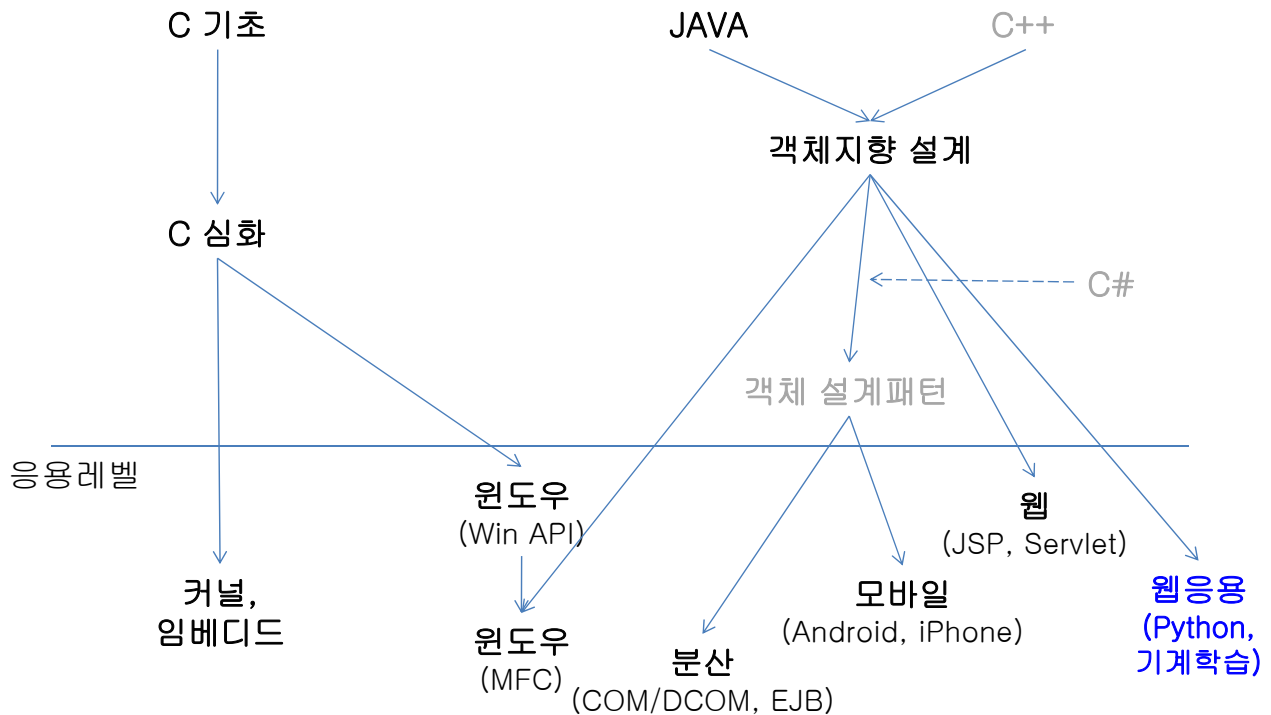


<절차적 프로그래밍>

<객체지향 프로그래밍>



• 참고교재

- 집단지성 프로그래밍, 토비 세가란 저, 윤종완 역, 한빛미디어, 2008
- 인공지능, 머신러닝, 딥러닝 입문, 김의중 저, 위키북스, 2016
- 머신러닝 이론 입문, 나카이 에츠지 저, 위키북스, 2016
- 실전예제로 살펴보는 집단지성 프로그래밍, 사트남 알랙 저, 전희원 역, 인사이드, 2010
- 어서와 파이썬은 처음이지, 천인국 저, 인피니티북스, 2016

인공지능 개요 (Artificial Intelligence)

“인공지능, 머신러닝, 딥러닝 입문”, 김의중 저, 위키북스, 2016

알파고와 구글 딥마인드

- **DeepMind**

- 2011년 Demis Hassabis, Shane Legg, Mustafa Shuleyman이 공동 설립한 회사
- Reinforcement Learning(강화학습) 분야에서 최고의 전문성을 갖추
- 2014년 1월 구글이 4억 달러(약 4400억원)에 인수됨
 - 구글의 자율주행 자동차, 로봇 사업에 시너지를 내기 위해

– 구글에 인수된 후, 강화학습 알고리즘인

DQN(Deep Q-Network) 프로젝트 추진

- 고전적 컴퓨터 게임 Atari 2600(핑퐁, 벽돌깨기, 스페이스 인베이더, 시퀀스트, 빔라이더, 엔듀로, 큐버트)와 비교
- 벽돌깨기, 핑퐁, 엔듀로에서 전문가보다 월등한 결과 얻음
- DQN이 받은 유일한 교육은 점수를 얻는 방법, 조이스틱을 조작하는 요령이 전부
- DQN은 모니터에 나오는 상황을 인식하고 점수를 더 많이 얻는 방법을 스스로 터득
- 컴퓨터가 스스로 학습하고 진화한 사실을 입증

인간과의 게임

• IBM의 Deep Blue

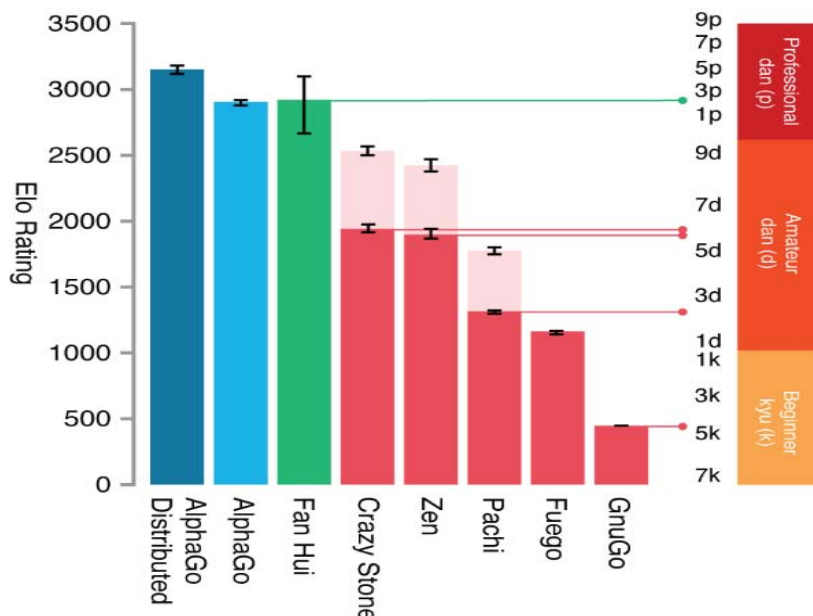
- 1996년 세계 체스 챔피언 Garry Kasparov와의 대결에서 3승 2패로 사람이 승리
- 1997년 재대결에서는 2승3무1패로 컴퓨터가 승리

• IBM의 Watson

- 2011년 Jeopardy 퀴즈 프로그램에서 역대 최강 후보자들과 대결에서 Watson이 우승

- **Google의 AlphaGo**

- 2015년 10월 유럽 바둑챔피언 Fan Hui 2단을 5대0으로 이김
- 2016년 3월 프로 바둑기사 이세돌 9단과의 대결에서 4승 1패로 AlphaGo가 승리
- 바둑에서 가능한 전체 게임의 수는 $19 \times 19 \Rightarrow 361!$
- 호주 심리학자 버마이스터가 제시한 경우의 수
 - 바둑 10^{575} , 체스 10^{123}
- SW: AlphaGo ver.18, HW: CPU 1202개, GPU 176개



- **Elo rating**

- 각종 경기에서 각 팀 또는 선수의 실력을 상대적으로 평가해서 계산한 수치
- Elo 값이 클수록 이길 수 있는 확률이 높아짐
- 이세돌 9단과의 대결 이후 지속적인 학습을 통해 4,500점에 도달

인공지능의 역사

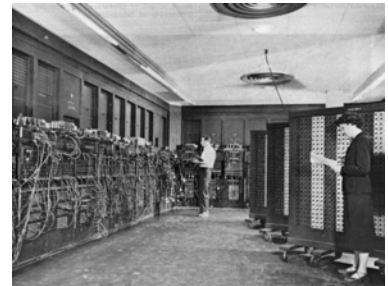
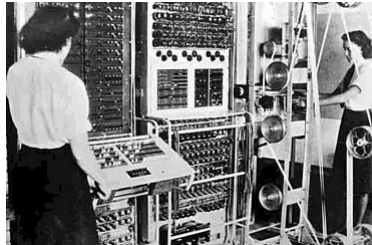
- 여러 학문이 연계된 융합학문
 - 컴퓨터과학, 수학, 통계학, 철학, 심리학, 의학, 언어학 등
- Stuart Russel & Peter Norvig의 정의
 - 인간처럼 생각하는 시스템
 - **인간처럼 행동하는 시스템**
 - 보고, 듣고, 움직이고, 운전하기
 - Ex) 자연어처리, 자동추론, 지식표현, 음성인식, 머신러닝, 컴퓨터비전, 로봇틱스 등
 - 이성적으로 생각하는 시스템
 - 이성적으로 행동하는 시스템

인공지능 검사

- 컴퓨터가 지능을 갖고 있음을 어떻게 판단?
- 1950년 Alan Turing이 다음의 테스트를 제안
- **Turing Test**
 - 독립된 공간에 컴퓨터, 사람, 심사원이 들어가서,
 - 설치된 컴퓨터 화면을 통해 문자를 주고받은 후,
 - 컴퓨터와 사람 둘 다에게 문자를 주고받은 심사원이 어떤 상대가 컴퓨터인지 구별하는 테스트

인공지능 뇌(HW)의 진화

- 1939년 제 2차 세계 대전의 발발
 - 대학교 중심으로 진행되던 개념적인 인공지능 알고리즘이 실제로 구현되는 플랫폼이 등장하게 된 계기
- 1941년 독일 Z3 컴퓨터 개발
- 1944년 영국 Colossus 컴퓨터 개발
 - 독일군 암호 해독
- 1945년 미국 ENIAC 컴퓨터 개발



- **1세대 컴퓨터**
 - 1946년부터 1959년 사이에 개발된 컴퓨터
 - **진공관**, 펀치카드, 종이/자기 테이프 사용
 - 발열, 전력 소모 등의 문제
 - Assembler로 프로그래밍
 - ENIAC, EDVAC, UNIVAC, IBM 650, IBM 701 등
- **2세대 컴퓨터**
 - 1959년부터 1965년 사이에 개발된 컴퓨터
 - **트랜지스터**, 자기테이프, 디스크 사용
 - Fortran, Cobol 등으로 프로그래밍
 - IBM 1620, IBM 7094, CDC 1604, CDC 3600, UNIVAC 1108 등

- **3세대 컴퓨터**

- 1965년부터 1971년 사이에 개발된 컴퓨터
- **집적회로**, 원격명령실행, 멀티프로그래밍 사용
- Pascal, Basic 등으로 프로그래밍
- IBM 360, Honeywell 6000 등

- **4세대 컴퓨터**

- 1971년부터 1980년 사이에 개발된 컴퓨터
- **고밀도 집적회로**, 마이크로프로세서 사용
- 컴퓨터의 대중화 시대 열림
- C, C++ 등으로 프로그래밍
- Cray X-MP, Y-MP, DEC 10, STAR 1000, PDP 11 등

- **5세대 컴퓨터**

- 1980년부터 지금까지 개발된 컴퓨터
- 실리콘 칩의 초고밀도 집적화
- 병렬컴퓨팅의 발전

- 초고집적 프로세서, 고성능 네트워크 장비, 대용량 메모리, 입출력 장치를 갖춘 컴퓨팅 환경이 완성됨

→ Deep Learning과 같은 기술이 실용화되는 단계에 이르게 됨

최초의 인공지능경망

- **Perceptron**

- 1958년 Frank Rosenblatt가 발표
- 인공지능경망을 실제로 구현한 최초의 모델



- 1956년 John McCarthy가 다트머스 대학에서 열린 컨퍼런스에서 "**인공지능**"이라는 말을 처음 사용
 - 1958년 MIT에서 인공지능 분야에서 사용되는 고급 프로그래밍 언어인 LISP 개발

인공지능의 첫 번째 겨울

- 1969년 Marvin Minsky와 Seymour Papert는 Perceptron 이론의 한계점을 수학적으로 증명함
 - Perceptron은 **AND**나 **OR** 같은 선형 분리가 가능한 문제에서는 적용가능하나, **XOR** 문제에는 적용할 수 없다는 내용
- 60년대 말, 70년대 초에 학계의 많은 사람들이 인공지능에 대한 관심을 잃게 됨
- 인공지능의 겨울 : 1974 ~ 1980

인공지능의 재도약

- 1980년대 산업계에서 **Expert System**의 활약
 - 사람이 보유한 전문적인 지식과 경험을 잘 정리해 적재적소에 가장 알맞은 정보를 제공하는 시스템
 - 비전문가도 전문가 수준으로 업무를 처리할 수 있게 함
- 전문가 시스템의 구성
 - 전문가의 지식과 경험을 사실과 규칙 형태로 저장한 지식 데이터베이스
 - 자문과 의사결정을 생성하는 추론엔진
 - IF – THEN 구조
 - 사용자와 시스템 간 입력/출력 인터페이스

- XCON (eXpert CONfigurer)
 - 1978년 카네기멜론 대학에서 DEC의 요청으로 개발한 전문가 시스템
 - 고객의 요구사항에 맞게 주문생산하기 위해
- 1980년대 말까지 미국 내 500대 기업의 절반이 전문가 시스템 개발

두 번째 인공지능의 겨울

- 1980년대 말부터 1990년대 초까지
- 전문가시스템은 성능 대비 가격 측면에서 문제
 - 새로운 정보의 업데이트 절차가 매우 까다롭고 복잡함
 - 성능이 뛰어난 개인용 컴퓨터의 보급이 시작됨

현재의 인공지능

- 최근의 핫 트렌드는 인공신경망 기반의 딥러닝과 로봇틱스
- 80년대 중반, 단일 퍼셉트론에서 해결 못한 비선형 문제는 **다층 신경망에서의 역전파 알고리즘**으로 해결
- 좀 더 복잡한 문제를 풀기 위해 신경망의 층수를 늘리다가 국부적은 최소값을 벗어나지 못하거나 수많은 반복 계산 발생 문제 직면
 - **경사감사소멸문제**: 신경망 층이 늘어나면서 출력층의 정보가 역전파되는 과정에서 판별력이 소멸되기 때문

- 90년대 **장단기 기억법(LSTM: Long Short-Term Memory)** 등장
 - 출력층의 정보를 계속 메모리에 기억해서 역전파에 이용하는 방법
 - 필기체 인식, 음성인식, 자연어처리에서 주요 알고리즘으로 활용
- 이미지 인식 분야
 - 80년대 말 역전파 기반의 **Convolutional 신경망**
 - 2006년 RBM(Restricted Boltzmann Machine) 기반의 **DBN(Deep Belief Network)**
 - 학습초반에 이미지의 특성을 효과적으로 분석하는 장점

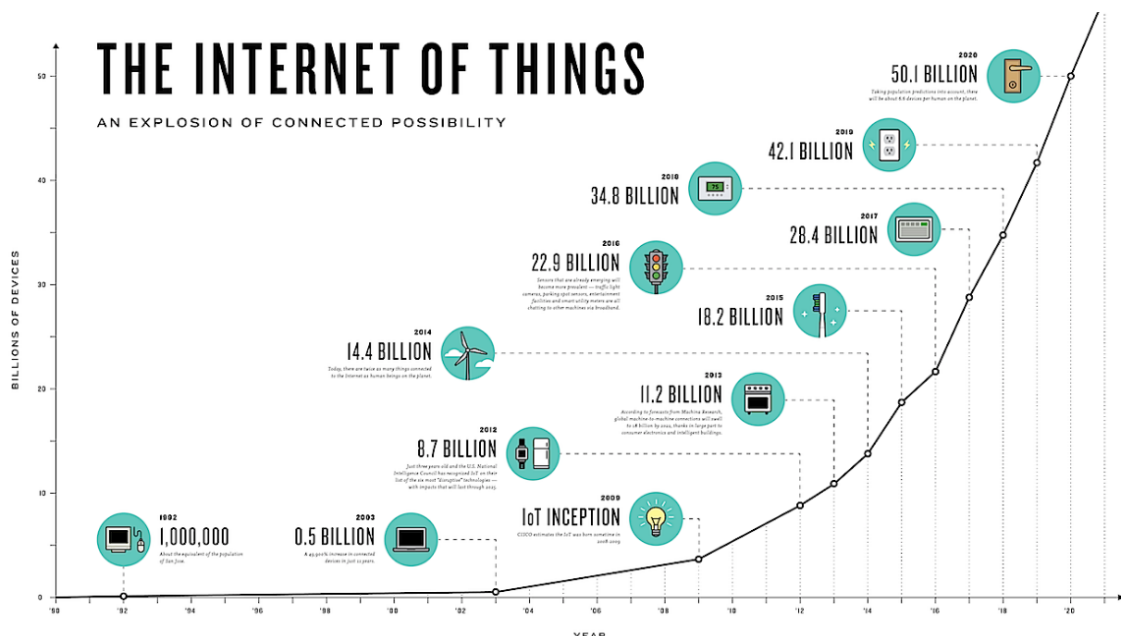
- **로보틱스**
 - 산업용 로봇
 - 공장 자동화
 - 전문서비스 로봇
 - 재난용, 물류운송용, 군사용
 - 개인서비스 로봇
 - 장애인 도우미, 가정용 로봇청소기, 개인비서, 완구류 등
- **Google**
 - 2013년부터 로보틱스를 구글의 새로운 성장 동력으로 선정하고 본격적인 투자
 - Schaft, Boston Dynamics 등 10여 개 이상의 로봇 전문회사 인수

IoT 시대의 인공지능

- 인간의 몸은 감각 세포를 통해 실시간으로 주변상황을 인지하고, 매 순간 적절한 반응 작용을 통해 신체를 가장 최적으로 유지함
 - 자동적인 반응, 의도적인 반응 등
- 인간의 몸을 IT 환경으로 비유하면,
 - 감각신경 → 센서
 - 생성되는 데이터는 너무 크고, 빠르고, 다양한 빅 데이터
 - 중추신경 → 네트워크
 - 근육 → 액추에이터(actuator)
 - 뇌 → 컴퓨터

• IoT (Internet of Things)

- 99년 영국의 Kevin Ashton이 처음 사용
- 2015년말 기준 약 64억 개의 IoT 기기 설치됨
- 2020년 약 500억 개의 센서가 설치될 것으로 예상



- IoT가 등장하게 된 배경

- RFID

- 바코드 대신 사물에 부착되어 전파를 송수신하여 사물의 고유번호와 필요 정보를 탐지하는 기술

- IPv6로의 전환

- 2^{128} 개의 인터넷 주소 사용 가능
 - 500억 개의 IoT 기기 충분히 사용 가능

- IoT의 구성

- 단말센서층

- IoT 단말기는 센서, CPU/메모리, 네트워크 장치로 구성
 - 배터리 내장
 - 기기 동작을 위한 매우 가벼운 OS 설치

- 통신망

- 여러 센서를 관리하는 **마스터 센서**
 - 마스터 센서를 관리하는 **로컬 서버**
 - 로컬 서버는 데이터를 중앙컴퓨터 또는 클라우드로 전송

- 응용층

- 센서에서 전달된 데이터 저장, 분석해서 적재/적소/적시에 의사결정을 내림
 - 로컬 서버 또는 클라우드에 존재
 - IoT의 핵심이며 인공지능이 결정적인 역할을 함

- 2020년이 되면 44ZB의 디지털 데이터가 생성된다고 예상
 - IoT에서 생성된 데이터가 80% 이상

1 kilobyte	1,000,000,000,000,000,000
1 megabyte	1,000,000,000,000,000,000,000
1 gigabyte	1,000,000,000,000,000,000,000,000
1 terabyte	1,000,000,000,000,000,000,000,000,000
1 petabyte	1,000,000,000,000,000,000,000,000,000,000
1 exabyte	1,000,000,000,000,000,000,000,000,000,000,000
1 zettabyte	1,000,000,000,000,000,000,000,000,000,000,000,000

- 1TB 용량의 노트북으로 44ZB 저장하려면 440억대 필요
- 머신러닝이나 딥러닝 같은 인공지능 기술이 IoT 분야에서 중요한 이유

인공지능의 미래

- 인간의 뇌를 모방한 딥러닝 기법이 광범위하게 인공지능 분야에 적용되면서 다양하고 파괴적인 성공 사례 보여줌
- 인공지능 분야의 미래학자인 Ray Kurzweil은 2045년 정도가 되면 인공지능이 제어할 수 없을 정도로 폭발적으로 발전할 것이라 예측
 - Recursive self-improvement 능력을 가진 초인공지능
 - 스스로 학습하는 인공지능

머신러닝 개요 (Machine Learning)

“인공지능, 머신러닝, 딥러닝 입문”, 김의중 저, 위키북스, 2016

Machine Learning

- **Machine**
 - 프로그래밍 가능한 컴퓨터를 의미
- 머신러닝의 3가지 접근법
 - 신경 모형 패러다임
 - Perceptron → Deep Learning
 - 심볼 개념 패러다임
 - 숫자나 통계이론 대신 논리학이나 그래프 구조를 사용
 - 현대지식 집약적 패러다임
 - 백지상태에서 학습을 시작하는 신경 모형을 지양하고, 이미 학습된 지식은 재활용하는 방식

머신러닝 정의

- 카네기멜론 대학의 Tom Mitchell 교수
 - “만약 컴퓨터 프로그램이 특정한 태스크 T를 수행할 때, P만큼 개선되는 경험 E를 보이면 그 컴퓨터 프로그램은 태스크 T와 성능 P에 대해 경험 E를 학습했다고 할 수 있다.”
 - 예) 필기체 인식의 경우
 - 태스크 T : 필기체를 인식하고 분류하는 것
 - 성능 P : 필기체를 정확히 구분한 확률
 - 경험 E : 필기체와 정확한 글자를 표시한 데이터세트

- 실무적인 관점에서의 러닝
 - **Learning = Representation + Evaluation + Optimization**
 - **표현 (Representation)**
 - 어떤 태스크를 수행하는 에이전트가 입력값을 처리해 어떻게 결과값을 만들지를 결정하는 방법
 - 예) SVM, 의사결정 트리, k-mean 모델 등
 - **평가 (Evaluation)**
 - 에이전트가 얼마만큼 태스크를 잘 수행했는지 판정하는 방법
 - 예) 최소제곱법 등
 - **최적화 (Optimization)**
 - 평가에서 설정한 기준을 최적으로 만족하는 조건을 찾는 방법
 - 예) 경사감소법 등
 - 최적화 과정이 끝나 학습 모델에 사용된 가중치가 결정되면 **학습이 완료됨**
 - 이후, 새로운 데이터에 대한 예측을 하는 것이 **일반화 (Generalization) 과정**

- **머신 러닝 vs 데이터 마이닝**

- **머신 러닝 (Machine Learning)**

- 컴퓨터 과학에서의 관점
 - 기존 데이터를 통해 학습을 시킨 후, 새로운 데이터에 대한 예측값을 알아내는 데 목적

- **데이터 마이닝 (Data Mining)**

- 통계학에서의 관점
 - 가지고 있는 데이터에서 현상 및 특성을 발견하는 것이 목적

머신러닝의 분류

- **Supervised Learning (지도학습)**

- 학습데이터에 레이블이 있는 경우

- Classification

- 결과값이 고정
 - KNN, Support Vector Machine, Decision Tree

- Prediction

- 결과값이 데이터세트의 범위 내 어떠한 값도 가능
 - Regression

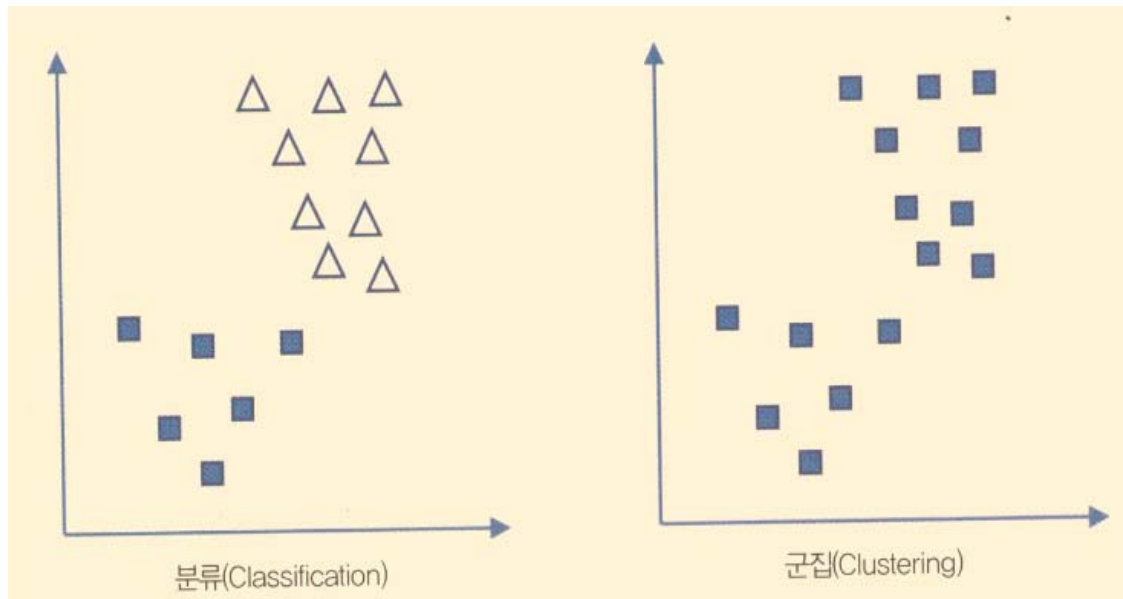
- **Unsupervised Learning (비지도학습)**

- 학습데이터에 레이블이 없는 경우

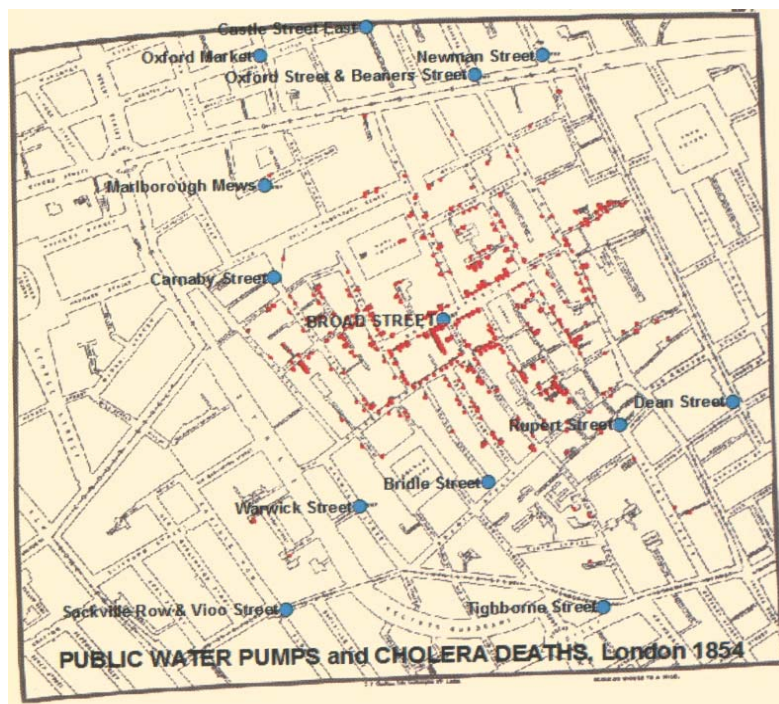
- Clustering

- 분할 기반 군집 모델: k-means, k-medoids, DBSCAN 등
 - 계층적 군집 모델

- 분류(Classification)와 군집(Clustering) 비교
 - 분류 모델에서는 레이블이 있으나,
 - 군집 모델에서는 레이블이 없음



- 1854년 영국 런던에서 콜레라가 극심했던 지역
- John Snow 박사가 최초로 역학 조사를 하여 원인 발견
- 머신러닝 기법 가운데 군집에 해당



최초의 머신러닝 모델

- 1936년 Ronald Fisher는 붓꽃의 4가지 특성으로 붓꽃을 구별하는 방법 제시
 - Setosa, Versicolor, Virginica 등 3가지 붓꽃 종류를 각각 50개씩 표본 추출
 - 꽃받침의 길이와 넓이, 꽃잎의 길이와 넓이를 기록
 - SVM, 의사결정 트리 등 다른 머신러닝 모델의 레퍼런스 데이터로 사용됨

<i>Iris setosa</i>				<i>Iris versicolor</i>				<i>Iris virginica</i>			
Sepal length	Sepal width	Petal length	Petal width	Sepal length	Sepal width	Petal length	Petal width	Sepal length	Sepal width	Petal length	Petal width
5.1	3.5	1.4	0.2	7.0	3.2	4.7	1.4	6.3	3.3	6.0	2.5
4.9	3.0	1.4	0.2	6.4	3.2	4.5	1.5	5.8	2.7	5.1	1.9
4.7	3.2	1.3	0.2	6.9	3.1	4.9	1.5	7.1	3.0	5.9	2.1
4.6	3.1	1.5	0.2	5.5	2.3	4.0	1.3	6.3	2.9	5.6	1.8
5.0	3.6	1.4	0.2	6.5	2.8	4.6	1.5	6.5	3.0	5.8	2.2
5.4	3.9	1.7	0.4	5.7	2.8	4.5	1.3	7.6	3.0	6.6	2.1

강화학습 (Reinforcement Learning)

- 지도학습 하나로 분류하거나, **3번째 머신러닝 모델로 분류**하기도 함
- 에이전트가 취한 모든 행동에 대해 환경으로부터 보상과 벌칙을 피드백 받아 학습
- 사람과 동물이 시행착오를 통해 학습하는 원리를 머신러닝에 적용한 것
- 사건이 전개되면서 받는 피드백을 통해 학습하기 때문에 순차적 사건에 대한 의사결정을 내릴 때 주로 사용
- 게임이나 로봇틱스에 가장 효과적으로 활용됨

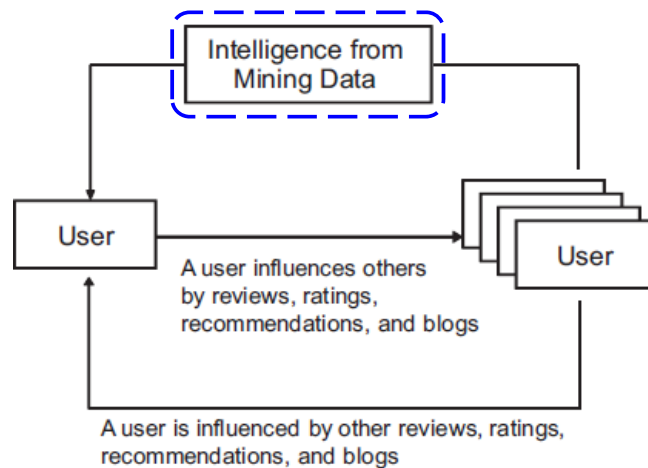
머신러닝에 필요한 사전 처리

- 데이터의 사전 처리 (Preprocessing)
 - 아래의 데이터를 계산 가능한 정량적인 단위로 변환하고 벡터나 행렬 형태로 저장
 - 텍스트로 입력된 레이블,
 - 자연어로 구성된 문장,
 - 음성 신호, 디지털 이미지 및 동영상 등의 입력 데이터
- 컴퓨터는 상기 데이터를 가지고 학습을 함
 - 즉, 판단 규칙을 만듦

Chapter 1. Introduction to Collective Intelligence

• 집단지성 (Collective Intelligence)

- 사용자들의 소통과 참여에 의해 만들어지는 지능 정보에서 추출되는 지능
- 사용자들이 진정으로 원하는 가치 있는 정보를 필터링하는 데에 사용
- 애플리케이션을 개선하기 위해, 사용자들이 제공한 정보를 효과적으로 이용하는 것

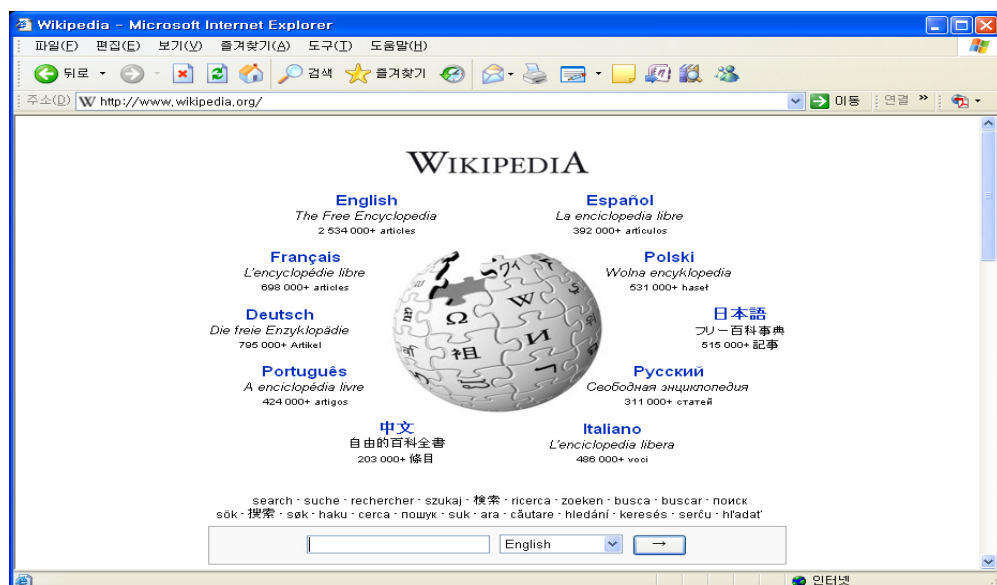


- 구글(Google)

- 웹 페이지 평가에 다른 페이지에서 얼마나 링크를 많이 받았는지를 적용한 첫 검색엔진

- 위키피디아(Wikipedia, <http://www.wikipedia.org>)

- 사용자의 공헌으로 만들어지는 온라인 백과사전



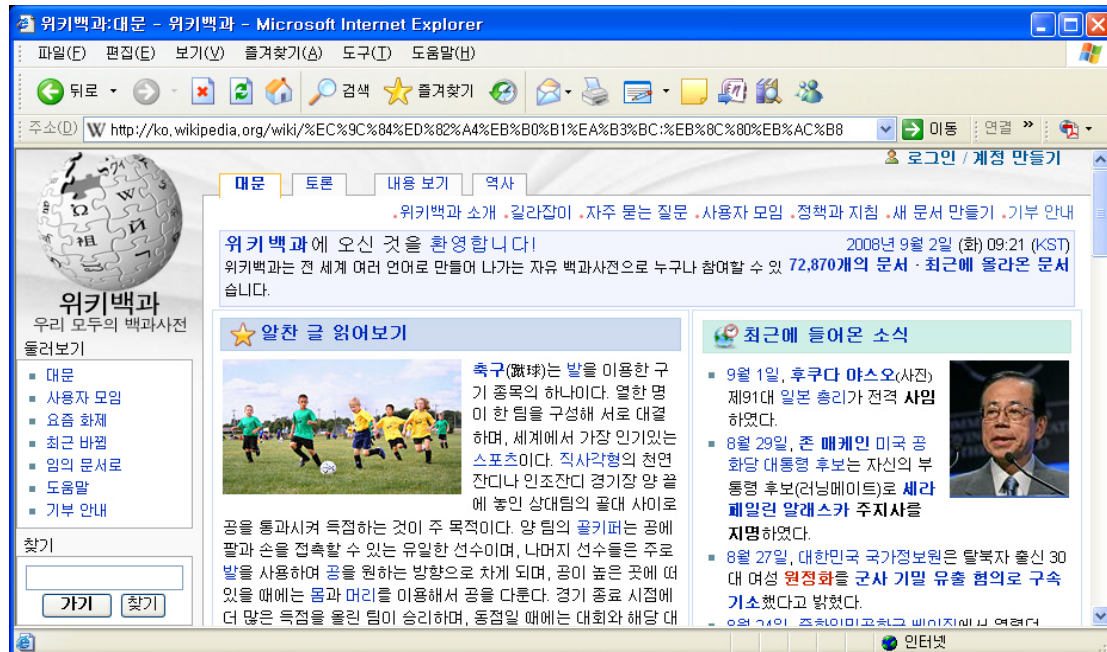
다른 언어로 된 위키백과: (모두 보기)

1,000,000 항목 이상: 영어(English)

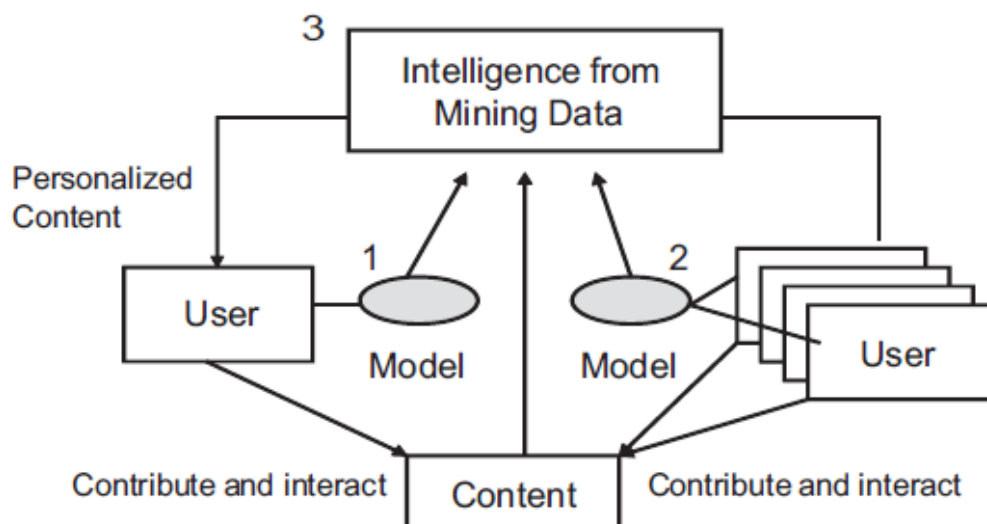
200,000 항목 이상: 네덜란드어(Nederlands) · 독일어(Deutsch) · 러시아어(Русский) · 스웨덴어(Svenska) · 스페인어(Español) · 이탈리아어(Italiano) · 일본어(日本語) · 중국어(中文) · 포르투갈어(Português) · 폴란드어(Polski) · 프랑스어(Français)

100,000 항목 이상: 노르웨이어(Norsk bokmål) · 루마니아어(Română) · 불가리아어(Български) · 슬로바키아어(Slovenčina) · 에스페란토(Esperanto) · 우크라이나어(Українська) · 체코어(Česky) · 카탈루냐어(Català) · 터키어(Türkçe) · 핀란드어(Suomi) · 헝가리어(Magyar)

50,000 항목 이상: 덴마크어(Dansk) · 리투아니아어(Lietuvių) · 베트남어(Tiếng Việt) · 불가리아어(Български) · 세르비아어(Српски / Srpski) · 슬로베니아어(Slovenščina) · 아랍어(العربية) · 마이티어(Kreyòl ayisyen) · 에스토니아어(Eesti) · 인도네시아어(Bahasa Indonesia) · 히브리어(עברית)



- 웹 애플리케이션에 집단지성을 적용하려면,
 1. 사용자들이 소통하게 하라.
 2. 사용자들에게 배워라.
 3. 소통 데이터와 수집된 데이터를 기반으로 개인화 콘텐츠를 제공하라.



- **집단지성의 이점**

- 높은 유지율

- 소통 ↑, 흥미 ↑, 재방문 가능성 ↑

- 사용자가 관심 있는 정보를 찾고 거래할 가능성 ↑

- 검색 엔진 순위 상승

- 사용자 참여 ↑, 콘텐츠 ↑, 다른 사용자의 방문 가능성 ↑

- **웹 2.0 애플리케이션의 7가지 원칙**

- 1. 네트워크가 플랫폼이다.

- 브라우저에서 서비스를 사용하기만 하면 됨

- 소프트웨어 라이선스 및 업그레이드 불필요

- 2. **집단지성을 이용한다.**

- 개인화된 서비스 통하여 보다 알맞은 콘텐츠 제공

- 3. 데이터를 복제하기 힘들다.

- 방대한 데이터 집합

- 4. 평생 베타 버전

- 릴리스 주기를 짧게 하여 사용자를 일찍 끌어들이м

- 5. 간단한 프로그래밍 모델

- 보다 넓은 활용성과 재사용성 위해

- 6. 다양한 환경에서 작동하는 소프트웨어

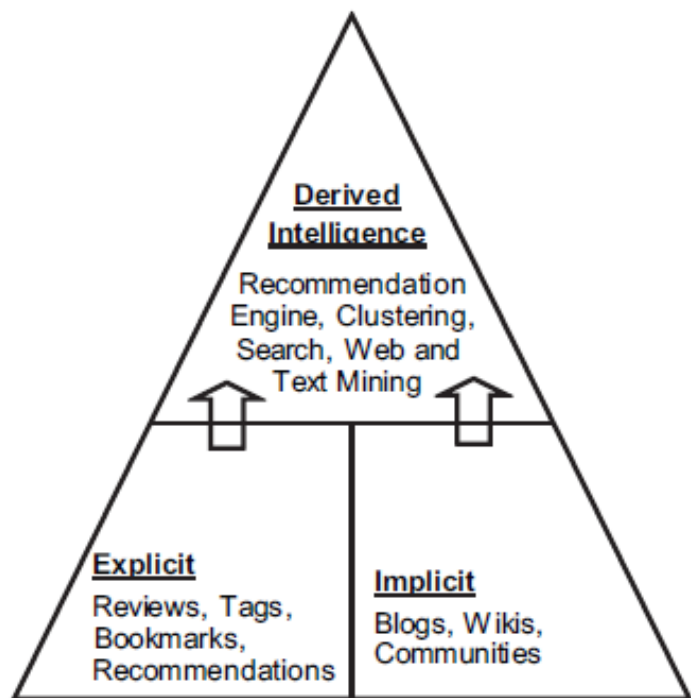
- 7. 풍부하고 다양한 사용한 인터페이스

- AJAX, Flash/Flex, Silverlight, ...

• 지능정보*의 분류

- 명시적 지능 정보
- 암묵적 지능 정보
- 추출된 지능 정보

* 정제된 정보



• 명시적 지능 정보

- 리뷰와 추천
- 태깅
 - 태그 클라우드
- 투표

- **암묵적 지능 정보**

- 애플리케이션에 관한 정보가 비 구조화된 형태로 존재
- 블로그, 위키, ...

- **추출된 지능 정보**

- 데이터 마이닝 & 텍스트 마이닝
- 클러스터링과 예측 분석
 - 데이터 마이닝의 2가지 주요 요소
- 지능형 검색
- 추천 엔진

- 추천 예1: Netflix (<http://www.netflix.com>)

- 온라인 DVD 대여 회사
- 2006년 말부터 추천시스템 정확도 10% 향상시키는 사람에게 100만 달러 상금
- 매년 우승자에게 5만 달러 상금



- 추천 예2: Amazon.com

http://www.amazon.com/ - Windows Internet Explorer

http://www.amazon.com/

즐거찾기

Frequently Bought Together



Price For All Three: \$79.81

[Add all three to Cart](#) [Add all three to Wish List](#)

Some of these items ship sooner than the others. [Show details](#)

- ☒ **This item:** Collective Intelligence in Action by Satnam Alag Paperback **\$29.69**
- ☒ Algorithms of the Intelligent Web by Haralambos Marmanis Paperback **\$23.73**
- ☒ Programming Collective Intelligence: Building Smart Web 2.0 Applications by Toby Segaran Paperback **\$26.39**

Customers Who Bought This Item Also Bought



Algorithms of the Intelligent Web by Haralambos Marmanis
★★★★☆ (12)
\$23.73



Programming Collective Intelligence: Building Smart Web 2.0 Applications by Toby Segaran
★★★★★ (69)
\$26.39



Mining the Social Web: Analyzing Data from... by Matthew A. Russell
★★★★★ (13)
\$26.39