

Why AI?

"결정"

Learning

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

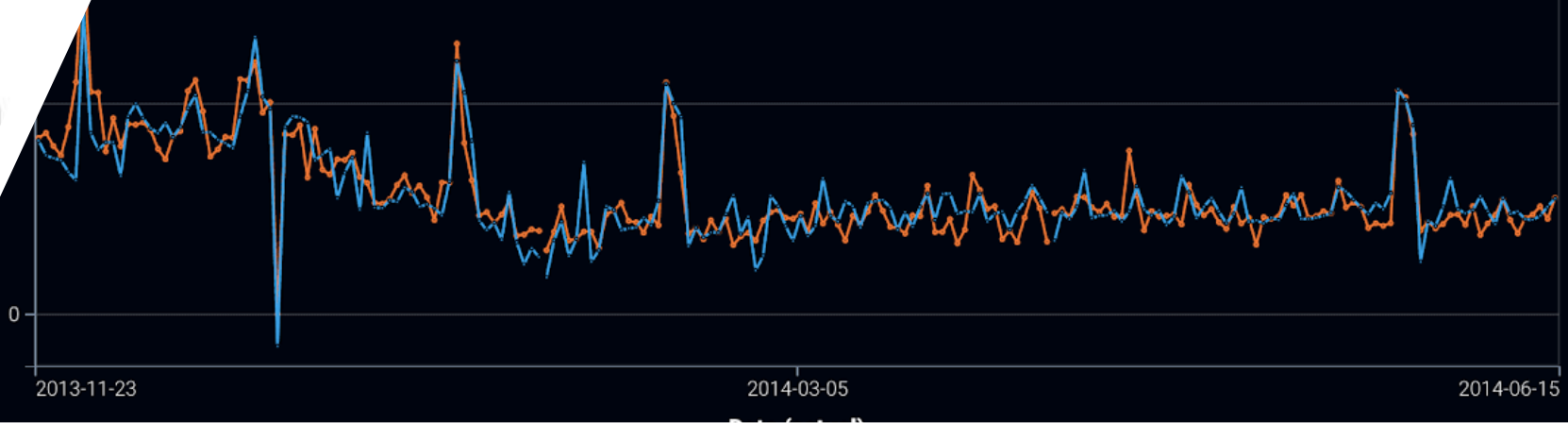
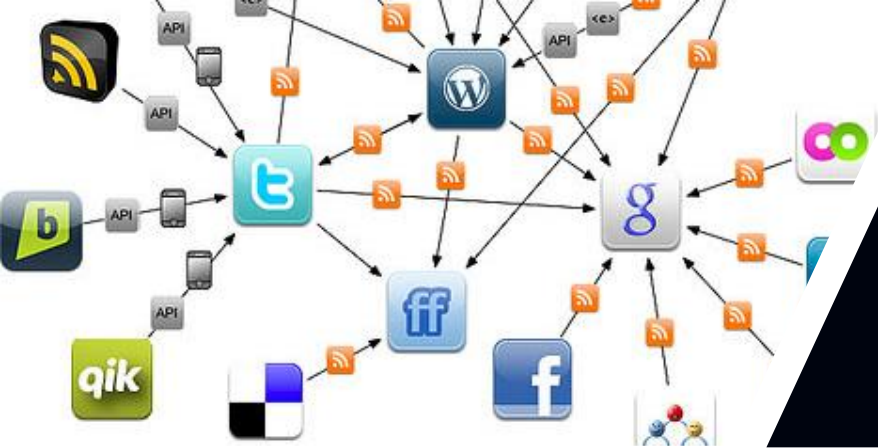
Rule Based Learning

하늘	기온	습도	바람	수온	예보	Picnic
맑음	따뜻	적당	적당	따뜻	변동 없음	가능
맑음	따뜻	높음	강함	따뜻	변동 없음	가능
구름	추움	높음	적당	차가움	변화	불가능
비	추움	높음	강함	차가움	변동 없음	불가능

If Else 문장의 나열

Picnic|하늘 = True, Picnic|기온 = True, Everything is True!!

➔ **Function Approximation**



In real World



Probability Theory

불확실성(uncertainty)는 현실에서 중요한 개념 중 하나

TO BE OR NOT TO BE, THAT IS THE QUESTION.
- William Shakespeare

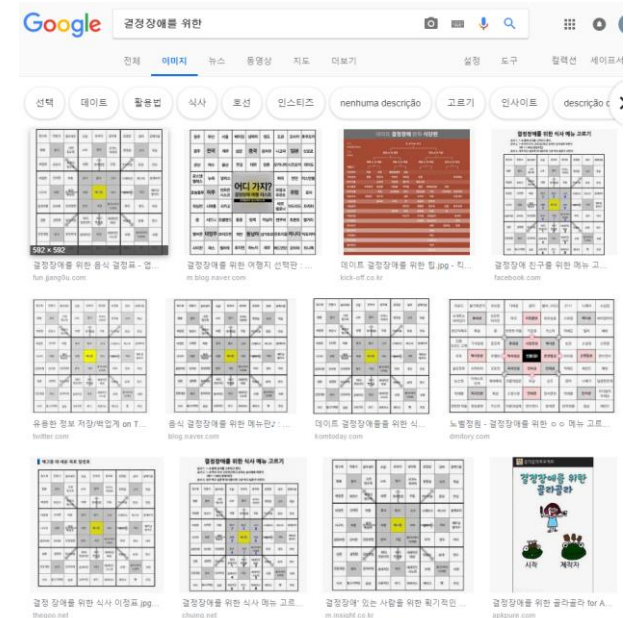


결정장애(햄릿 증후군)

당신은 메이비족인가?

- ☐ 1980년대에 태어나 1990년대에 학창 시절을 보냈다.
- ☐ 아날로그가 추억 속으로 사라지고, 손에 닿는 대부분의 사물들이 눈 깜짝할 새에 디지털화하는 광경을 지켜보았다.
- ☐ TV보다는 컴퓨터나 스마트폰이 더욱 친숙하다.
- ☐ 사생활은 보호받아야 한다고 생각하지만, 나의 일상을 SNS에 기록하는 것에는 별 부담이 없다.
- ☐ 만나본 적 없는 SNS 친구들도 여럿 알고 있으며, 현실의 친구들과 만나서도 종종 SNS를 뒤적거린다.
- ☐ "Yes" 혹은 "No" 대신, "Maybe(글쎄/그런 듯)"라고 답하는 경우가 많다.

<http://blog.naver.com/kerry76>



불확실성이 발생하는 이유?

- 너무 적은양의 데이터 (위험 요소 회피)
- 너무 많은 데이터 (너무 많은 데이터)
- 관찰데이터에 포함된 노이즈

Probability Theory

- 불확실성을 정확하고 정량적으로 표현
- 신뢰 할 만한 수학적인 프레임워크를 제공

Term Frequency - Inverse Document Frequency(TF-IDF)

TF-IDF는 TF X IDF 연산 결과

TF : (단어 빈도, Term Frequency)

IDF : (역문서 빈도, Inverse Document Frequency)

DF : (문서 빈도, Document Frequency)

TF 표현식

Boolean Frequency : $tf(t, d) = 0 \text{ or } 1$

Log scale Frequency : $tf(t, d) = \log(f(t, d) + 1)$

IDF 표현식

$$idf(t, D) = \log \frac{|D|}{|\{d \in D : t \in d\}| + 1}$$

Measuring similarity

