

# Machine Learning Fueled by Open Source

안명호

[james@deepnumbers.com](mailto:james@deepnumbers.com)

[www.DeepNumbers.com](http://www.DeepNumbers.com)

# 발표자 안명호는



- KAIST SW석사과정을 마쳤다.
- 어느 날 알게 된 머신러닝에 흠뻑 빠져 그동안 애지중지하던 클라우드를 버리고 머신러닝으로 전향하였다. 이제 더는 다른 기술은 관심을 두지 않고 머신러닝 한길만으로 정했기에 머신러닝을 공부하며 어려운 수식들을 다시 보느라 고생하고 있지만, 하루하루 배워가는 지식에 행복해하며 지내고 있다. 머신러닝으로 가마우지를 만들어 인생을 즐기려 노력하고 있으며, 그 결실이 완성되는 날 완벽한 경제적 자유를 누리하고자 한다.
- Facebook : <https://www.facebook.com/james.ahn.9>
- Homepage : <http://www.deepnumbers.com>

# AlphaGO!!!



# IT에서 DT(Data Technology) 시대로



- 마윈 알리바바 회장은 지난주 말 베이징에서 열린 빅데이터산업 설명회에서 “알리바바는 빅데이터의 최대 수혜자로 지금은 정보화시대(IT)에서 데이터시대(DT)로 급속히 이양하고 있다”고 말했다.
- 마 회장은 “데이터를 활용해 돈을 버는 일이 미래의 핵심가치가 될 것”이라며 “비용을 줄여 이윤 내는 사업방식은 앞으로 잘 되지도 않을 뿐더러 성장성도 없을 것”이라고 내다봤다

# 2016 다보스포럼, 제 4차 산업혁명을 논하다.

## 눈앞에 온 제4차 산업혁명

혁명	연도	정보
	1 1784	증기, 물, 기계 생산 설비
	2 1870	노동분화, 전기, 대량생산
	3 1969	전자, IT, 자동생산
	4 ?	사이버-물리적 시스템

[자료: WEF]

- 인공지능이나 머신러닝 같은 신기술들이 촉발할 제4차 산업혁명으로 인해 생산, 경영, 산업 지배구조를 완전히 바꿔놓게 된다는 것이다.
- 물론 그 바탕이 되는 것은 무선 통신이다. 이를 통해 전 세계인들이 연결되면서 직업의 기본 개념과 작동 방식 자체가 획기적으로 바뀌게 된다는 것이다.

# 구글의 에릭 슈미츠, 머신러닝의 거대한 미래를 얘기하다.



- 머신러닝을 컴퓨팅 분야에 광범위하게 적용하면 웹에서부터 스마트폰 앱에 이르기까지 엄청난 전환이 예상되며 우버나 스냅챗과 같은 업체의 성공을 가져올 것이다
- 이는 다가오는 미래에 거대한 새로운 플랫폼, 기업, IPOs, 부, 그리고 엄청난 가능성의 출현을 의미한다

# 구글, 머신러닝 텐서플로우 오픈소스로 공개



- 2015년 11월 10일 공개
- Deepmind 알파고에 활용
- 구글 클라우드에서 활용가능

# 어쩌지, 내 일자리에 인공지능이!

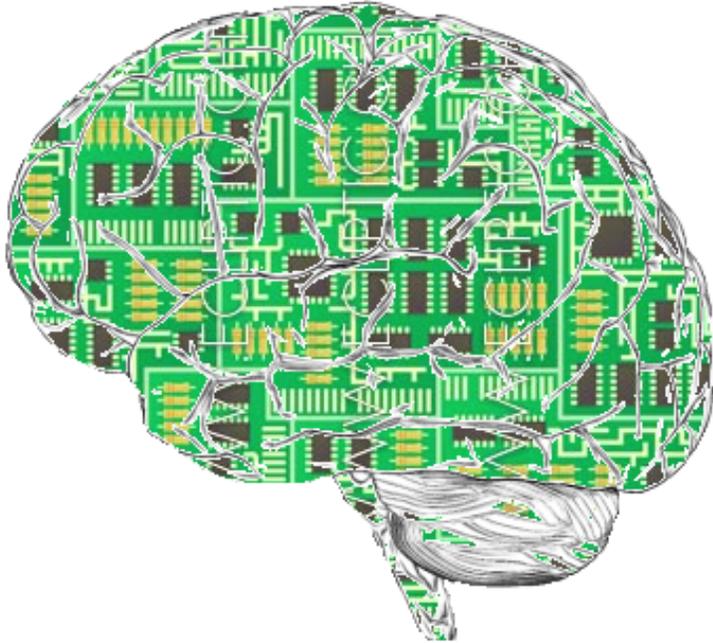
직종별 미래 일자리 증감 (단위=천개)



- 영국 BBC는 AI 로 인해 사라질 직업을 예측한 '직업별 자동화 대체율' 추정치를 발표했다.
- 단순 서비스나 정보 처리 직종이 가장 먼저 없어짐. 텔레마케터, 법률비서, 우체국 직원 등 단순 직종은 위험한 상황.
- 전문직이라고 안전한 것도 아니다. 회계사나 세무사 등의 대체 가능성도 높음

**인공지능이란 무엇인가?**

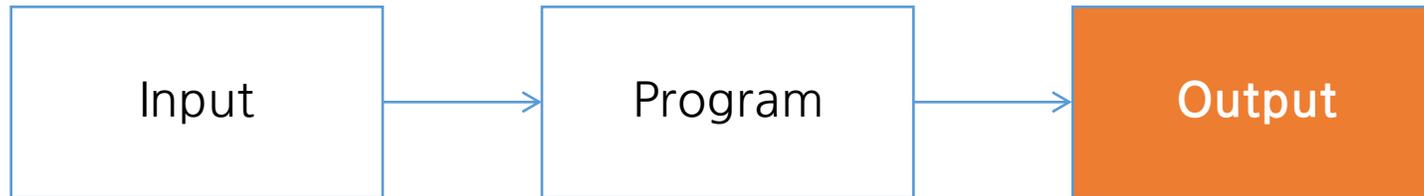
# 인공지능의 정의



- 인공지능( [영어](#): artificial intelligence, AI)은 철학적으로 인간성이나 지성을 갖춘 존재, 혹은 시스템에 의해 만들어진 지능, 즉 인공적인 지능을 뜻한다.
- 지능에 대한 명확한 과학적 정의는 아직 없다.
- Machine Learning이라고도 함

# 인공지능의 핵심개념

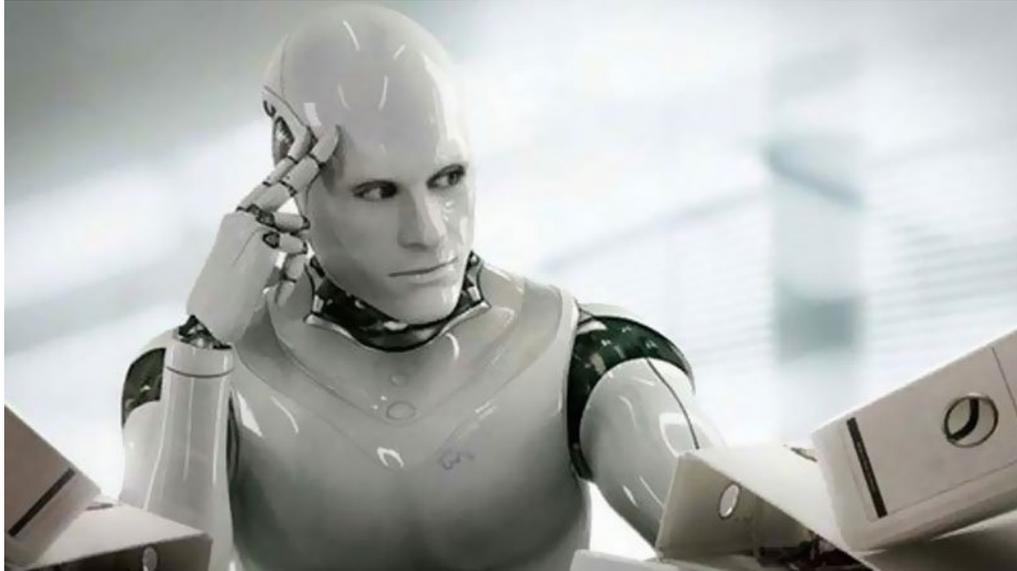
## Traditional Approach



## Machine Learning Approach

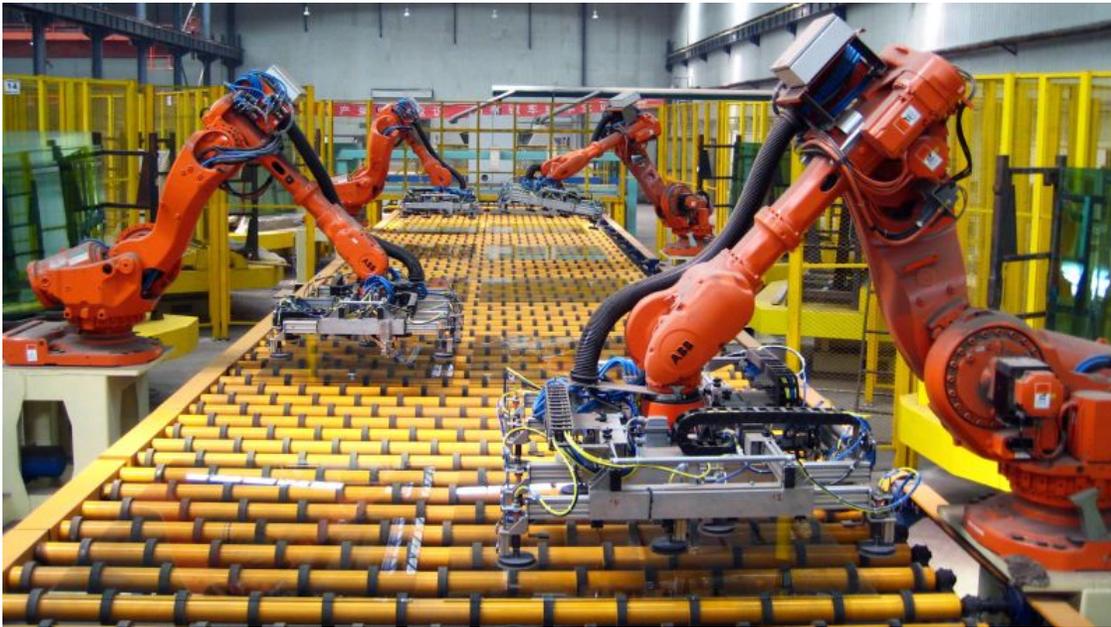


# 철학적 관점 - 강한 인공지능 (Strong AI)



- 강한 인공지능은 어떤 문제를 실제로 사고하고 해결할 수 있는 컴퓨터 기반의 인공적인 지능을 만들어 내는 것에 관한 연구다. 즉, 인공지능의 강한 형태는, 지각력이 있고, 스스로를 인식하는 것이라고 말할 수 있다.
- 1) 인간의 사고와 같이 [컴퓨터 프로그램](#)이 행동하고 사고하는 인간형 인공지능.
- 2) 인간과 다른 형태의 지각과 사고 추론을 발전시키는 컴퓨터 프로그램인 비인간형 인공지능.

# 철학적관점 - 약한 인공지능 (Weak AI)

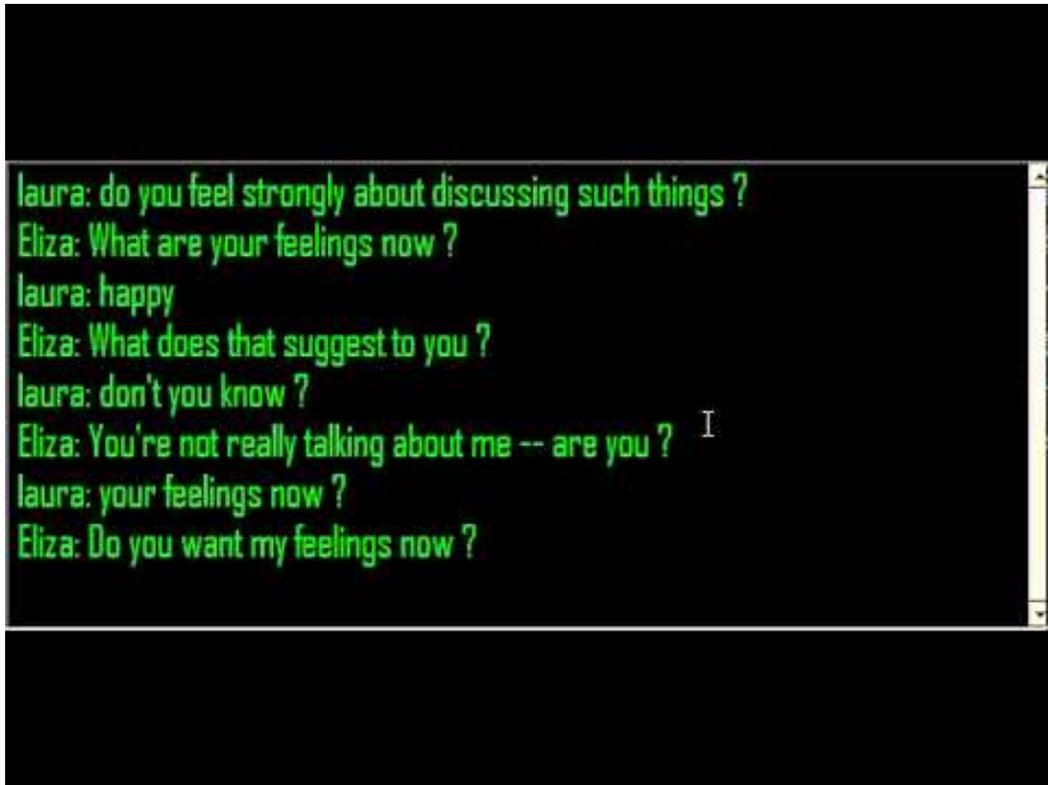


- 약한 인공지능(weak AI)은 어떤 문제를 실제로 사고하거나 해결할 수는 없는 컴퓨터 기반의 인공적인 지능을 만들어 내는 것에 관한 연구다.
- 이와 같은 시스템은 진짜 지능이나 지성을 갖추고 있지는 못하지만, 어떤 면에서 보면 지능적인 행동을 보일 것이다.
- 오늘날 이 분야의 연구는 주로 미리 정의된 규칙의 모음을 이용해서 지능을 흉내내는 컴퓨터 프로그램을 개발하는 것에 맞추어져 있다.

# AI와 연관된 영역들

<b>Philosophy</b>	Logic, methods of reasoning, mind as physical system, foundations of learning, language, rationality.
<b>Mathematics</b>	Formal representation and proof, algorithms, computation, (un)decidability, (in)tractability
<b>Statistics</b>	Modeling uncertainty, learning from data
<b>Economics</b>	Utility, decision theory, rational economic agents
<b>Neuroscience</b>	Neurons as information processing units
<b>Psychology / Neuro Science</b>	How do people behave, perceive, process cognitive information, represent knowledge
<b>Computer Engineering</b>	Building fast computers
<b>Control Theory</b>	Design systems that maximize an objective function over time
<b>Linguistics</b>	Knowledge representation, grammars

# AI 역사 - Golden Years(1956-1974)



A screenshot of a text-based chat interface with a black background and green text. The conversation is between a user named Laura and a program named Eliza. The text is as follows:

laura: do you feel strongly about discussing such things ?  
Eliza: What are your feelings now ?  
laura: happy  
Eliza: What does that suggest to you ?  
laura: don't you know ?  
Eliza: You're not really talking about me -- are you ?  
laura: your feelings now ?  
Eliza: Do you want my feelings now ?

- 1960
  - 인공지능 연구센터에 많은 연구자금이 투입
  - Intelligent Automata 개발
- 쓸만한 AI 프로그램의 탄생
  - Samuel's checker program
  - Newell and simon's logic theorist
  - Gelernter's geometry engine
  - Robinson's complete algorithm for logical reasoning
- AI를 위한 언어탄생
  - McCarthy-Lisp(1958)
- 첫번째 chat bot 개발
  - 인간과 비교적 유사한 대화가 가능한 Eliza(1966) 개발

# AI 역사 - The Winter

- Golden Years에 수행했던 연구들에 대한 회의감
- 발견되었던 문제들에 대한 해답을 찾지 못함
  - Limited Computer Power
  - Combinational Explosion(Exponential time)
  - Commonsense Knowledge and reasoning
  - Moravec's paradox
  - The chinese room argument undermined the goal of building intelligent machines
- 연구자금 중단됨

# AI 역사 - Rise of Expert System

- Mycin(1972)
  - Diagnosed infection blood diseases
  - Diagnosed infection blood diseases.
  - It had a set of about 600 rules and started to ask questions.
  - In some cases better than human experts In some cases, better than human experts.
- XCON (1980):
  - Production-rule-based system that assisted the ordering type of computers by automatically selecting the computer system components based on customer requirements
  - Saving \$40 million dollars
  - 2500 rules and processed 80000 orders with 95%-98% accuracy.
- New Funding!!!
  - 1986 sales of AI-based HW/SW were \$425 million

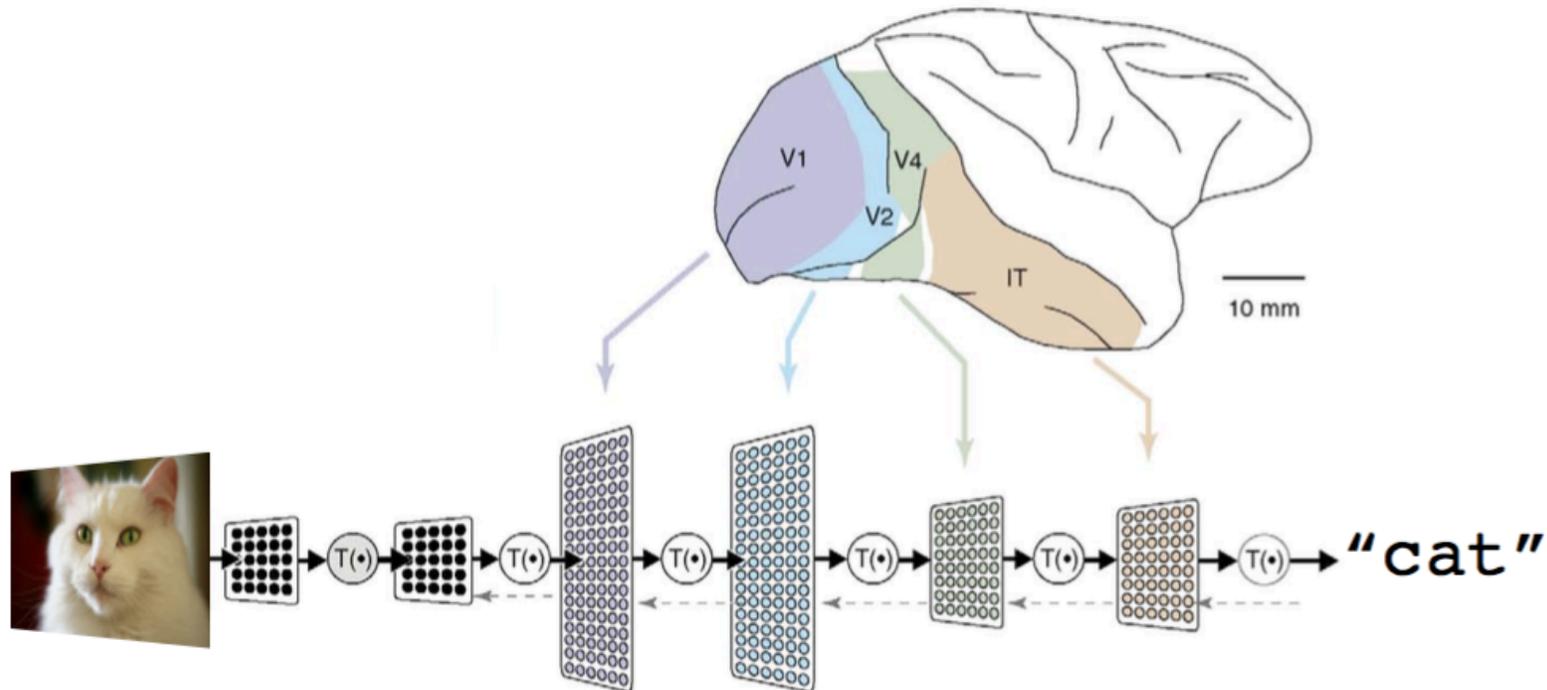
# AI 역사 - Quick Pace



- 90's
  - NCSA releases the first web browser, Mosaic
  - Deep Blue beats Gary Kasparov
  - Robotic soccer players in RoboCup
  - Sony corporation introduced the robotic dog AIBO
- 00's
  - iRobot introduces the vacuum cleaning robot Roomba
  - DARPA grand challenge
  - Honda ASIMO

# AI 역사 - Deep Learning

- 기존의 인공신경망(Artificial Neural Network)을 개량한 기술
- 간단한 수학적 모델을 여러개 활용해 복잡한 모델을 스스로 만들수 있는 방법
- 인간의 두뇌구조에 착안하였음

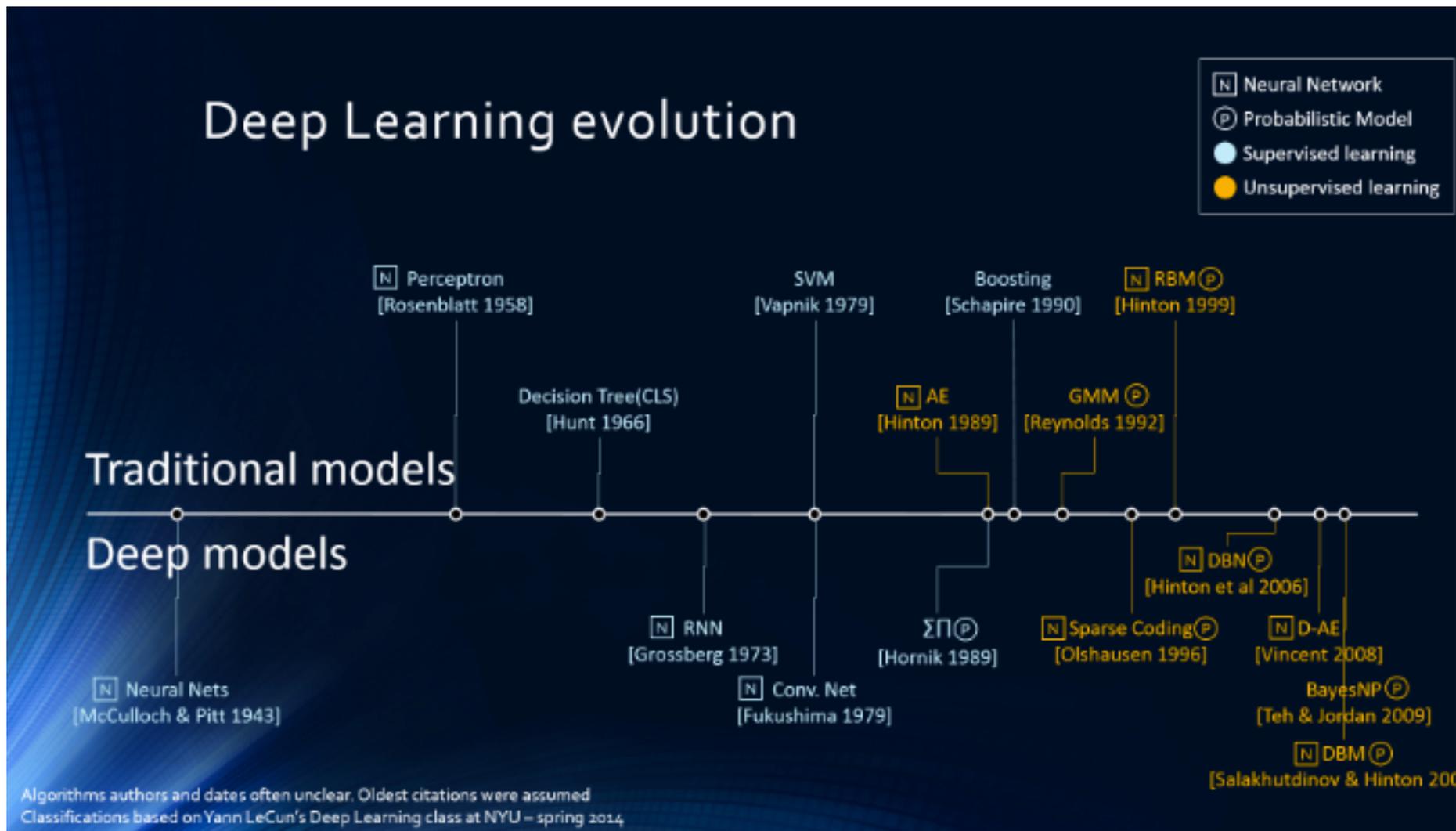


# AI 역사 - Deep Learning



- Imagenet Challenge 2015
  - Microsoft : 3.5%
  - Human : 5.1%
- 수많은 경진대회에서 Deep Learning이 우승을 하며 화려하게 재등장함

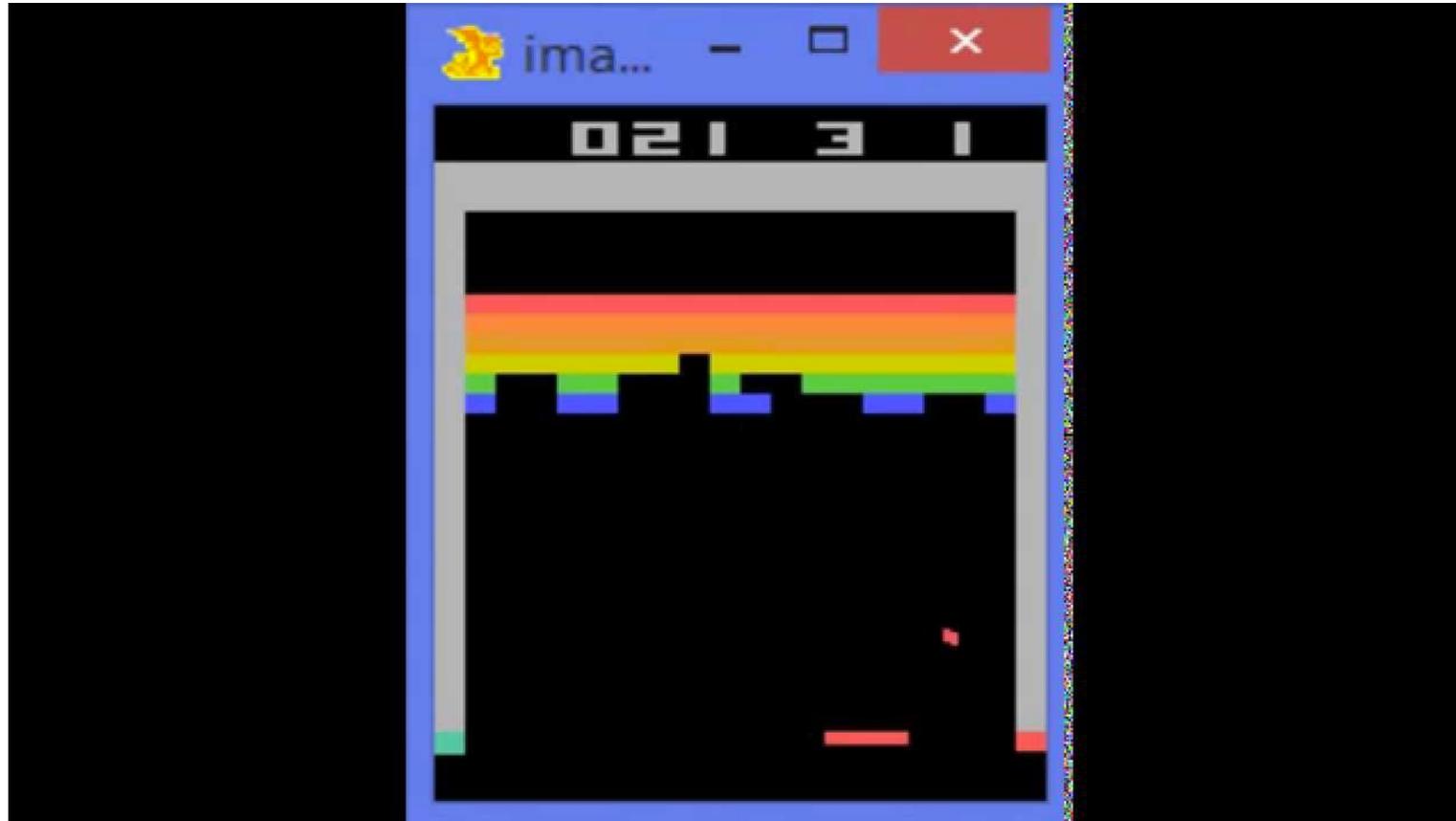
# AI 역사 - Deep Learning 발전사



**현 시점에서**

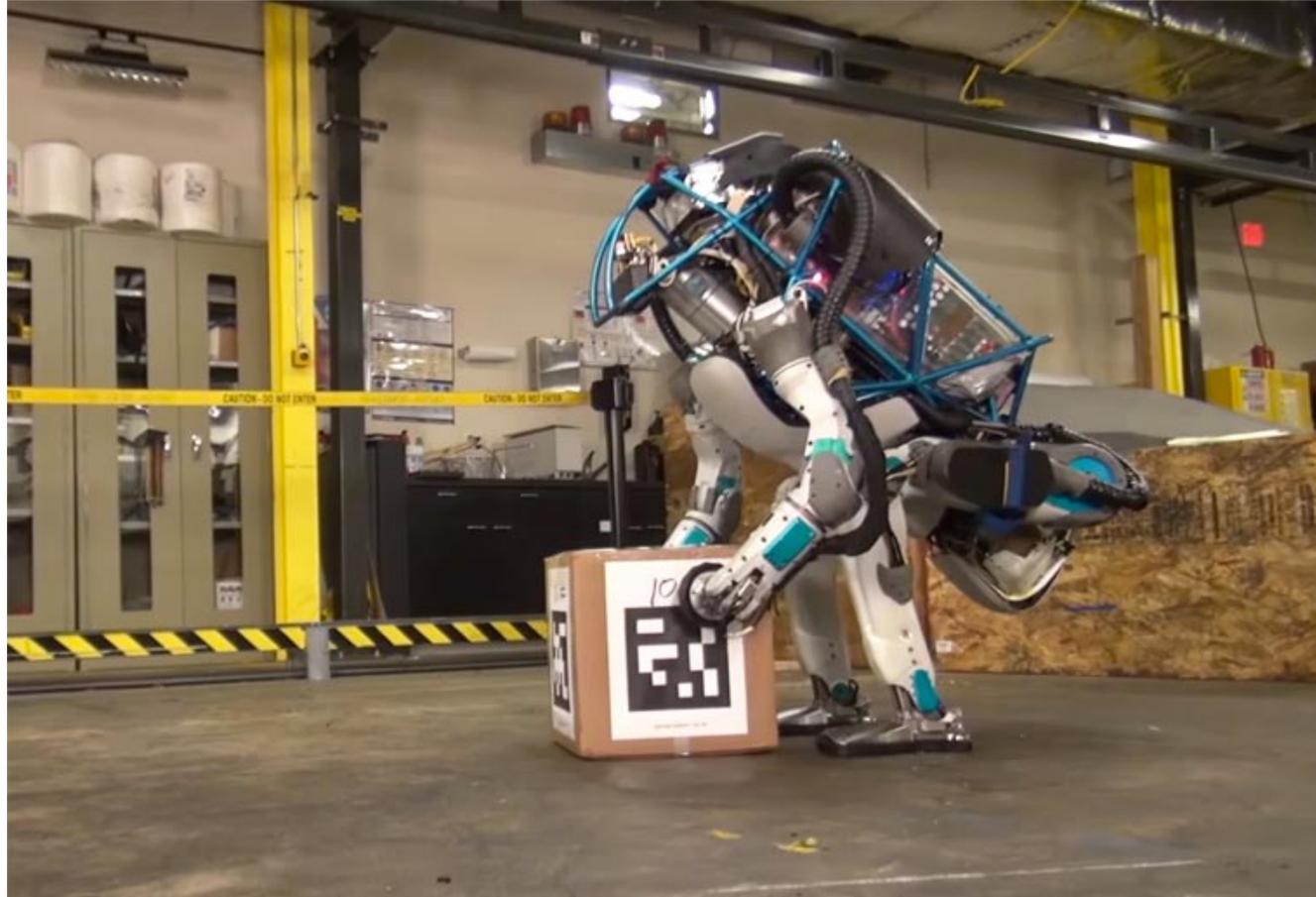
**인공지능은 무엇을 할 수 있는가?**

# Deepmind - Atari



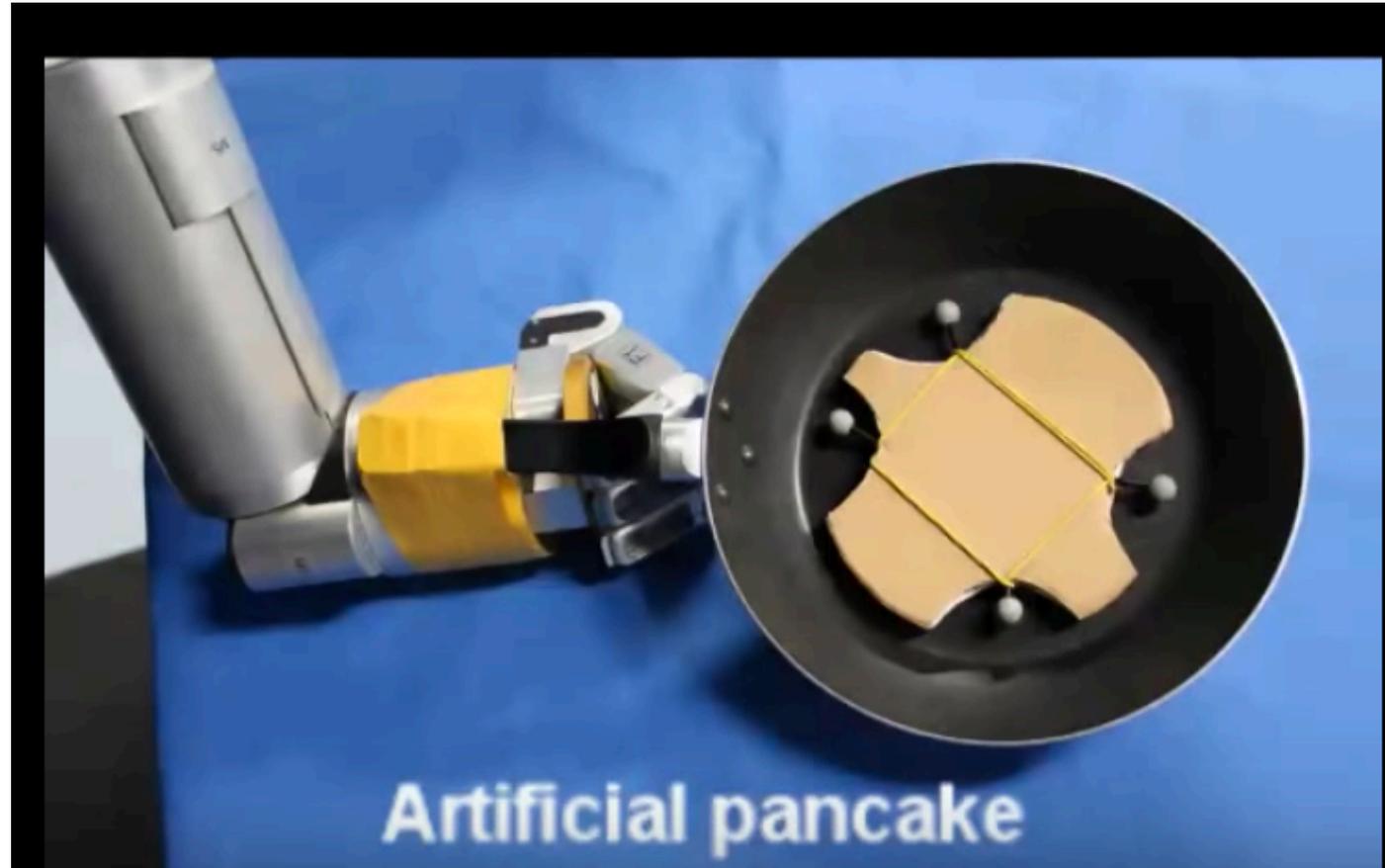
<https://www.youtube.com/watch?v=V1eYniJ0Rnk>

# Boston Dynamics – Atlas



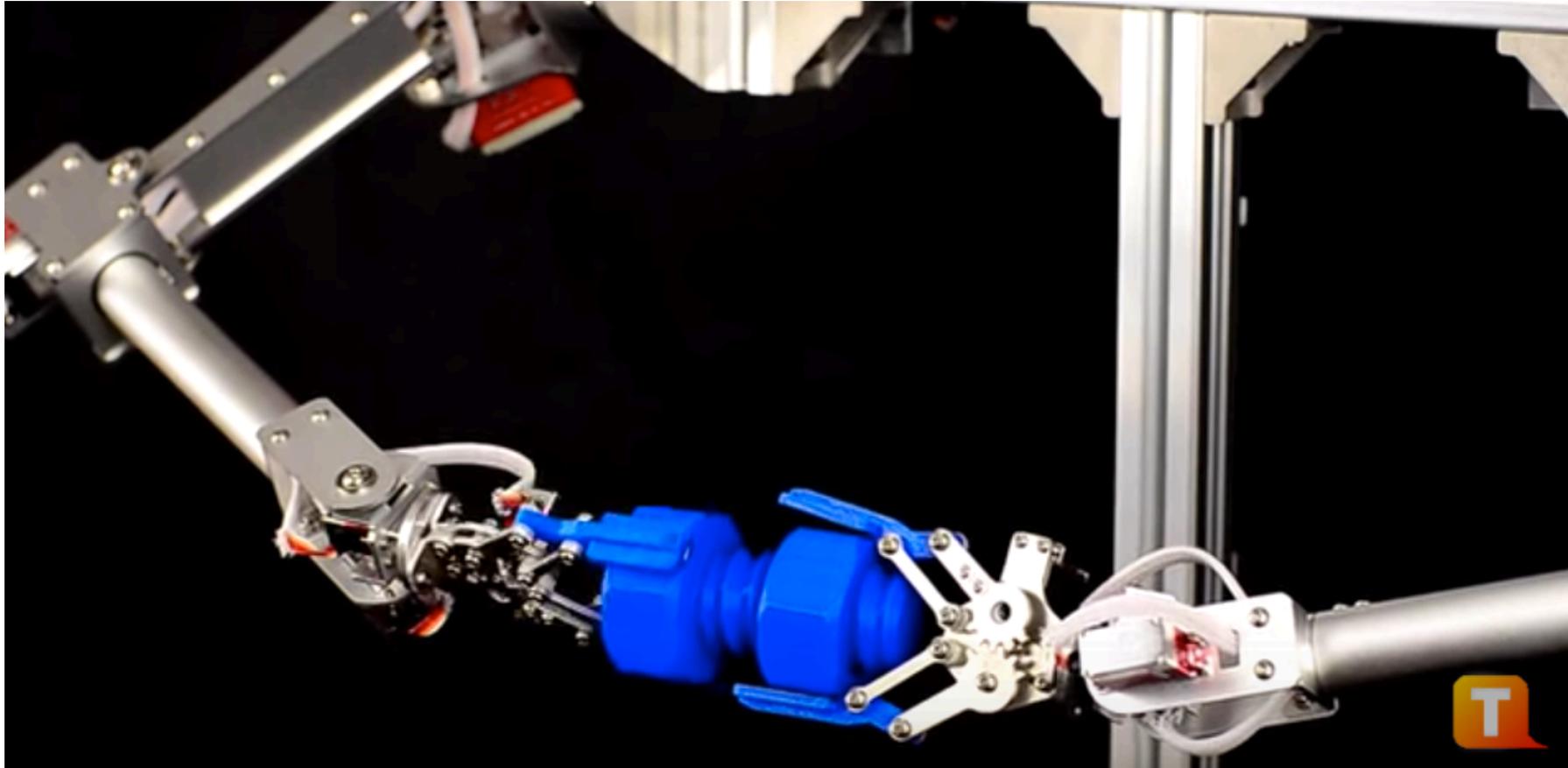
<https://www.youtube.com/watch?v=rVlhMGQgDkY>

# Reinforcement Learning – Pancake



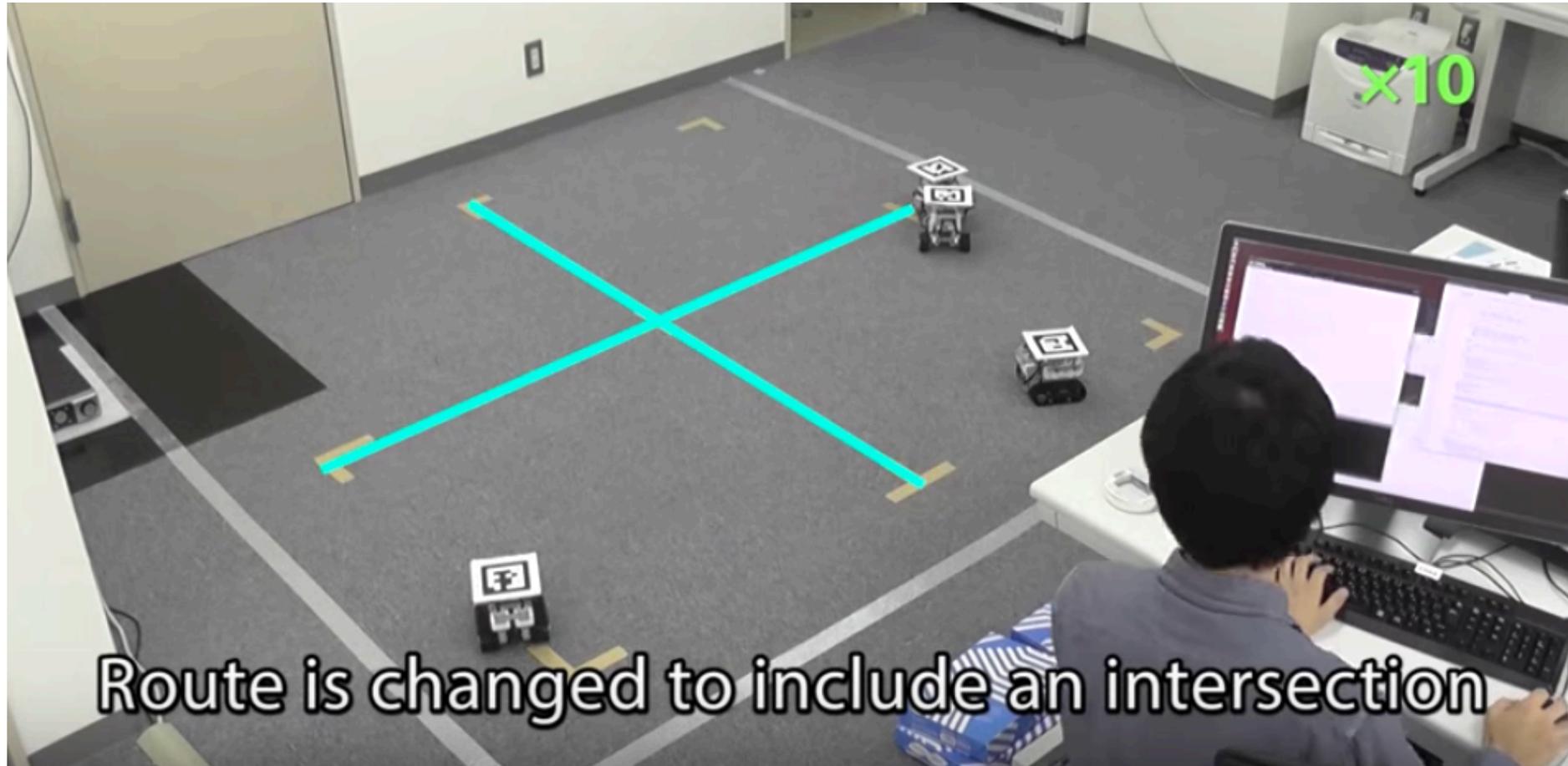
[https://www.youtube.com/watch?v=W\\_gxLKsSsIE&list=PL5nBAYUyJTrM48dViiby68urttMIUv7e](https://www.youtube.com/watch?v=W_gxLKsSsIE&list=PL5nBAYUyJTrM48dViiby68urttMIUv7e)

# 7Bot: \$300 Robot arm that can see, think and learn



<https://www.youtube.com/watch?v=KLP7vXVvp3w>

# Robot control with DRL



Route is changed to include an intersection

<https://www.youtube.com/watch?v=-YMfJLFynmA>

# 티센크루프 엘리베이터



<https://www.youtube.com/watch?v=nK1p-xhDilo>

# Amazon

BEAUTIFUL THINGS ON AMAZON UPDATED DAILY **EXPLORE**

**amazon** Try Prime

Prime Video Instant streaming of movies and TV shows **Get started**

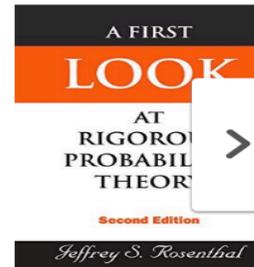
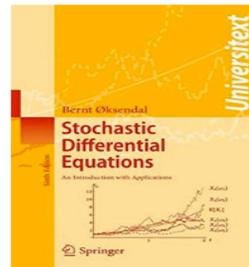
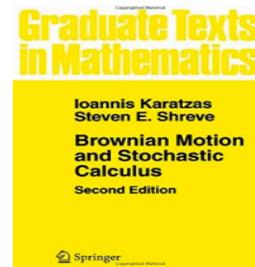
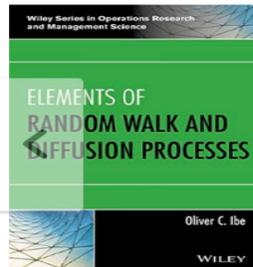
Departments Browsing History james's Amazon.com Today's Deals Hello, james **Your Account** Try Prime Lists

**Prime Video monthly membership now available**

Learn more

**Hi, james** On Order **0 items** Amazon Prime **Try Prime** Try Audible **Get 2 free audiobooks** Customer Since **2001**

Inspired by your browsing history [See more](#)



**amazon launchpad**

Today's brightest startups, now on Amazon

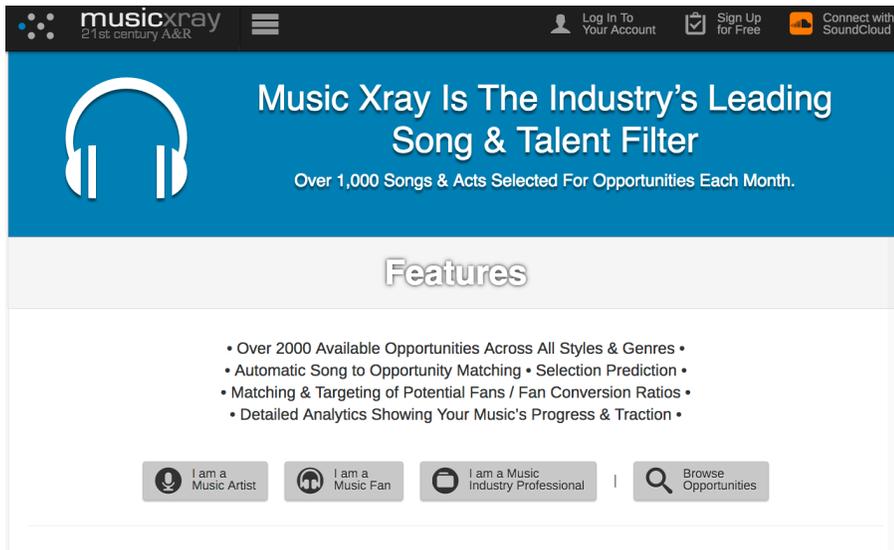
**> Shop now**

Ad feed

# Netflix



# Music XRay



- 인공지능을 이용한 음악 인기도 평가 서비스 제공
- 2011년 기준으로 1500개의 음반사와 기타 음악업자들이 서비스 활용중
- 컬럼비아, 타임워너, 게펜, EMI등 메이저 음반사들이 서비스 사용
- Come away with me, Norah Jones
  - 2000만장이상 판매

# Emmy



- David Cope가 개발한 클래식 음악 작곡 인공지능 프로그램
- “Classical Music Composed by Computer” 앨범발매
- 1997년 일반 청중을 대상으로 Emmy, Bach 그리고 라슨이라는 저명한 작곡가가 만든 곡을 바탕으로 테스트한 결과 Emmy가 우승함

# 캘리포니아 대학 샌프란시스코 캠퍼스 약국



<https://www.youtube.com/watch?v=oumIYbwfAsI>

- 2011년부터 로봇약국 운영
- 병원과 약국사이의 정보를 직접 수신해 약을 조제해 환자에게 제공
- 현재까지 200만개의 처방약을 조제하였고, 에러는 없음
- 미국약국의 에러율 4-10%

# eLoyalty



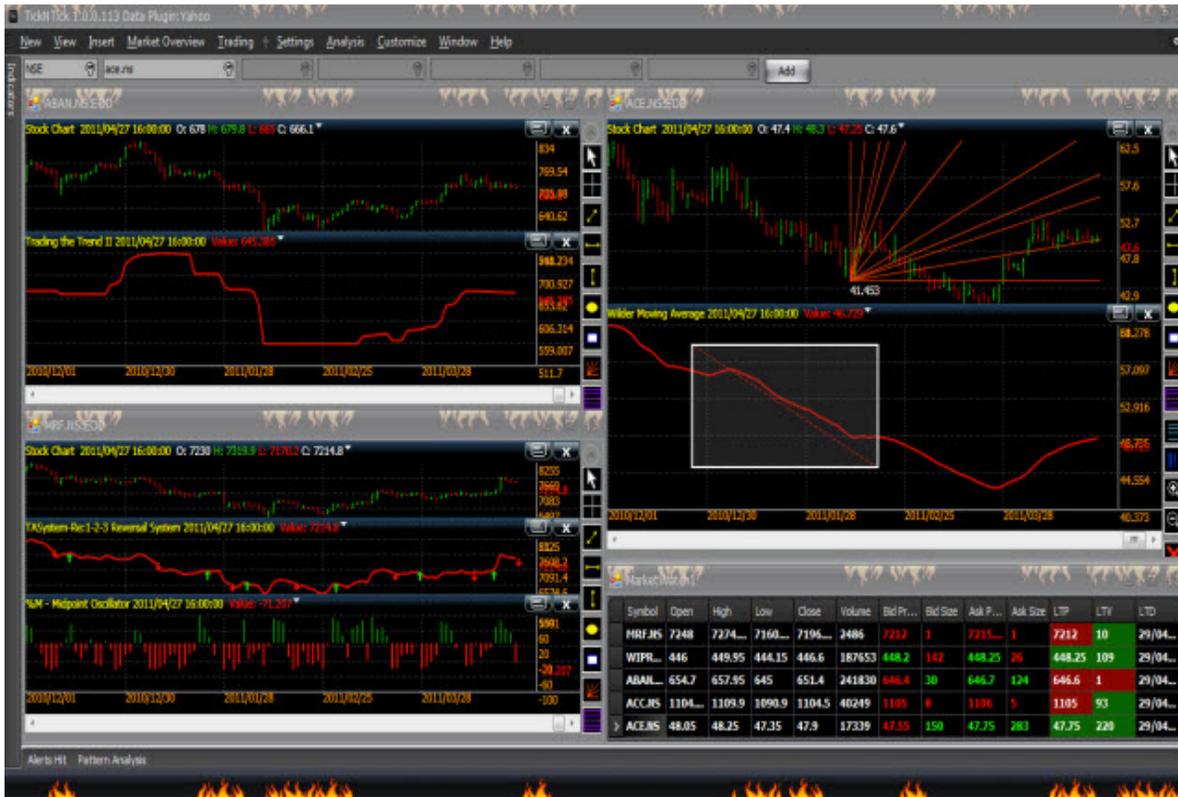
- 대화내용을 분석해 사람의 성향을 파악해 적절한 응대를 할 수 있도록 도와주는 인공지능 프로그램개발
- 콜센터에 적용함
- AT&T, 보다폰등등이 고객사
- 서비스사용후 콜센터 통화시간 5분단축, 문제 해결율 92%로 기존보다 2배이상 좋아짐
- AT&T는 콜센터비용 1/3 절감

# Epagogix



- 영화각본을 분석해 영화흥행여부를 예측할 수 있는 인공지능 프로그램 개발
- 2004년 개봉영화 9개에 대한 흥행도 예측 시험결과 6건에 대해 놀라울 정도로 정확히 예측
  - 영화 A: 실제수입 4000만달러, 예측수입 4900만달러
  - 영화 B: 120만달러 차이
- 20세기폭스사, 파라마운트등 메이저 영화사들이 사용중

# Wallstreet - Algorithm Trading

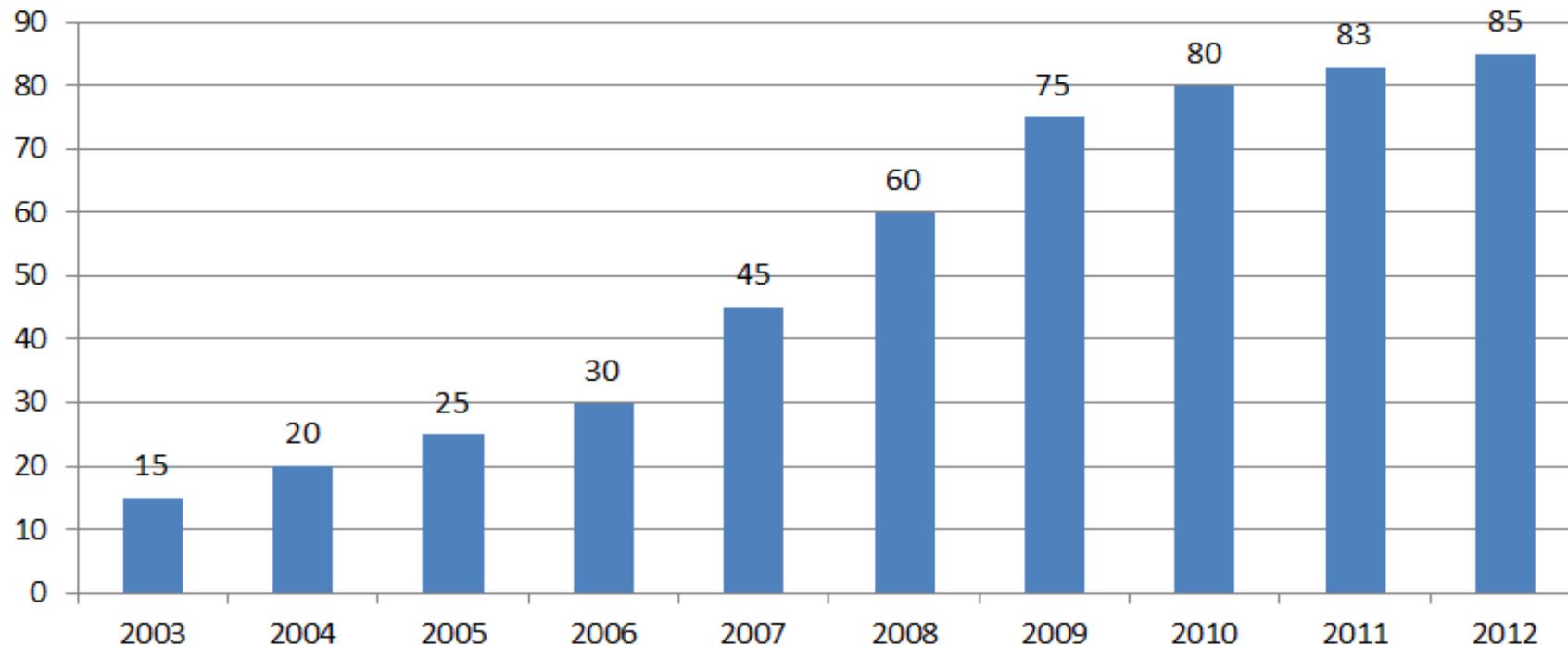


- 알고리즘 트레이딩은 수학적 계산과 IT 시스템을 이용해 트레이딩을 하는 것으로 시스템 트레이딩, Algo Trading 혹은 Blackbox 트레이딩이라고 한다.
- 알고리즘 트레이딩은 투자은행, 연기금, 헤지펀드, 증권회사등 많은 곳에서 사용되고 있으며, 최근 몇년사이에는 수학적지식과 IT지식을 가진 개인들도 많이 참여하고 있다.

# Wallstreet - Algorithm Trading

미국은 2012년 알고리즘 트레이딩 거래량이 85%에 달할만큼 알고리즘 트레이딩은 가파르게 증가

Algorithmic Trading. Percentage of Market Volume



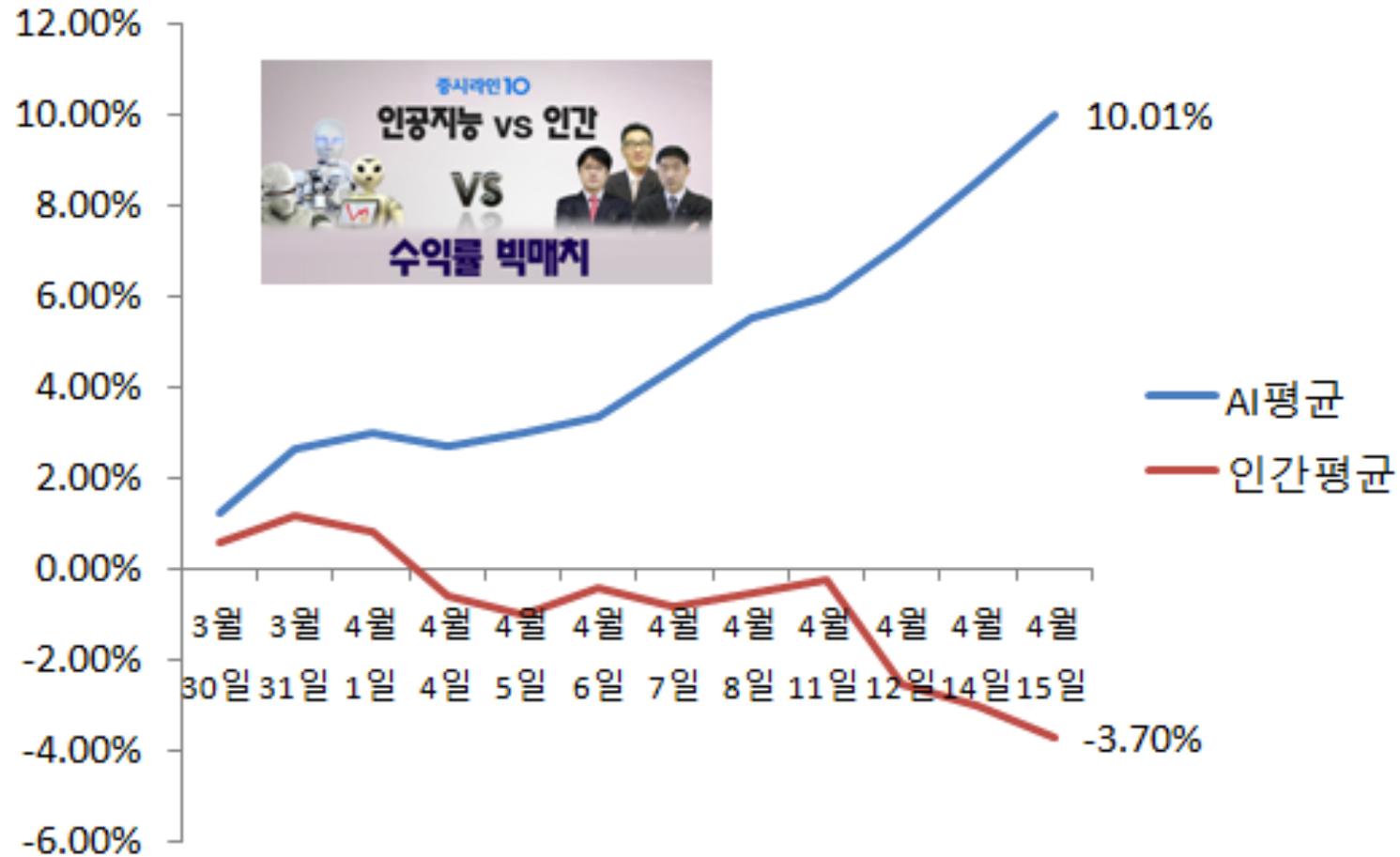
# Wallstreet - Algorithm Trading



- Goove WM, Zald DH등이 작성한 Clinical versus Mechanical Prediction: a meta-analysis 논문에 의하면 136건의 사례를 조사해보니 수학적 모델이 사람보다 비슷하거나 더 좋은 결과를 가져올 확률이 94%라고 한다.

(source - <http://www.ncbi.nlm.nih.gov/pubmed/10752360>)

# 기계 대 인간 주식투자대결 : 한국경제신문

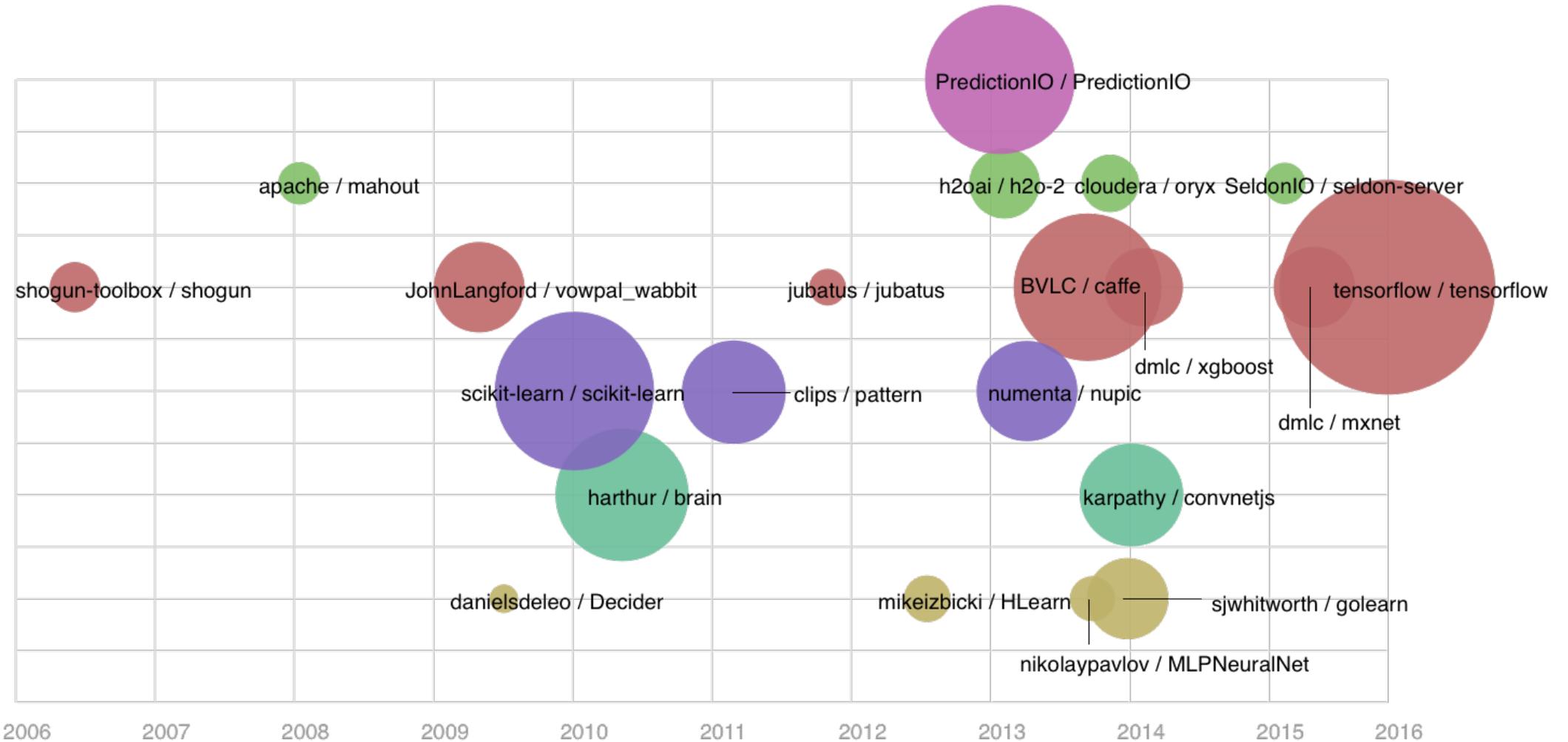


# Machine Learning & Open Source

# 10 Years of Open Source ML



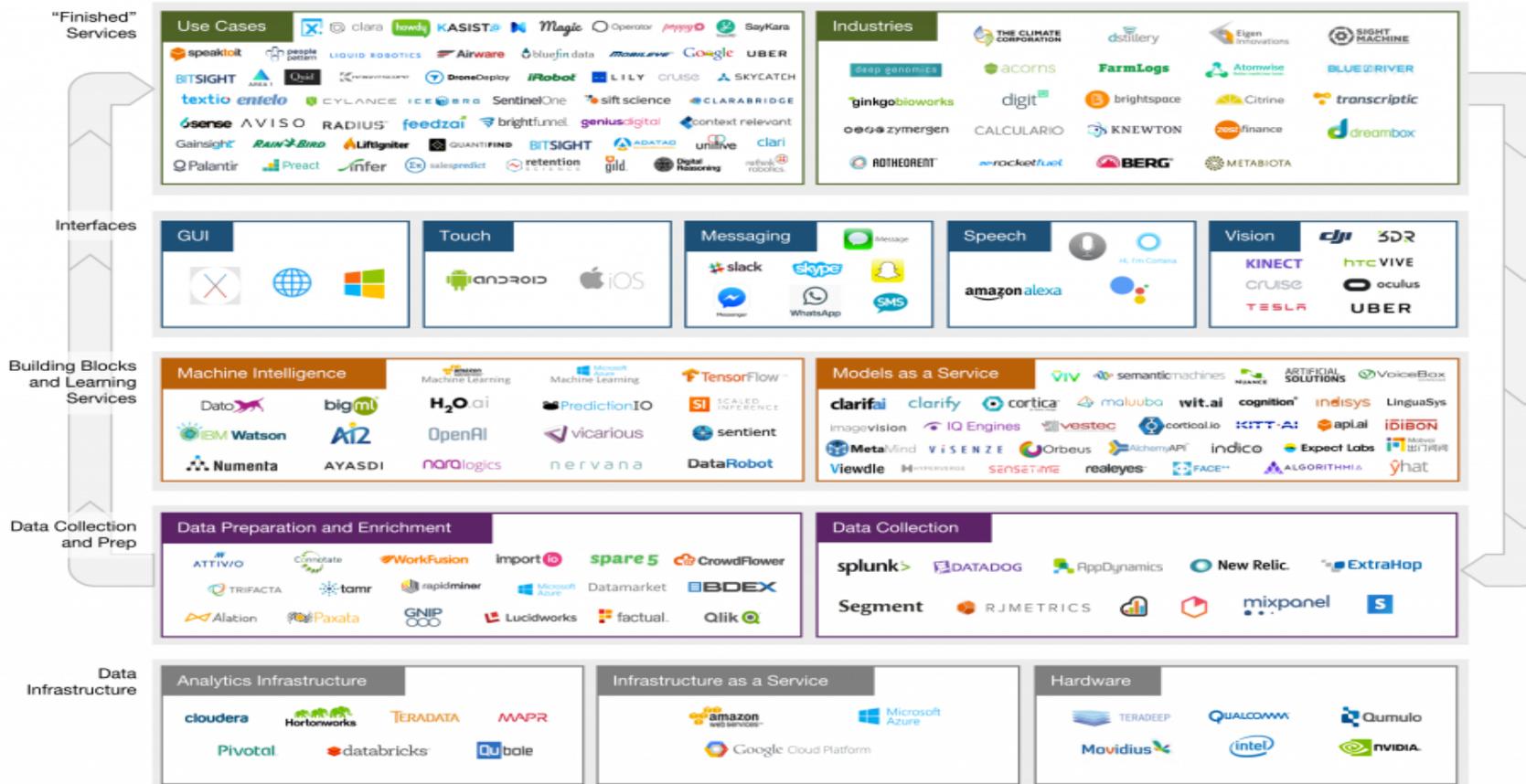
Other



# ML의 핵심기능은 오픈소스에서

오픈소스를 사용하지 않고는 ML이 사실상 거의 불가능

## Intelligent App Stack



# Deep Learning?

Caffe



DL4J  
Deeplearning4j

MINERVA

*mxnet*



**K**  
KERAS

Microsoft  
CNTK

MatConvNet

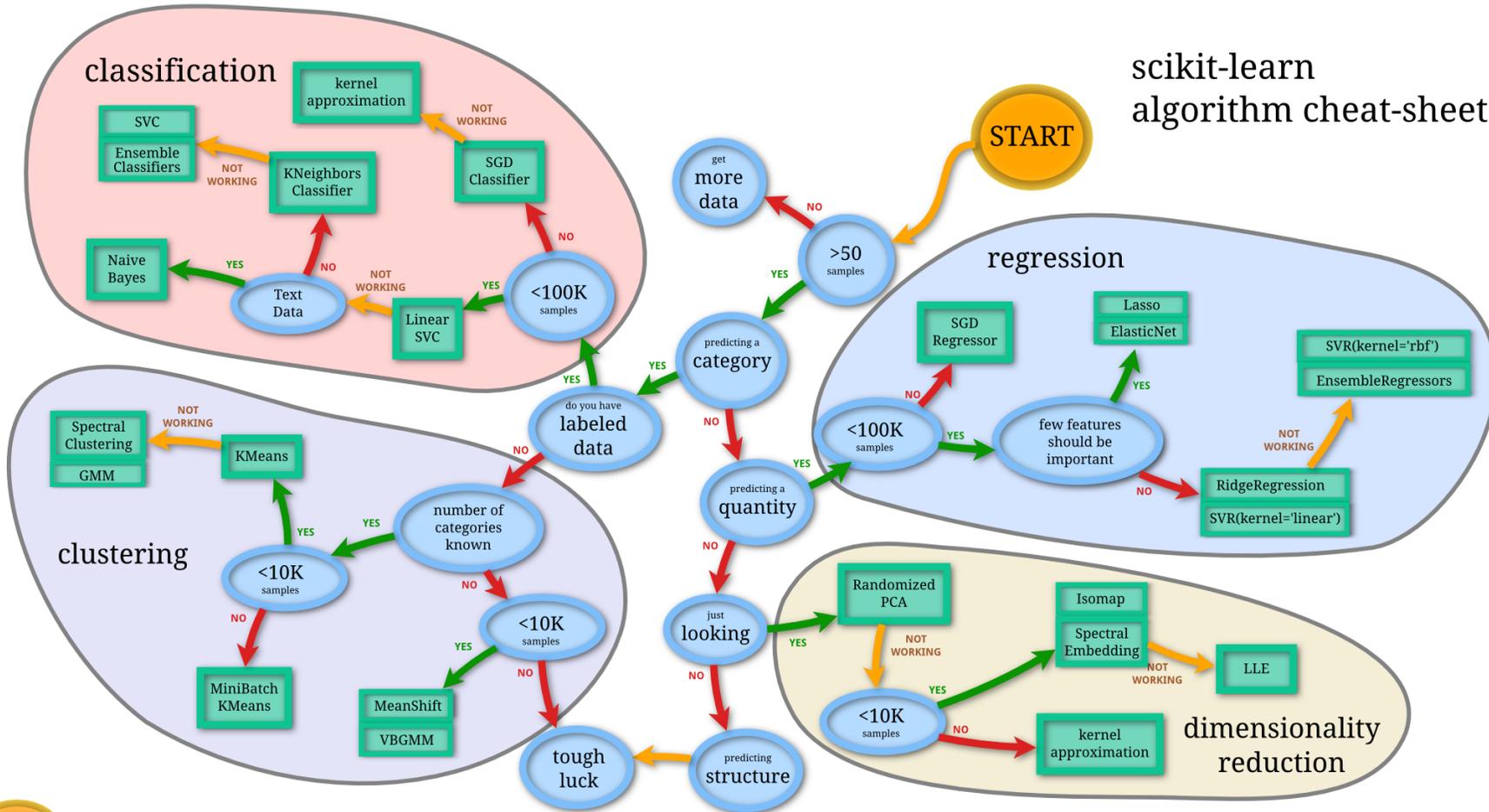
  
TensorFlow

theano

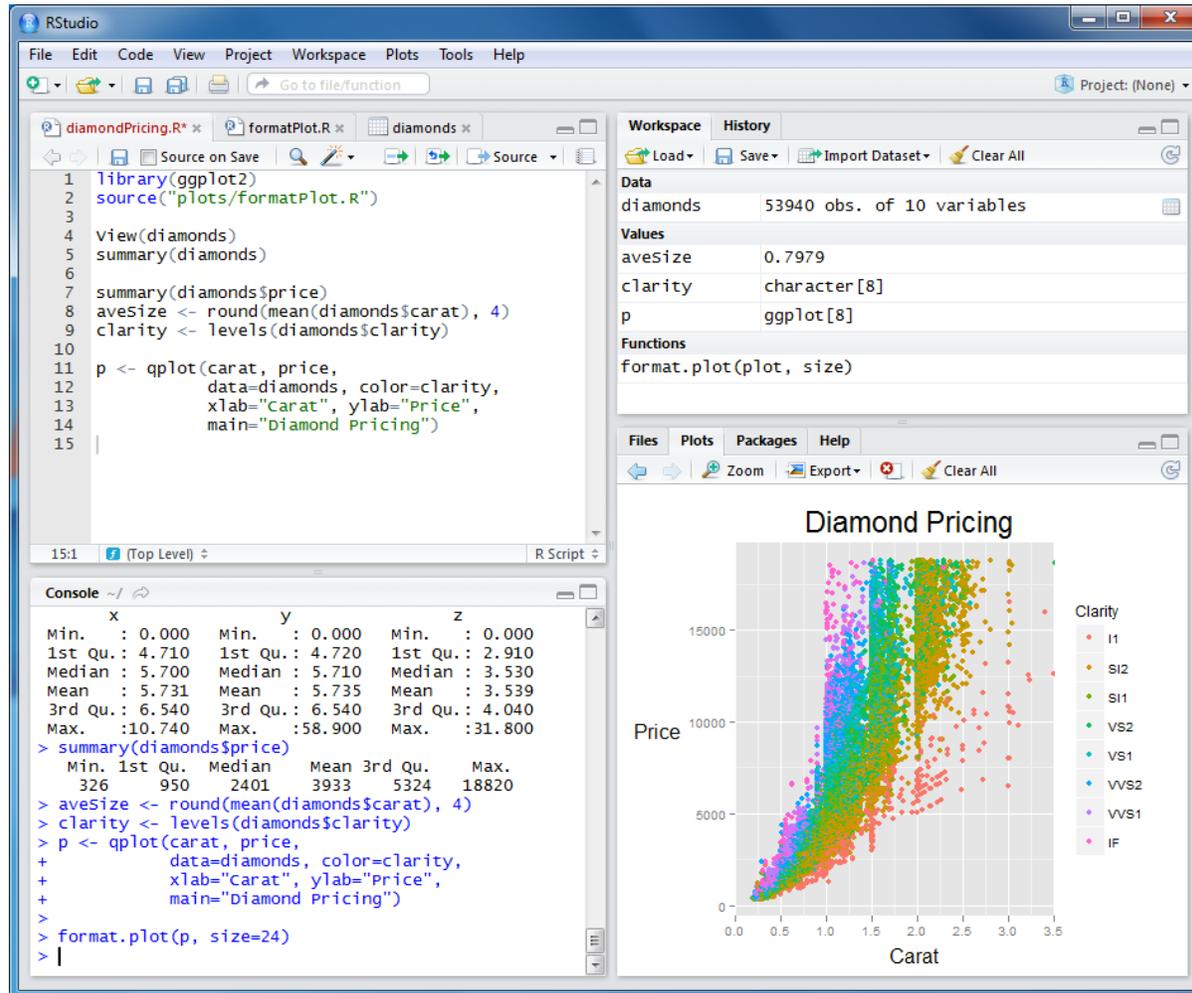


# Machine Learning?

scikit-learn  
algorithm cheat-sheet



# Data Analysis?



**Hype or Not?**

No Comments....

Cheap, Fast & Accurate

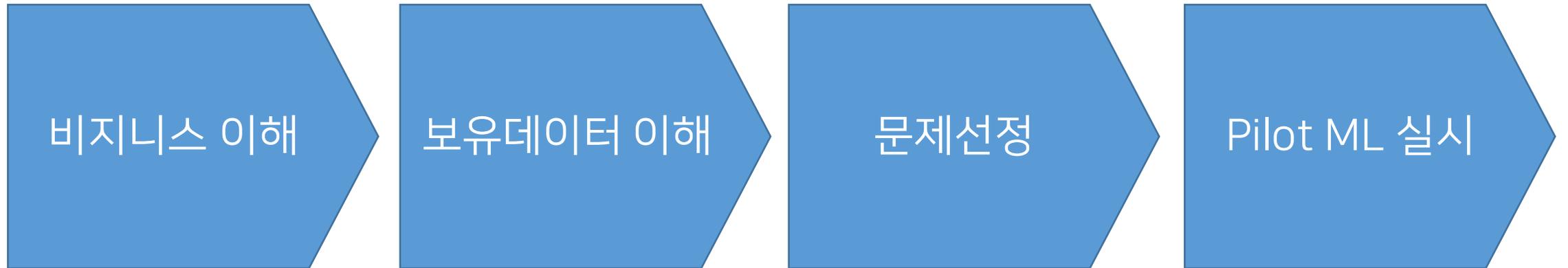
# Right Now!!!



- 일찍 시작하면 경쟁사보다 더욱 많은 고품질의 데이터 확보 가능
- 인공지능 기술 개발은 생각보다 오랜 시간이 필요함
- 오래시간 획득한 경험과 노하우가 많이 중요함
- 어차피 인공지능기술은 피할 수 없음

# Start Small

가볍게 시작할 수 있는 문제를 선정해, ML 기술개발 시작



**감사합니다.**