

## 1. 머신러닝

### 1) 정의

-> 명시적으로 프로그램 하지 않고 데이터로부터 학습하는 알고리즘

- The field of study that gives computers the ability to learn without being explicitly programmed (Author Samuel)
- A computer program is said to learn from experience E with respect to some class of Tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E. (Tom Mitchell)

### 2) 머신러닝과 통계

-> 머신러닝과 통계는 데이터를 바탕으로 실제 데이터의 함수( $y = f(x)$ )에 대해서 추정(estimate)하려는 시도이다. (근본적으로 둘은 아주 다르지 않다)

-> 하지만

- 머신러닝은 실제  $y = f(x)$ 를 추정함에 있어서 Loss Function을 도입하여 실제 y값과 가장 차이가 적은 함수  $\hat{f}$ 를 구한다.  
(emphasis on the use of computers to estimate complicated functions)
- 통계는  $y = f(x)$ 을  $\hat{f}$ 로 추정함에 있어서 이 함수가 정말 유의성을 검토하는 데 초점을 둔다. (hypothesis testing)  
(proving confidence intervals around these functions)

### 3) 머신러닝이 다루는 세계 - Uncertainty

-> Machine Learning은 Uncertainty가 존재하는 상황 하에서  $y = f(x)$ 를 추정한다.

-> 불확실성은 다음과 같은 이유로 존재한다.

- 우리가 추정하려는  $y = f(x)$ 에서 모든  $(x, y)$ 를 보고 추정하는 것이 아니기 때문이다.
- $y = f(x)$ 에서  $f$ 에 대한 가정 자체가 정보의 손실이 있다.

-> 따라서  $y = f(x)$ 를 추정함에 있어서  $y = \hat{f}(x) + \epsilon$ , 항상 error가 존재할 수 밖에 없다. (irreducible error)

$$\begin{aligned} E(Y - \hat{Y})^2 &= E[f(X) + \epsilon - \hat{f}(X)]^2 \\ &= \underbrace{[f(X) - \hat{f}(X)]^2}_{\text{Reducible}} + \underbrace{\text{Var}(\epsilon)}_{\text{Irreducible}}, \end{aligned}$$

### 4) 머신러닝의 기반

- Probability Theory - 불확실성을 수량화해서 표현한다.
- Decision Theory - 확률 이론을 바탕으로 최적의 모델을 선택하는 기준을 제시한다. (Minimize Expected Loss)
- Information Theory - 실제  $y = f(x)$ 와 내가 추정해서 얻은  $y = \hat{f}$ 과 차이에 대한 척도를 제공한다.  
-> Decision Theory의 Loss Function의 근간이 된다.

1-2)에서 머신러닝과 통계는  $y = f(x)$ 를 추정(estimate)하는 것이라고 했는데 도대체 추정(estimate)이란 무엇이고 추정량(estimator)이란 무엇인가?