

Machine Learning Lecture Note

This lecture note is based on the CS229 lecture materials from Stanford University.

Lecture 1. Supervised Learning

머신러닝(Machine Learning)이란?

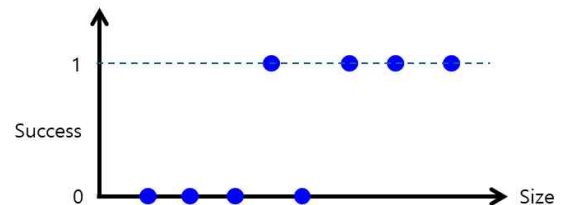
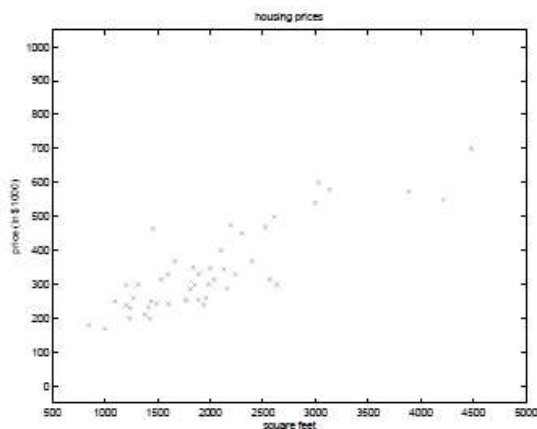
- 정의: Field of study that gives computers the ability to learn without being explicitly programmed. (Arthur Samuel 1959)
- 인공지능 > 머신러닝 > 딥러닝

머신러닝 종류

- 지도학습: 훈련데이터에 레이블(label)이 포함됨
- 비지도학습: 훈련데이터에 레이블이 없음
- 준지도학습: 일부 레이블이 있는 데이터를 사용
- 강화학습: 주어진 환경(environment)에서 행동(action)을 수행하고 관찰하여 그 결과 얻게 되는 보상(reward)을 통해 다음 행동을 결정

데이터에 레이블을 포함하는 것은 정확한 데이터 학습을 기대할 수 있지만 시간과 비용이 많이 발생한다.

Supervised Learning (지도학습)



Learning (학습) : 데이터 (x, y) 에 대하여, $x \rightarrow y$

Training example : (x^i, y^i)

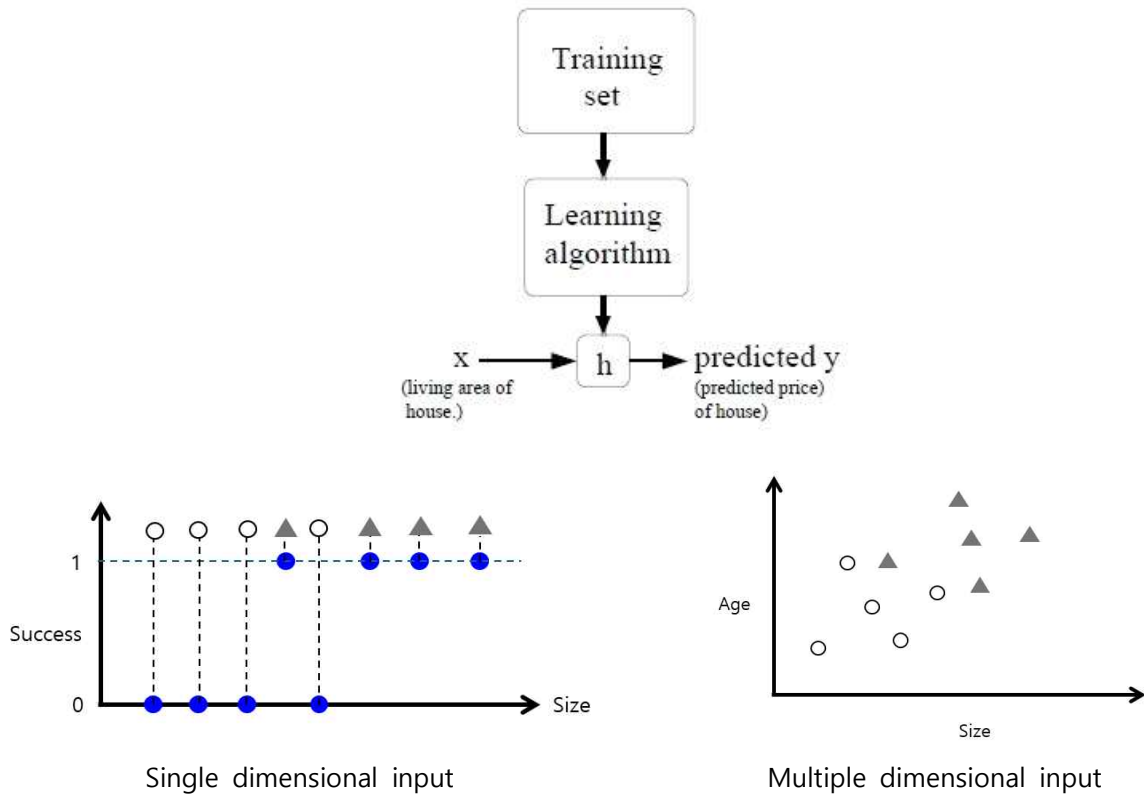
Training data set : $\{(x^i, y^i); i = 1, \dots, m\}$

"Regression" : 데이터 x 를 통해 y 를 예측

"Classification" : 데이터 x 를 통해 y 를 분류

데이터 셋 (Data set)

Supervised Learning (지도학습): 입력 x 와 레이블 y 를 함께 훈련에 사용하고, 학습 알고리즘을 통해 새로운 입력 x 가 주어질 때 y 를 획득함 (e.g., 컴퓨터 비전)



Machine Learning Strategy (Learning Theory) : 데이터 수집과 다양한 알고리즘 적용을 통한 응용에서 systematic engineering의 접근법이 요구된다.

Systematic engineering : 전체 수명 주기 동안 시스템을 설계, 개발 및 관리하기 위한 논리적이고 총체적인 접근법

Unsupervised Learning (비지도학습) : 레이블 y 없이 입력 데이터 셋을 통해 데이터의 특징 또는 의미를 추출하는 것 (e.g., 구글 news)

