사진 구분하기

Unsupervised learning (비지도학습)

구글 뉴스, 비슷한 단어 모음과 같이 비슷한 뉴스들을 Grouping 하는 경우에는 미리 Lable을 정해주기 어렵다.

Lable을 직접 만들어주는 것이 아닌 데이터를 보고 직접 학습

Types of Supervised Learning

공부하는데 소요된 시간을 기반으로 시험 성적 예측

- regression (회귀분석)

공부하는데 소요된 시간을 기반으로 시험에 통과했는지 실패했는지 예측

- binary classification (이진분류)

공부하는데 소요된 시간을 기반으로 학점 예측

- multi-label classification