패턴인식 및 컴퓨터비전

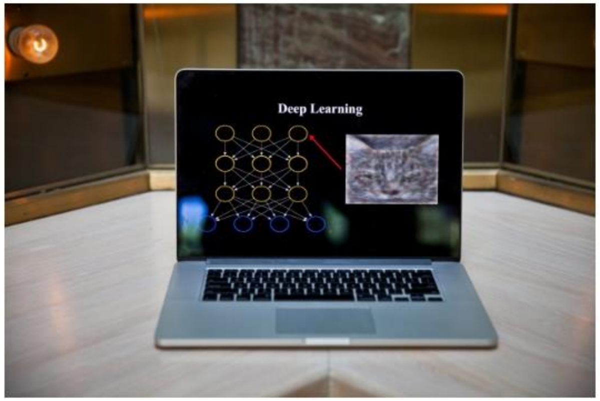
아주대학교

구형일

Course overview

- Introduction
  - Machine Learning/Deep Learning

- Machine Learning
  - Artificial Neural Network (ANN,MLP)
  - Convolution Neural Network (CNN)
  - Recurrent Neural Network (RNN)
  - (Deep) Reinforcement Learning

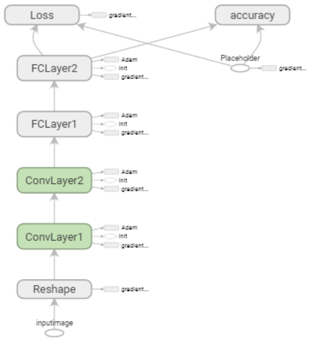


Andrew Ng's laptop explains Deep Learning. *Photo: Ariel Zambelich/Wired*

Course overview

with applications

```
tf.reset_default_graph()

x = tf.placeholder(tf.float32, [None, 784], name='inputimage')
y_ = tf.placeholder(tf.float32, [None, 10])

# [5, 5, 1, 32]: 5x5 convolution patch, 1 input channel, 32 output channel.

with tf.name_scope('Reshape'):
    x_image = tf.reshape(x, [-1,28,28,1])

with tf.name_scope('ConvLayer1'):
    W_conv = weight_variable([5, 5, 1, 32],name='1st_Layer_weight')
    b_conv = bias_variable([32],name='1st_Layer_bias')
    h_conv1 = conv2d(x_image, W_conv) + b_conv
    h_conv1_relu = tf.nn.relu(h_conv1)
    h_pool1 = max_pool_2x2(h_conv1_relu)

with tf.name_scope('ConvLayer2'):
    W_conv = weight_variable([5, 5, 32, 64],name='2nd_Layer_weight')
    b_conv = bias_variable([64],name='2nd_Layer_bias')
    h_conv2 = conv2d(h_pool1, W_conv) + b_conv
    h_conv2_relu = tf.nn.relu(h_conv2)
    h_pool2 = max_pool_2x2(h_conv2_relu)
    W_conv.name

with tf.name_scope('FCLayer1'):
    W_fc1 = weight_variable([7 * 7 * 64, 1024],name='FC1_layer_weight')
    b_fc1 = bias_variable([1024],name='FC1_layer_bias')
    h_pool2_flat = tf.reshape(h_pool2, [-1, 7*7*64])
    h_fc1 = tf.nn.relu(tf.matmul(h_pool2_flat, W_fc1) + b_fc1)

with tf.name_scope('FCLayer2'):
    W_fc2 = weight_variable([1024, 10],name='FC1_layer_weight')
    b_fc2 = bias_variable([10],name='FC1_layer_bias')
    y_conv=tf.nn.softmax(tf.matmul(h_fc1, W_fc2) + b_fc2)

with tf.name_scope('Loss'):
    cross_entropy = -tf.reduce_sum(y_*tf.log(y_conv))

with tf.name_scope('accuracy'):
    correct_prediction = tf.equal(tf.argmax(y_conv,1), tf.argmax(y_,1))
    accuracy = tf.reduce_mean(tf.cast(correct_prediction, "float"))

train_step = tf.train.AdamOptimizer(1e-4).minimize(cross_entropy)
init = tf.global_variables_initializer()
sess = tf.Session()

accuracy_list = []
with sess.as_default():
    sess.run(init)
    for i in range(20000): #20,000
        batch = mnist.train.next_batch(50)
        if i%1000 == 0:
            train_accuracy = accuracy.eval(feed_dict={x:batch[0], y_: batch[1]})
            print ("step %d, training accuracy %g" % (i, train_accuracy))
            print ("test accuracy %g" % accuracy.eval(feed_dict={x: mnist.test.images, y_: mnist.test.labels}))

        train_step.run(feed_dict={x: batch[0], y_: batch[1]})
```



with (some) codes

INTRODUCTION

교통사고원인

1. Driver distraction
2. Speeding
3. Drunk driving
4. Reckless driving
5. Rain
6. Running red lights
7. Running stop signs
8. Teenage drivers
9. Night driving
10. Design defects
11. Unsafe lane changes
12. Wrong-way driving
13. Improper turns
14. Tailgating
15. Driving under the influence of drugs
16. Ice
17. Snow
18. Road rage
19. Potholes
20. Drowsy driving
21. Tire blowouts
22. Fog
23. Deadly curves
24. Animal crossings
25. Street racing
26. Others

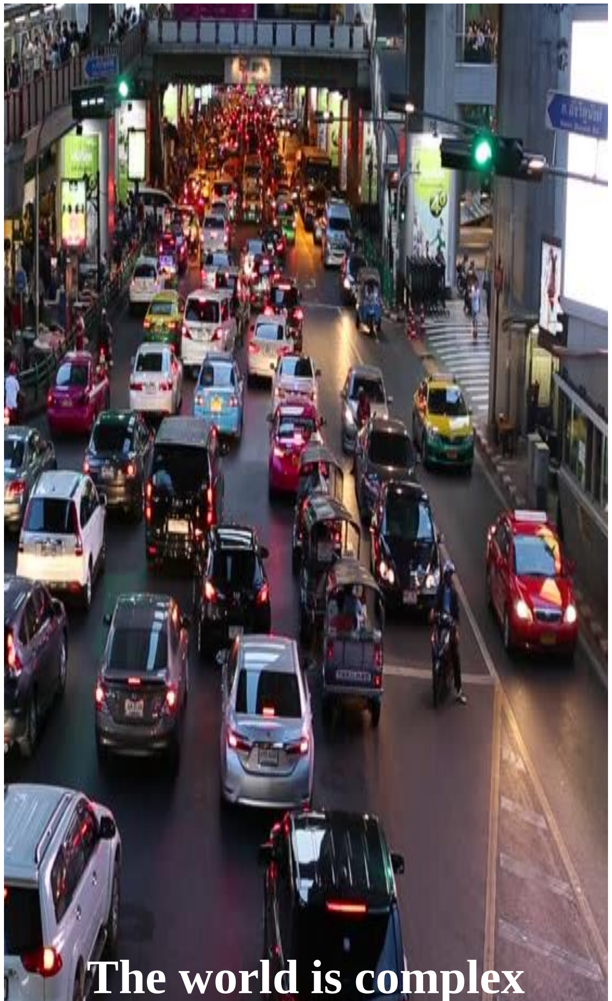
TED: Sebastian Thrun

SELF-DRIVING CARS

DARPA Grand Challenge II (2006)

DARPA Urban Challenge (2007)

Autonomous-driving is hard


The world is complex

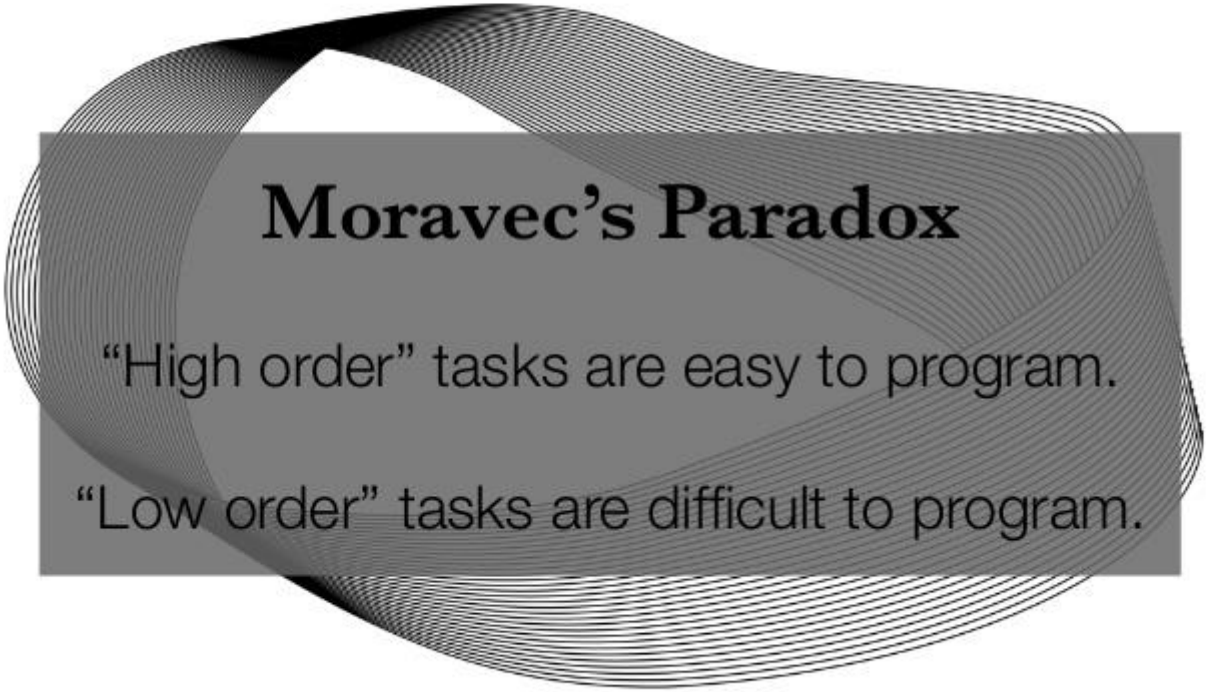

The world is unpredictable


The world is hazardous

Moravec's Paradox

- The main lesson of 35 years of AI research is that **the hard problems are easy and the easy problems are hard**. The mental abilities of a four-year-old that we take for granted – recognizing a face, lifting a pencil, walking across a room, answering a question – in fact solve some of the hardest engineering problems ever conceived... As the new generation of intelligent devices appears, it will be the stock analysts and petrochemical engineers and parole board members who are in danger of being replaced by machines. The gardeners, receptionists, and cooks are secure in their jobs for decades to come.
  - Pinker, Steven (September 4, 2007) [1994], The Language Instinct, Perennial Modern Classics, Harper, ISBN 0-06-133646-7

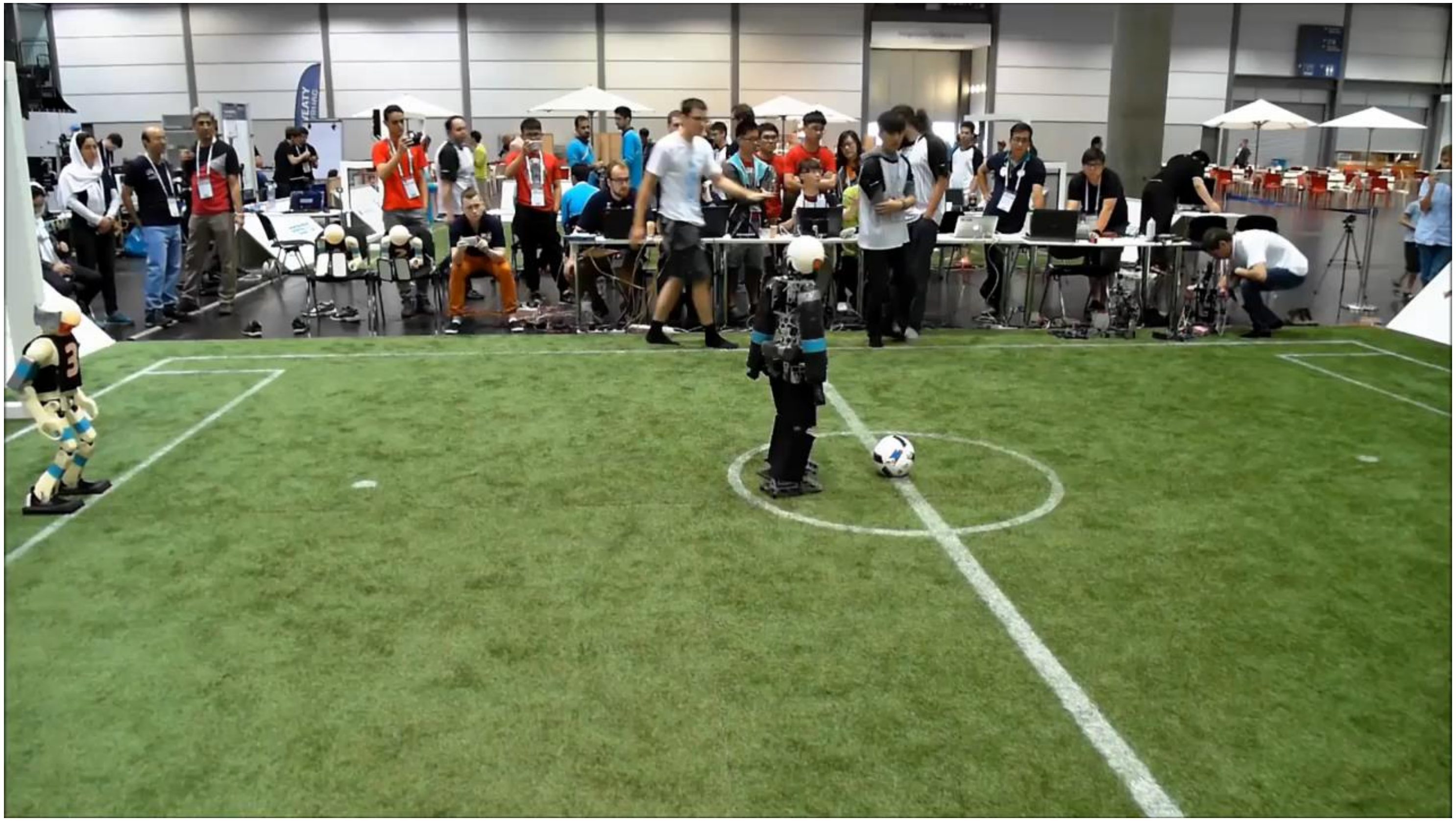
Moravec's Paradox



Moravec's Paradox

Moravec's Paradox

- RoboCup 2016: NimbRo vs AUTMan



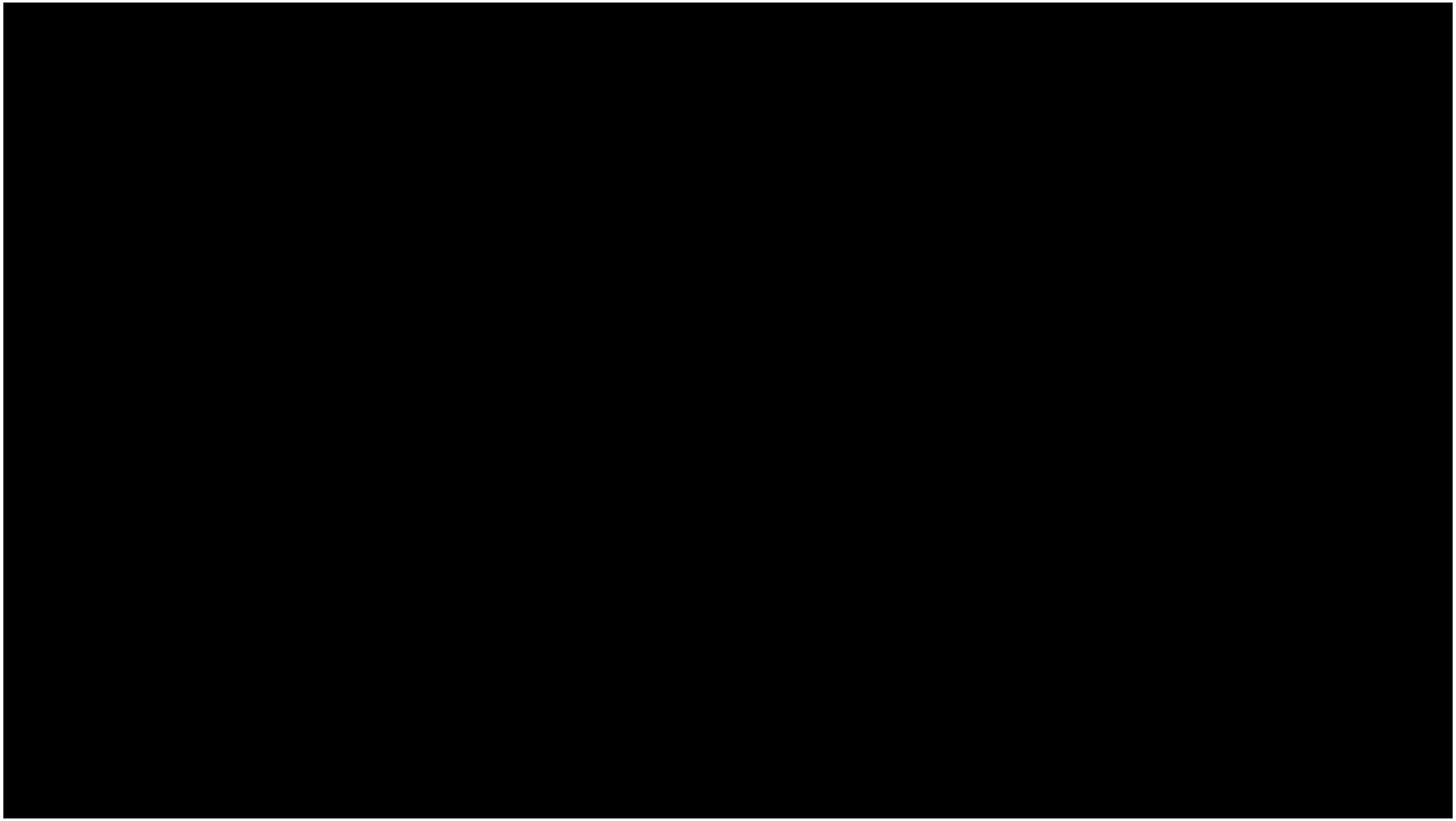
Moravec's Paradox

- A Compilation of Robots Falling Down at the DARPA Robotics Challenge

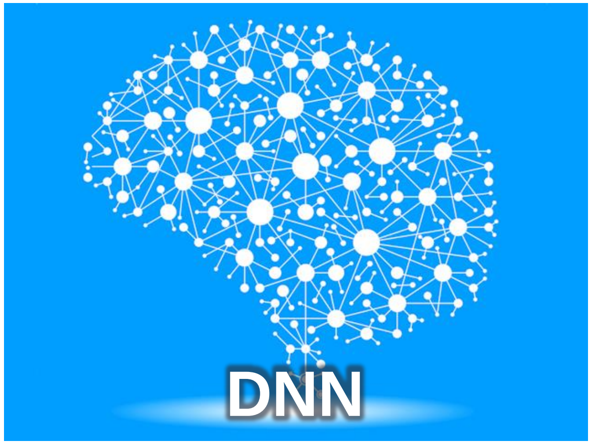
Why?

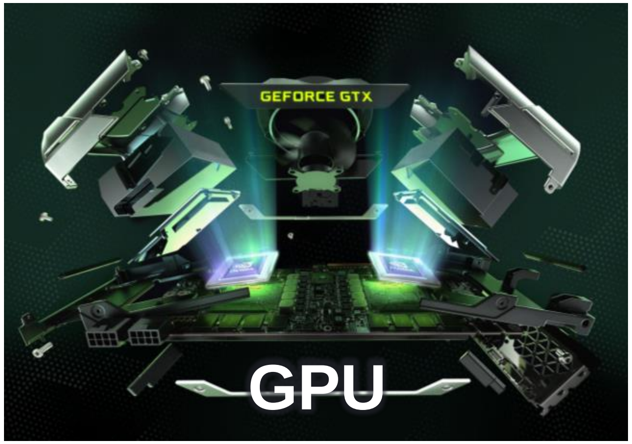
- "Encoded in the large, highly evolved sensory and motor portions of the human brain is a billion years of experience about the nature of the world and how to survive in it.
- We are all prodigious Olympians in perceptual and motor areas, so good that we make the difficult look easy. Abstract thought, though, is a new trick, perhaps less than 100 thousand years old. We have not yet mastered it. It is not all that intrinsically difficult; it just seems so when we do it"
  - Moravec, Hans (1988), Mind Children, Harvard University Press

How hard is driving?



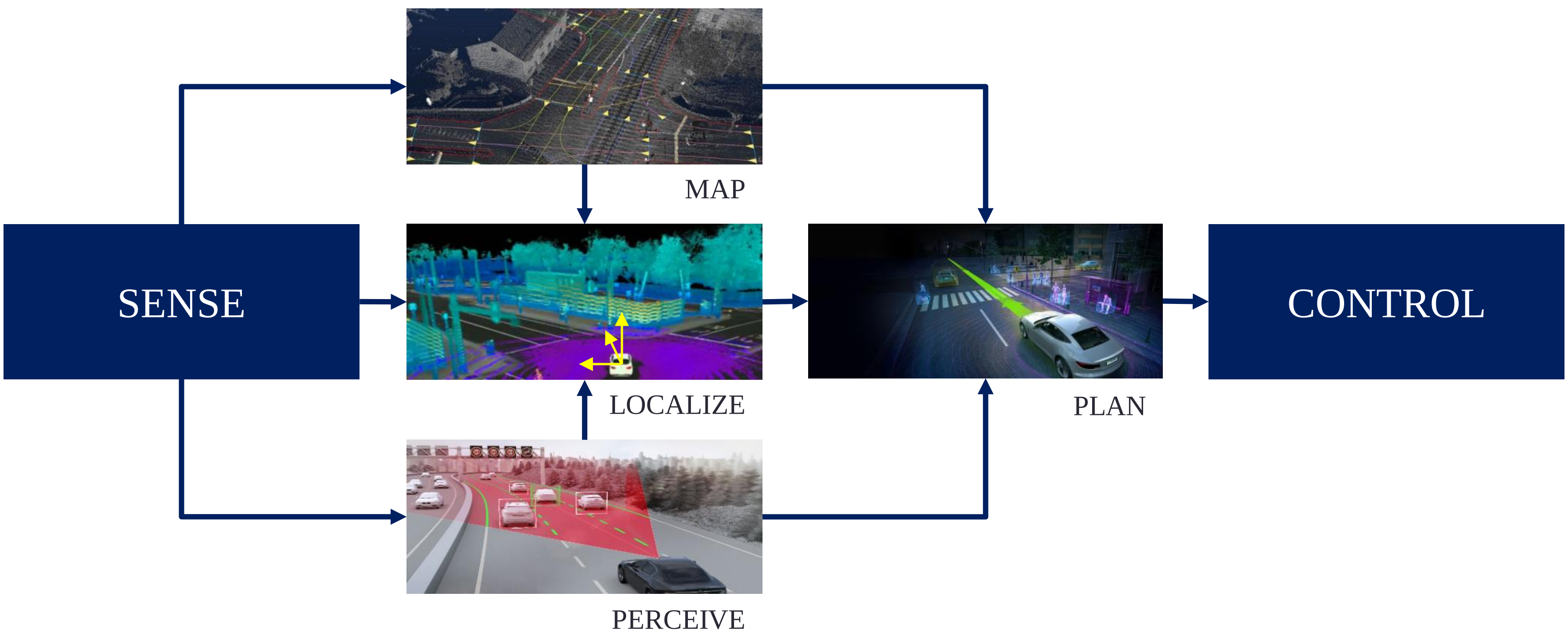
Deep learning to the rescue



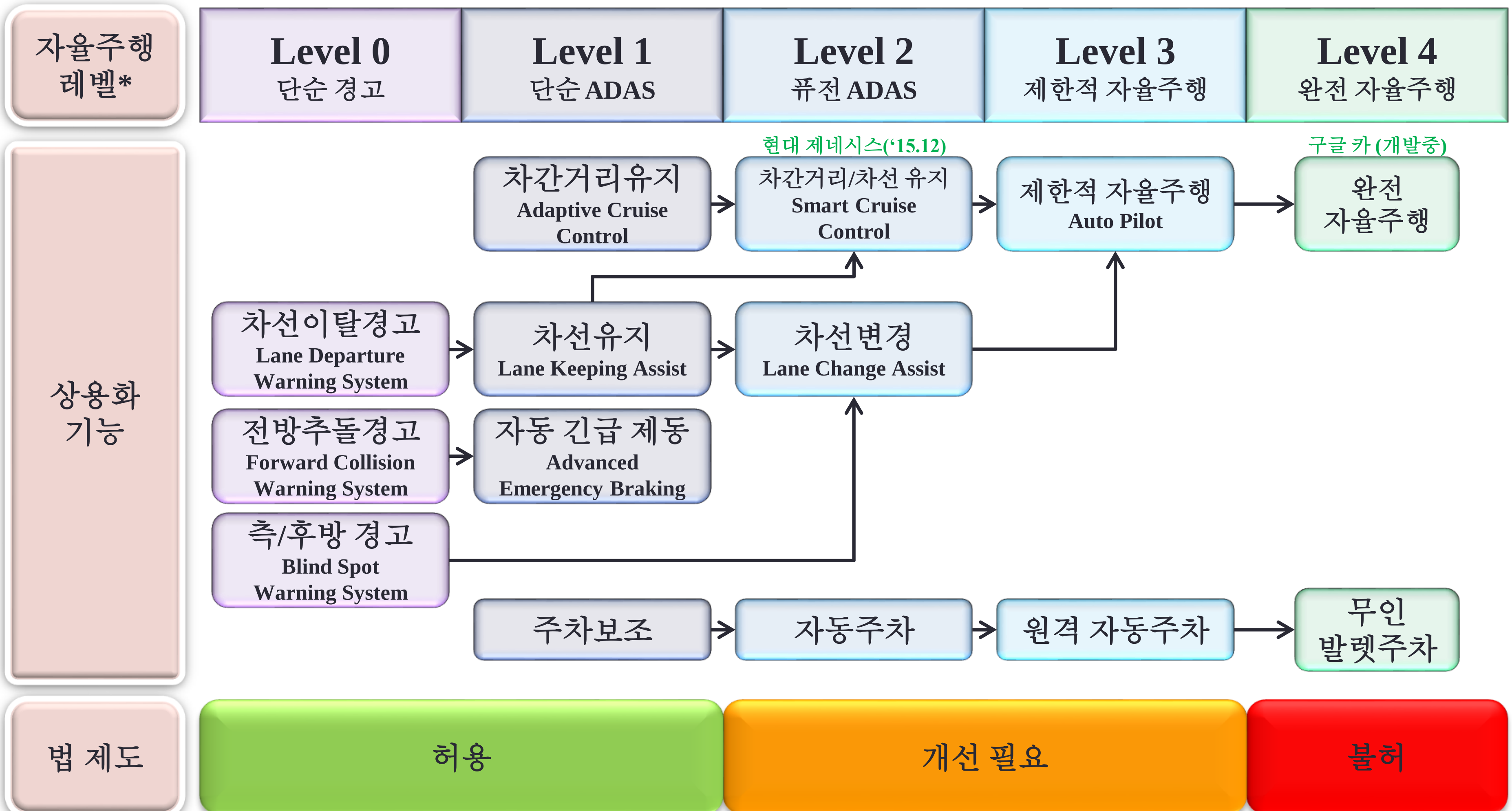


SELF DRIVING CARS

The basic self-driving loop

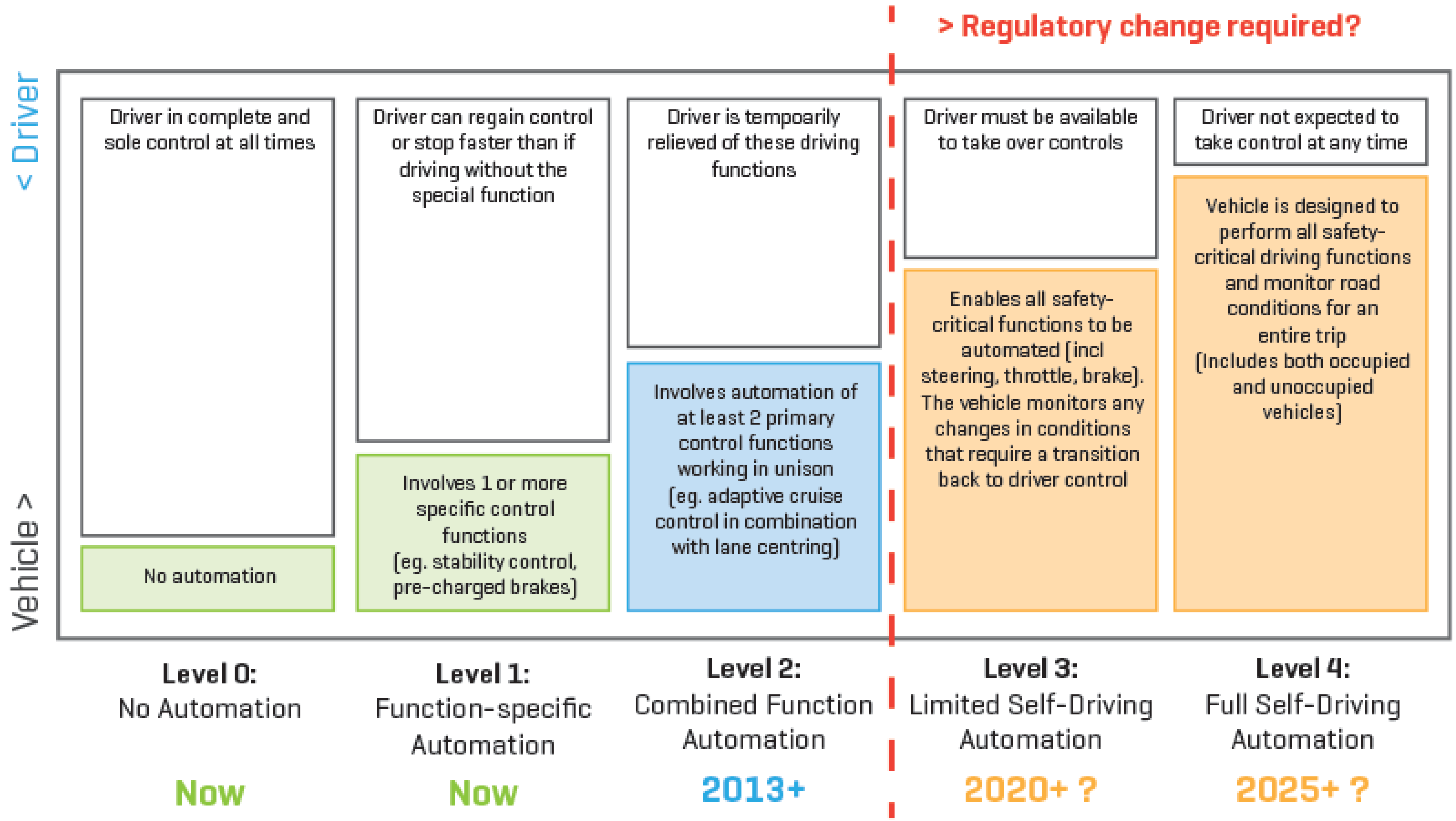


Autonomous Driving



*National Highway Traffic Safety Administration (NHTSA-미국고속도로교통안전청) 기준

Levels of driving automation (NHTSA)

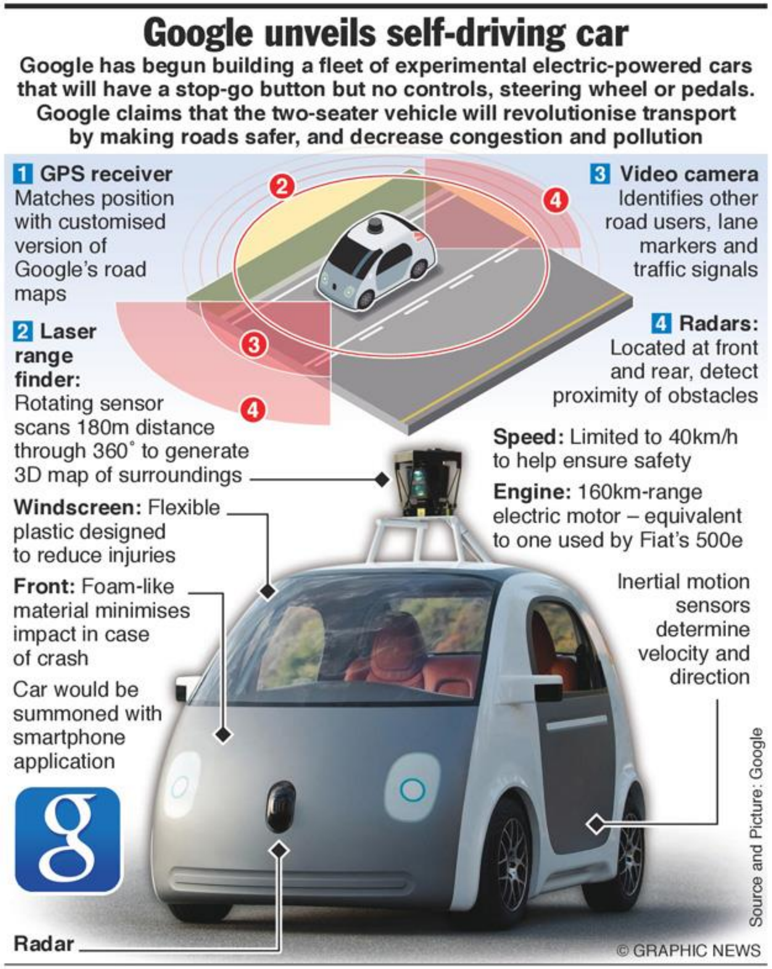


Source: NHTSA (Modified)

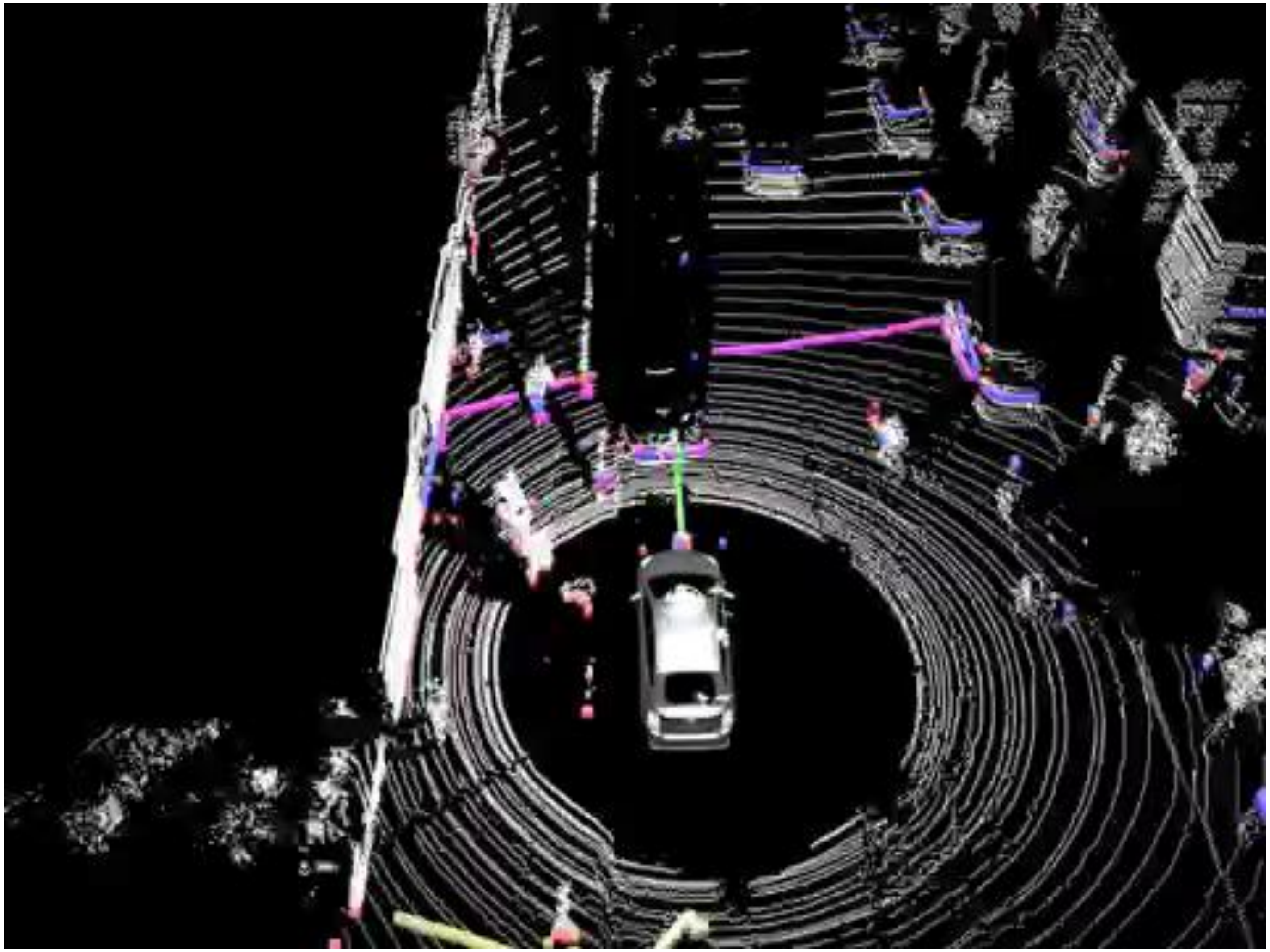
GOOGLE'S SELF DRIVING CAR

Google self-driving car

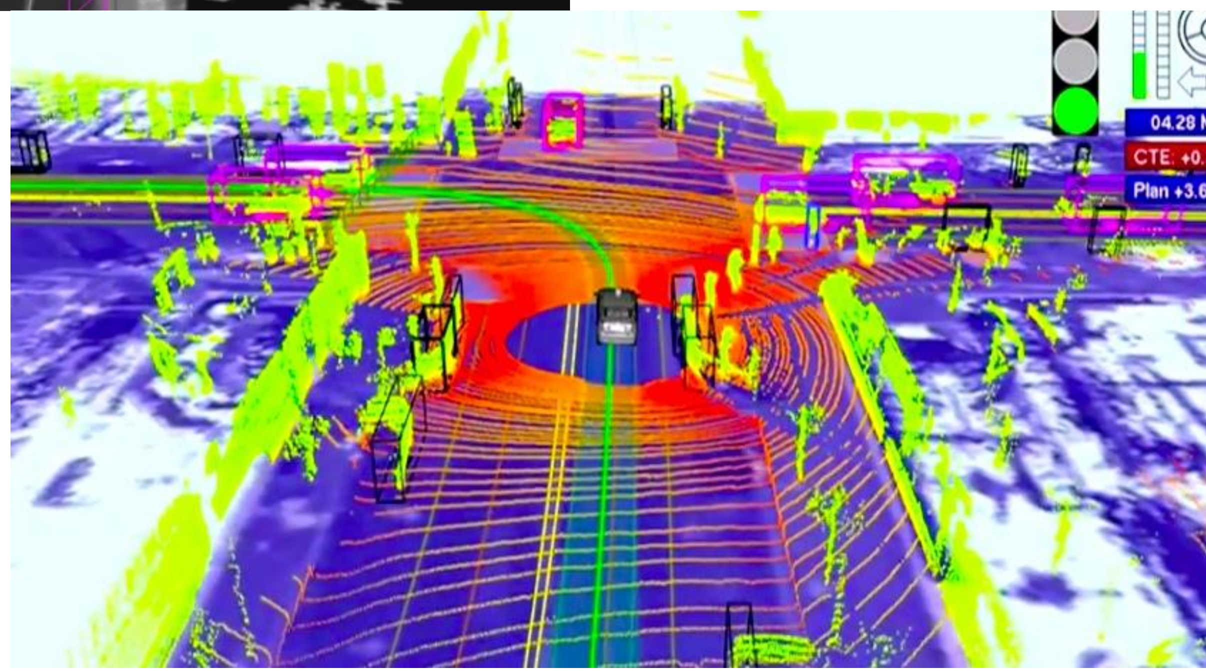
- GPS
- LiDAR
- Camera
- Radar



Visualization of LIDAR data




Cruise
14
SPEED
LIMIT
35


CTE: +0.2
Plan +3.6m

Disengagements Reports

- Disengagements: deactivations of the autonomous mode
  - when a failure of the autonomous technology is detected (272 cases)
  - when the safe operation of the vehicle requires that the autonomous vehicle test driver disengage the autonomous mode and take immediate manual control of the vehicle. (13+56 cases)

Disengagements related to detection of a failure of the autonomous technology

Month	Number Disengages	Autonomous miles on public roads
2014/09	0	4207.2
2014/10	14	23971.1
2014/11	14	15836.6
2014/12	40	9413.1
2015/01	48	18192.1
2015/02	12	18745.1
2015/03	26	22204.2
2015/04	47	31927.3
2015/05	9	38016.8
2015/06	7	42046.6
2015/07	19	34805.1
2015/08	4	38219.8
2015/09	15	36326.6
2015/10	11	47143.5
2015/11	6	43275.9
Total	272	424331

Driver-initiated disengagements related to safe operation of the vehicle

Month	Number Disengages	Autonomous miles on public roads
2014/09	2	4207.2
2014/10	5	23971.1
2014/11	7	15836.6
2014/12	3	9413.1
2015/01	5	18192.1
2015/02	2	18745.1
2015/03	4	22204.2
2015/04	4	31927.3
2015/05	4	38016.8
2015/06	4	42046.6
2015/07	10	34805.1
2015/08	3	38219.8
2015/09	1	36326.6
2015/10	5	47143.5
2015/11	10	43275.9
Total	69	424331

California Autonomous Testing Disengagements (2015)

Company	Miles Driven	DE*	Miles per DE	Miles per DE in 2015	Common Causes
Waymo (aka Google)	635868	124	5128	1244.4	Software discrepancy; unwanted vehicle manuver
VW/Audi	N/A	N/A	N/A	74.8	N/A
Mercedes-Benz	673.4	336	2	1.8	Driver discomfort; technology evaluation management
Delphi	3125.3	178	17.6	41.9	Completing lane change in heavy traffic; traffic light detection
Tesla Motors	550	182	3	N/A	Planner output invalid; follower output invalid
Bosch	983	1442	0.7	1.5	Planned test of technology
Nissan	4099	28	246.7	14	AV system failure; AV is about to collide with vehicle or obstacle
Cruise (GM)	9846.5	414	9.3	N/A	To avoid unexpected behavior
BMW	638	1	638	N/A	Lane marking unclear
Honda	N/A	N/A	N/A	N/A	N/A
Ford	590	3	196.7	N/A	Aborted lane change due to vehicle overtaking at high speed

*DE = Disengagements

https://www.wired.com/2017/02/california-dmv-autonomous-car-disengagement/

TESLA'S AUTOPILOT

Autopilot

Autopilot combines a forward looking camera, radar, and 360 degree sonar sensors with real time traffic updates to automatically drive Model S on the open road and in dense stop and go traffic. Changing lanes becomes as simple as a tap of the turn signal. When you arrive at your destination, Model S will both detect a parking spot and automatically park itself. Standard equipment safety features are constantly monitoring stop signs, traffic signals and pedestrians, as well as for unintentional lane changes.

TESLA

https://www.tesla.com/

Tesla vs Google

How Tesla's Technology Works

The Tesla uses a **computer vision-based** vehicle detection system, but according to the company, it is not intended to be used hands-free and parts of the system are unfinished.

The accident may have happened in part because the crash-avoidance system is designed to engage only when radar and computer vision systems agree that there is an obstacle, according to an industry executive with direct knowledge of the system.

Tesla Model S Michael Nagle for The New York Times

Forward-facing camera
Image-processing software can detect lane stripes, signs, stoplights, road signs and other objects.

Forward radar
Reflected microwaves can identify location and speed — but not always type — of nearby vehicles.

Ultrasonic sensors
Reflected sound waves detect distance to nearby objects.

GPS
Combined with high-precision mapping, GPS determines the car's position on the road.

How It Compares

Google does not intend to make its own cars but to partner with carmakers, and recently announced plans to adapt 100 Chrysler minivans for autonomous driving. Google's cars primarily use a laser system known as **Lidar** (light detection and ranging), a spinning range-finding unit on top of the car that creates a detailed map of the car's surroundings as it moves.

Lidar is also used on many of the experimental autonomous vehicles being developed by Nissan, BMW, Apple and others, but not by Tesla. Some experts speculate that a Lidar-driven car might have avoided this fatal crash.

Google's self-driving car on its test track. Gordon De Los Santos/Google

Google과 Tesla의 자율주행 자동차 기술 차이

	Google	Tesla
Computer Vision	LIDAR 사용 (높은 위치 인식 능력)	Camera 사용 테슬라의 CEO 앨런 머스크는 구글의 LIDAR 센서에 관해 **"그렇게 비싼 센서를 사용한 자율주행 자동차를 개발하는 것은 과하다**(overkill)"라고 비판
Car Control	완전 자율 주행 기술 목표 2013년 구글은 일부 직원의 출퇴근에 자율주행차를 타도록 했는데 차 안 비디오카메라를 모니터링한 결과 **운전자가 잠이 드는 등 운전에 집중하지 않음. 이후 완전 자율 주행 기술을 목표로 개발**	Autopilot 기능 제공 자동차의 비행기화 '비행기에서 돌발 상황에만 파일럿이 개입하듯 돌발 상황에서 운전자의 조작이 필요'

테슬라 자율주행 모드 첫 사망사고 발생…美당국 조사 착수

자동주행 모드로 운행 중이던 테슬라 모델 S 전기자동차의 운전자가 충돌사고로 사망했다고 테슬라가 30일(현지시간) 밝혔다.

테슬라는 미국 고속도로교통안전청(NHTSA)에 사고 내용을 통보했으며 NHTSA가 이에 대한 예비조사를 개시했다며 이렇게 밝혔다.

이 사고는 플로리다주 윌리스턴에서 올해 5월 7일 발생했다.

예비조사 보고서에 따르면 충돌사고가 발생한 시점은 옆면이 하얀색으로 칠해진 대형 트레일러트럭이 테슬라 앞에서 좌회전할 때였다. 사고 지점은 양방향이 중앙분리대로 분리된 고속도로의 교차로였으며, 신호등은 없었다.

충돌 당시 모델 S의 앞쪽 창문이 트레일러의 바닥 부분과 부딪혔으며 이때 당한 부상으로 모델 S 운전자가 사망했다.

테슬라에 따르면 운전자와 자동주행 센서 양쪽 모두 트레일러의 하얀색 면을 인식하지 못했고 브레이크를 걸지 않았다.

이 회사는 사고 당시 '밝게 빛나고 있던 하늘'이 배경에 깔려 있어 운전자나 자동주행 센서가 트레일러의 하얀색 면을 인식하지 못했던 것으로 보인다고 설명했다.

이 회사는 이번 사고가 '비극적 손실'이었으며 자동주행 모드가 작동되고 있는 상태에서 발생한 첫 사망사고라고 설명했다.

테슬라는 자사 자동차들이 자동주행 모드로 운행한 누적 거리는 2억900만 km에 이른다며, 미국과 세계의 모든 자동차를 놓고 따지면 사망사고가 각각 주행거리 1억5천 km, 9천700만 km에 한 차례 꼴로 일어난다고 말했다.

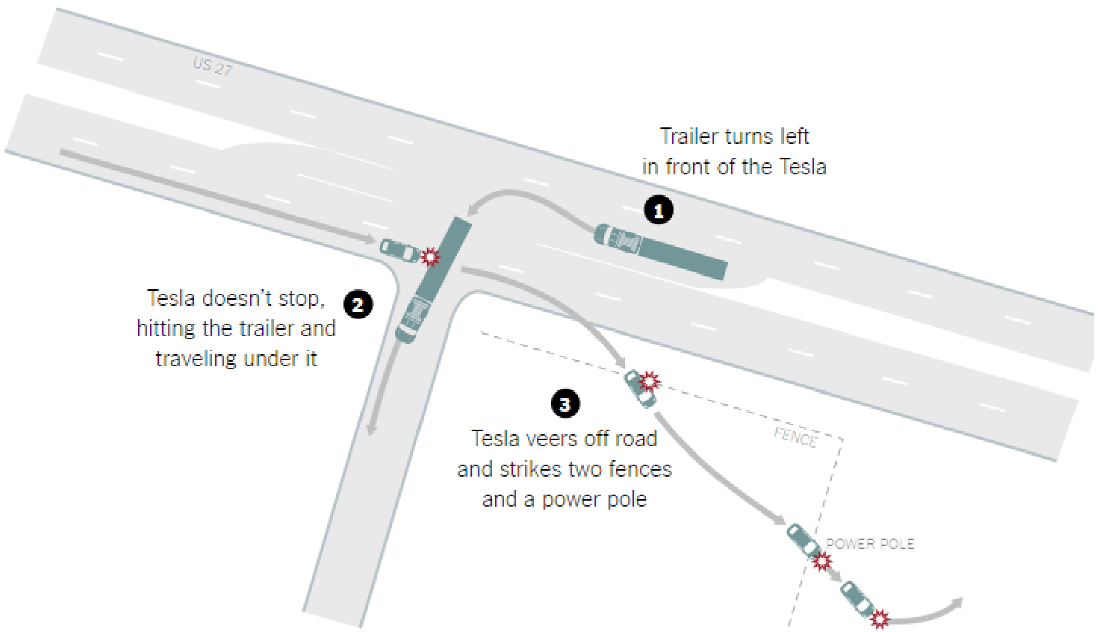
테슬라는 NHTSA의 예비조사가 지금 단계에서는 시스템이 기대대로 작동했는지 판별하기 위한 예비조사에 불과하다고 강조했다.

한경닷컴 뉴스룸 open@hankyung.com

http://www.hankyung.com/news/app/newsview.php?aid=201607018385g

How the Accident happened

The Tesla Model S crashed in northern Florida into a truck that was turning left in front of it. The Tesla then ran off the road, hitting a fence and a power pole before coming to a stop.



The New York Times | Source: Florida traffic crash report

Traffic Fatalities

- Total miles driven in U.S. in 2014:
  - 3,000,000,000,000 (3 million million)
  - Fatalities: 32,675 (1 in 90 million)

- Tesla Autopilot mile driven since October 2015:
  - 300,000,000 (300 million)
  - Fatalities: 1

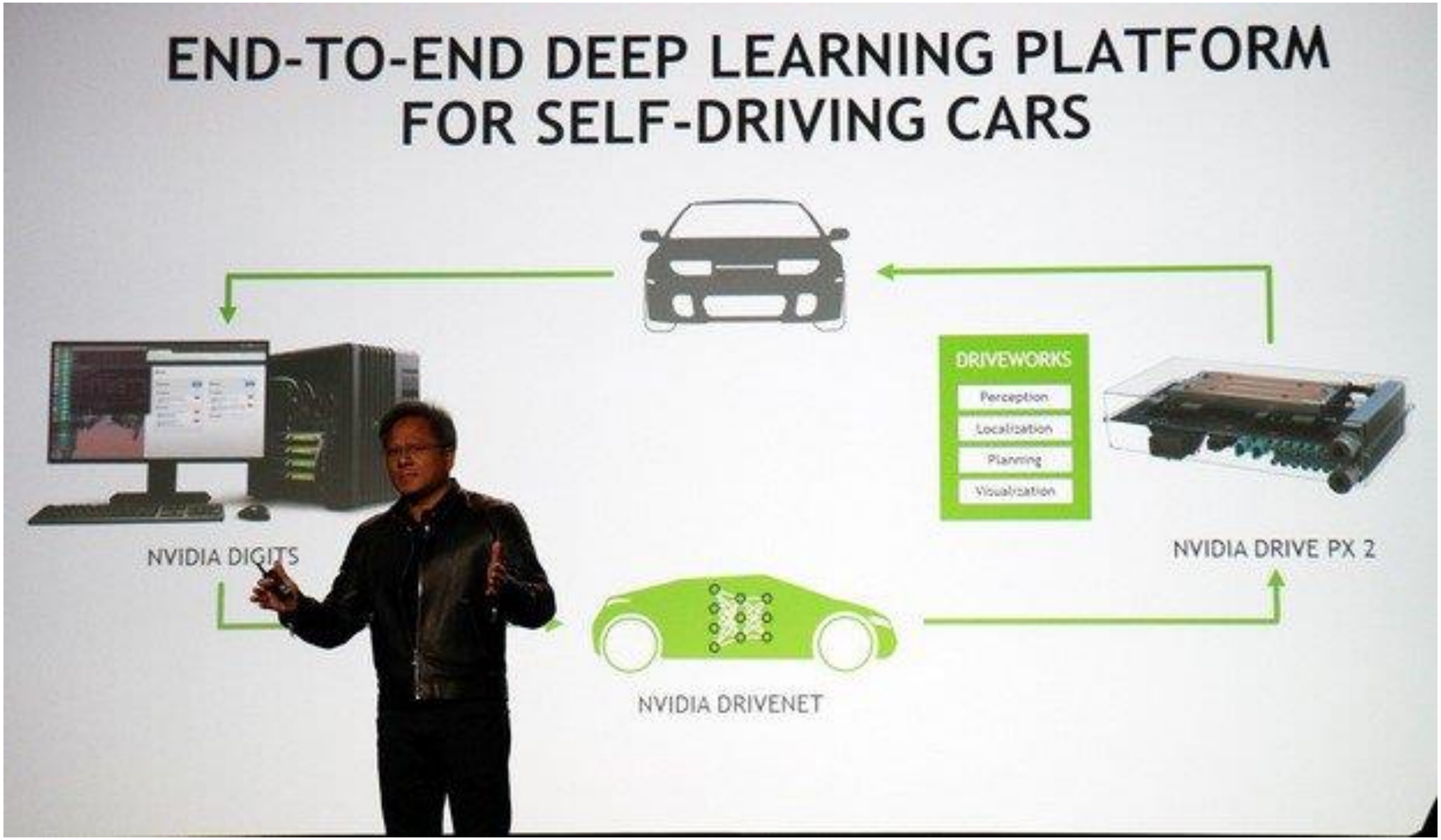
UBER…

Pittsburgh, your self-driving Uber is arriving now

UBER/nuTonomy

NVIDIA'S DRIVERWORKS

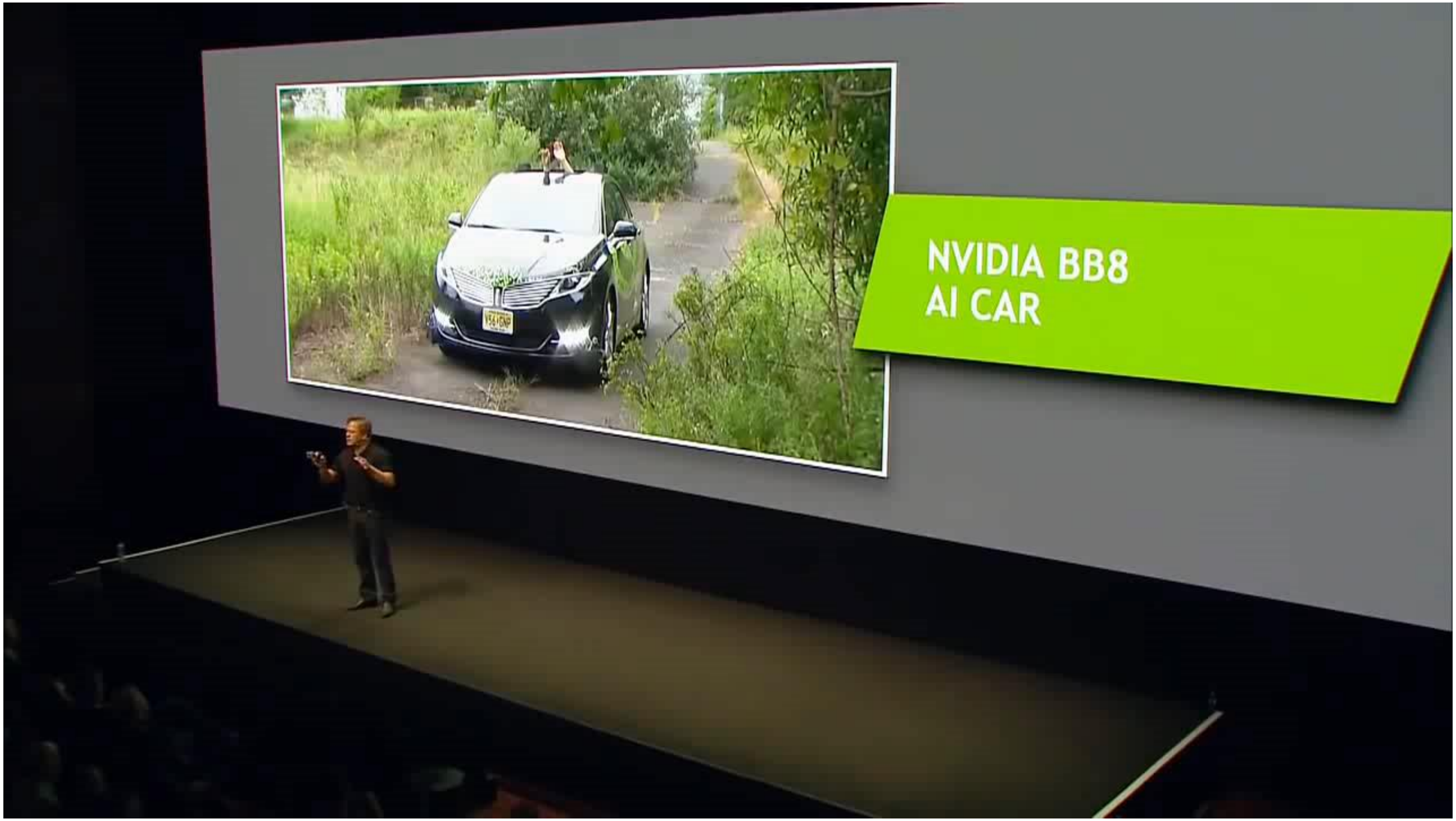
NVIDIA


END-TO-END DEEP LEARNING PLATFORM
FOR SELF-DRIVING CARS
DRIVEWORKS
Perception
Localization
Planning
Visualization
NVIDIA DIGITS
NVIDIA DRIVE PX 2
NVIDIA DRIVENET

Perception

Visualization

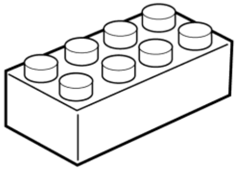
Planning

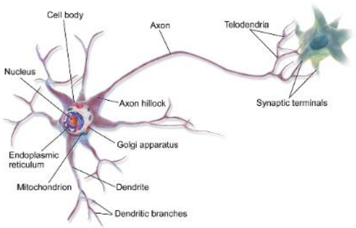


ARTIFICIAL NEURON

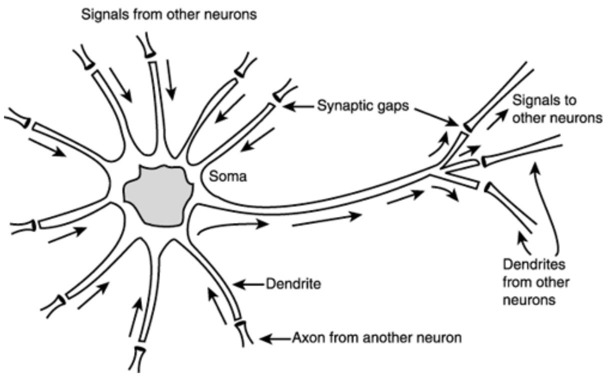
					Tasks					
					ADAS					
					Self Driving					
					Localization	Perception	Planning/ Control	Driver state	Vehicle Diagnosis	Smart factory
Methods	Traditional	Non-machine Learning			GPS, SLAM		Optimal control			
Methods	Traditional	Machine-Learning based method	Supervised	SVM MLP		Pedestrian detection (HOG+SVM)				
Methods	Deep-Learning based	Machine-Learning based method	Supervised	CNN		Detection/ Segmentation/Classification	End-to-end Learning			
Methods	Deep-Learning based	Machine-Learning based method	Supervised	RNN (LSTM)		Dry/wet road classification	End-to-end Learning			
Methods	Deep-Learning based	Machine-Learning based method	Supervised	DNN						
Methods	Deep-Learning based	Machine-Learning based method	Reinforcement							
Methods	Deep-Learning based	Machine-Learning based method	Unsupervised							

뉴런: 신경망의 기본 단위

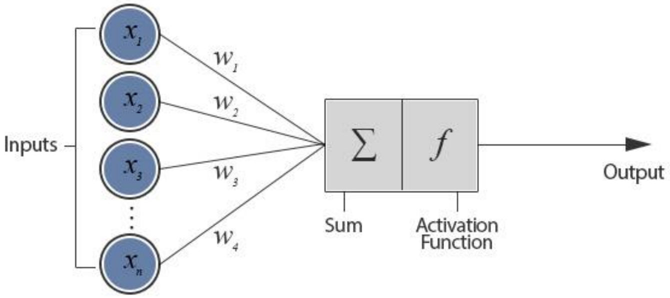



Cell body
Axon
Telodendria
Nucleus
Axon hillock
Synaptic terminals
Golgi apparatus
Endoplasmic reticulum
Mitochondrion
Dendrite
Dendritic branches

인공 뉴런(Artificial Neuron)



실제 뉴런



뉴런의 수학적 모델

예시: 연어와 농어의 구별

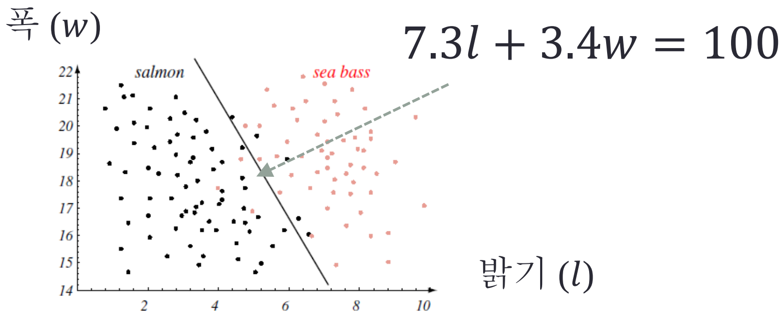
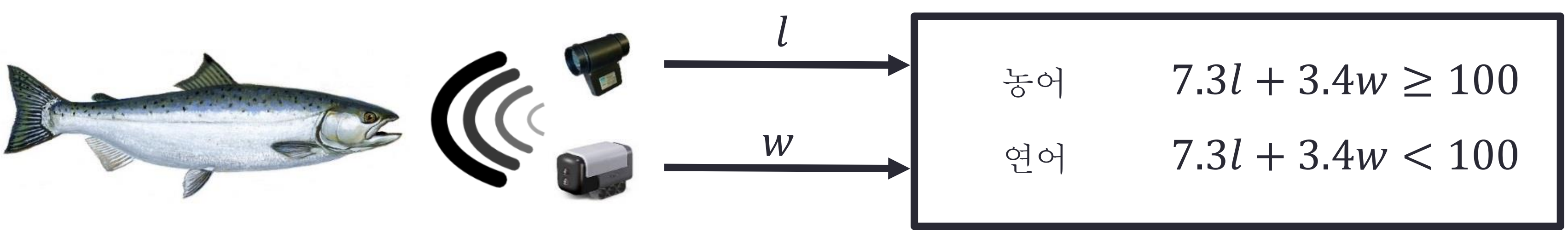
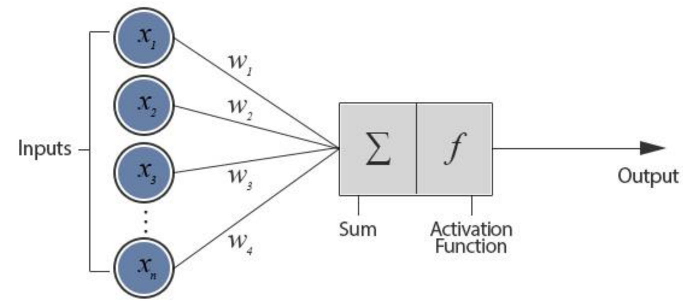


FIGURE 1.4. The two features of lightness and width for sea bass and salmon. The dark line could serve as a decision boundary of our classifier. Overall classification error on the data shown is lower than if we use only one feature as in Fig. 1.3, but there will still be some errors. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.



예시: 연어와 농어의 구별



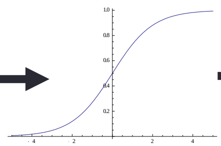
l

$7.3 \times l$

w

$3.4 \times w$

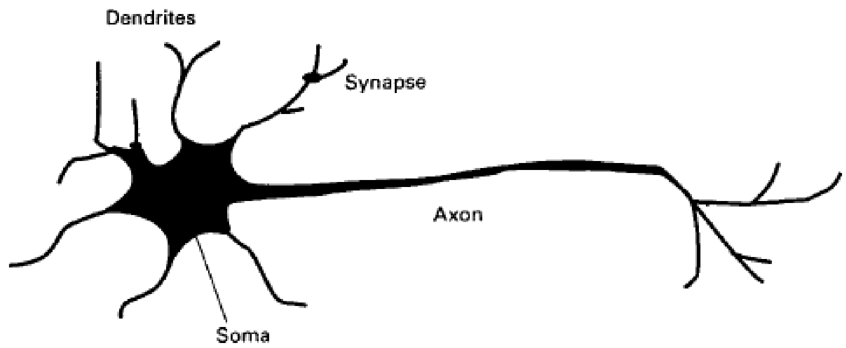
Σ

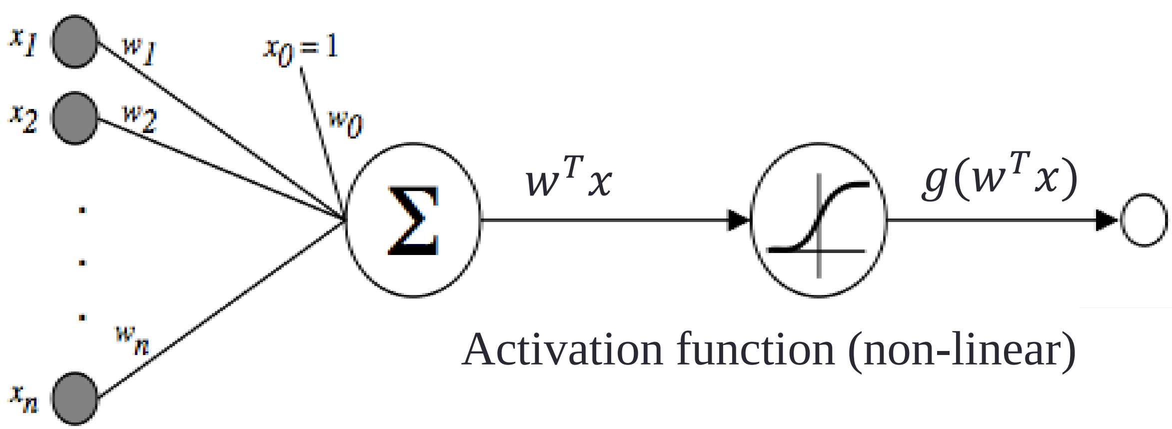


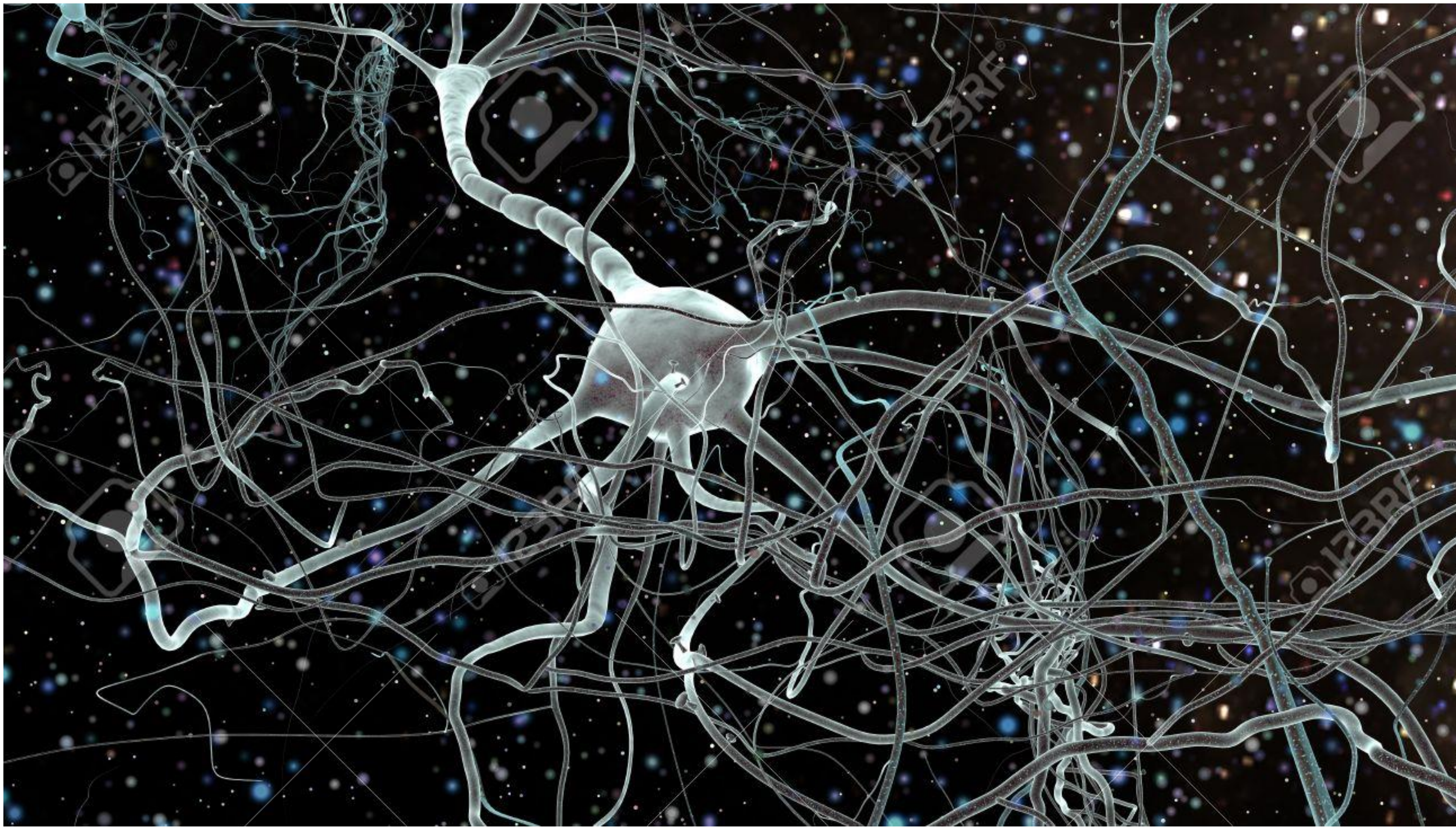
100

연어/농어

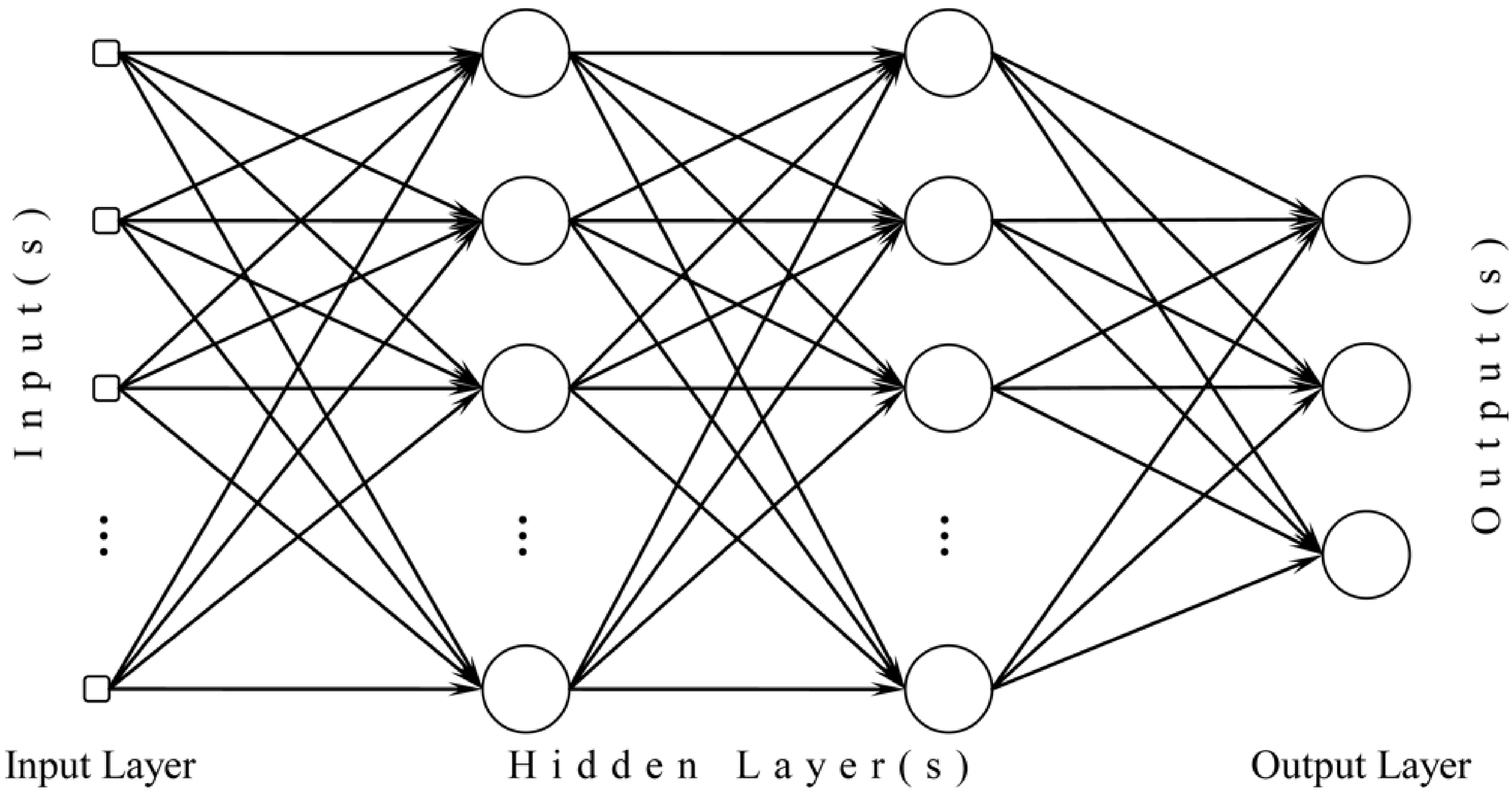
Artificial Neuron







Multi-layer Perceptron



TYPES OF MACHINE LEARNING

Neural Network 기반 방법을 중심으로

				Tasks					
				ADAS					
				Self Driving					
				Localization	Perception	Planning/ Control	Driver state	Vehicle Diagnosis	Smart factory
Methods	Traditional	Non-machine Learning		GPS, SLAM		Optimal control			
	Traditional	Machine-Learning based method	Supervised: SVM MLP		Pedestrian detection (HOG+SVM)				
	Deep-Learning based	Machine-Learning based method	Supervised: CNN		Detection/ Segmentation/Classification	End-to-end Learning			
	Deep-Learning based	Machine-Learning based method	Supervised: RNN (LSTM)		Dry/wet road classification	End-to-end Learning	Behavior Prediction/ Driver identification		*
	Deep-Learning based	Machine-Learning based method	Supervised: DNN					*	*
	Deep-Learning based	Machine-Learning based method	Reinforcement			*			
	Deep-Learning based	Machine-Learning based method	Unsupervised						*

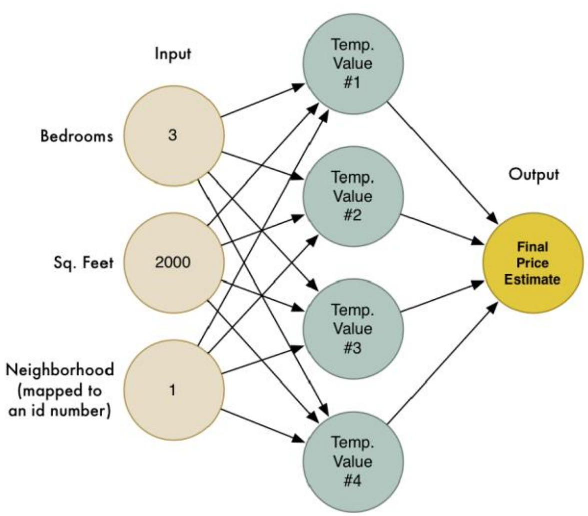
Why neural networks?

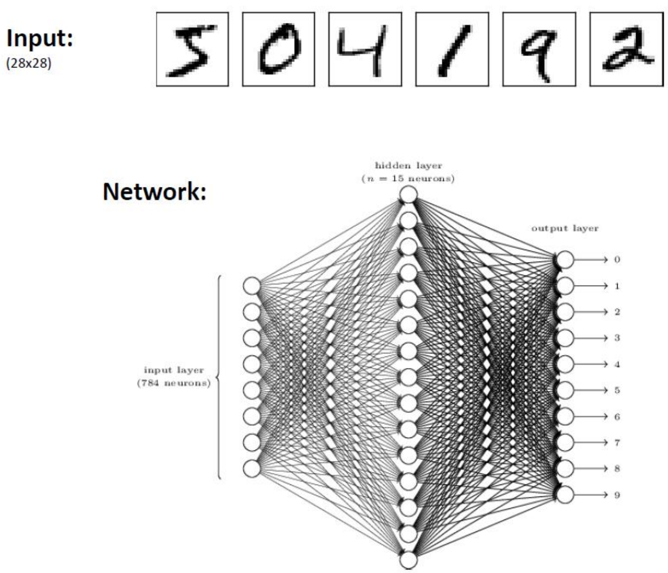
- Universal function approximator

Universal approximation theorem

From Wikipedia, the free encyclopedia

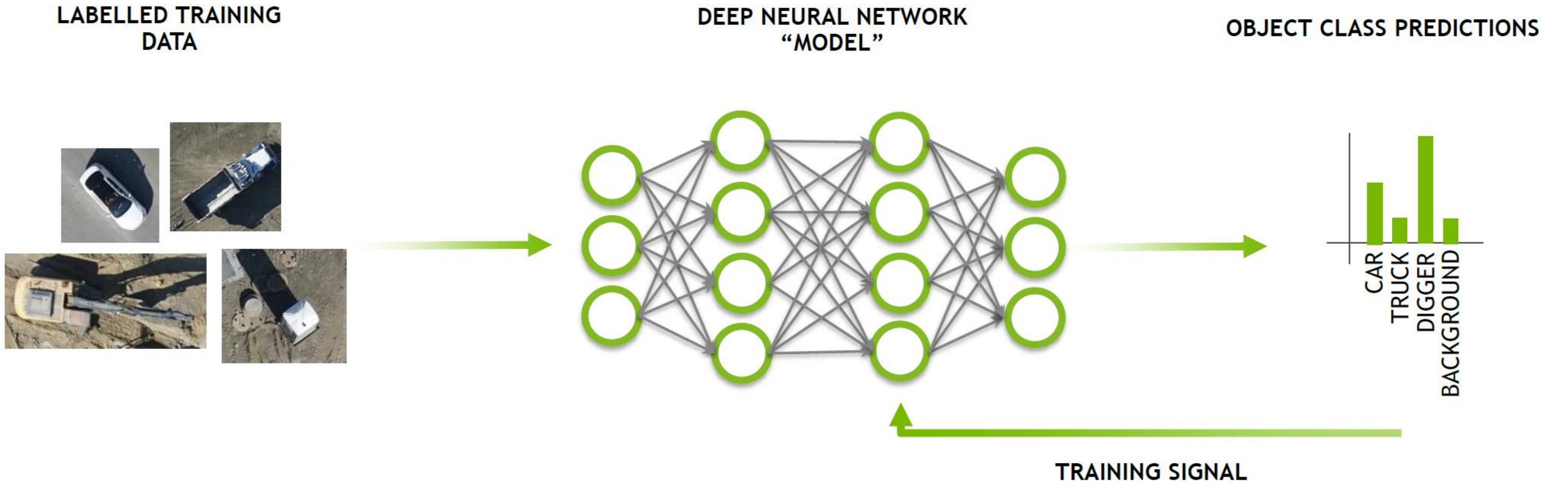
In the mathematical theory of artificial neural networks, the **universal approximation theorem** states[1] that a feed-forward network with a single hidden layer containing a finite number of neurons (i.e., a multilayer perceptron), can approximate continuous functions on compact subsets of $\mathbf{R}^n$, under mild assumptions on the activation function. The theorem thus states that simple neural networks can *represent* a wide variety of interesting functions when given appropriate parameters; however, it does not touch upon the algorithmic learnability of those parameters.





Why neural networks?

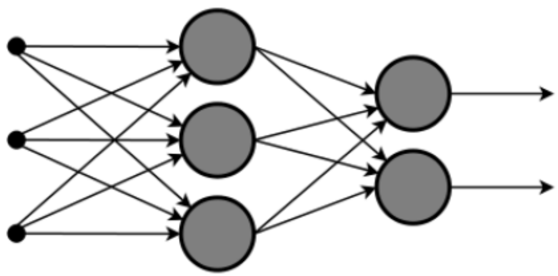
- It can learn…



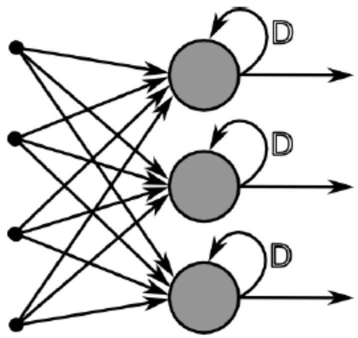
- Given sufficient training data an artificial neural network can approximate very complex functions mapping raw data to output decisions

Why neural networks?

- There can be lots of variations (layouts)…



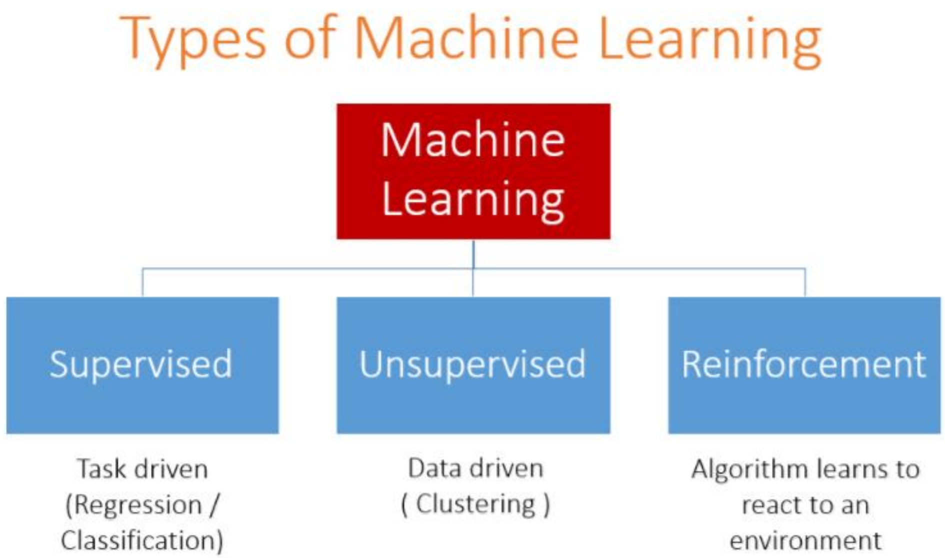
Feed Forward Neural Network



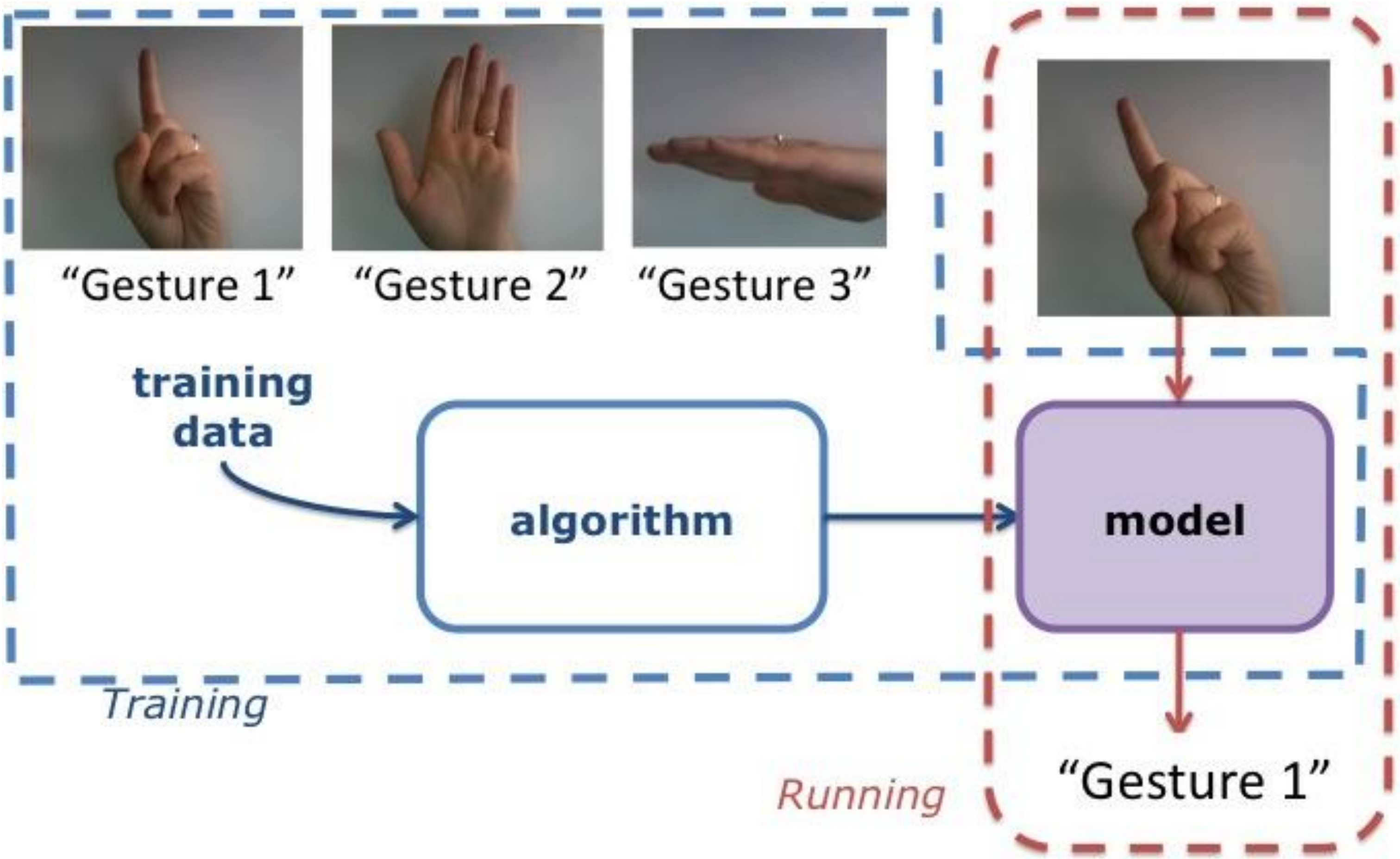
Recurrent Neural Network

Types of Machine Learning

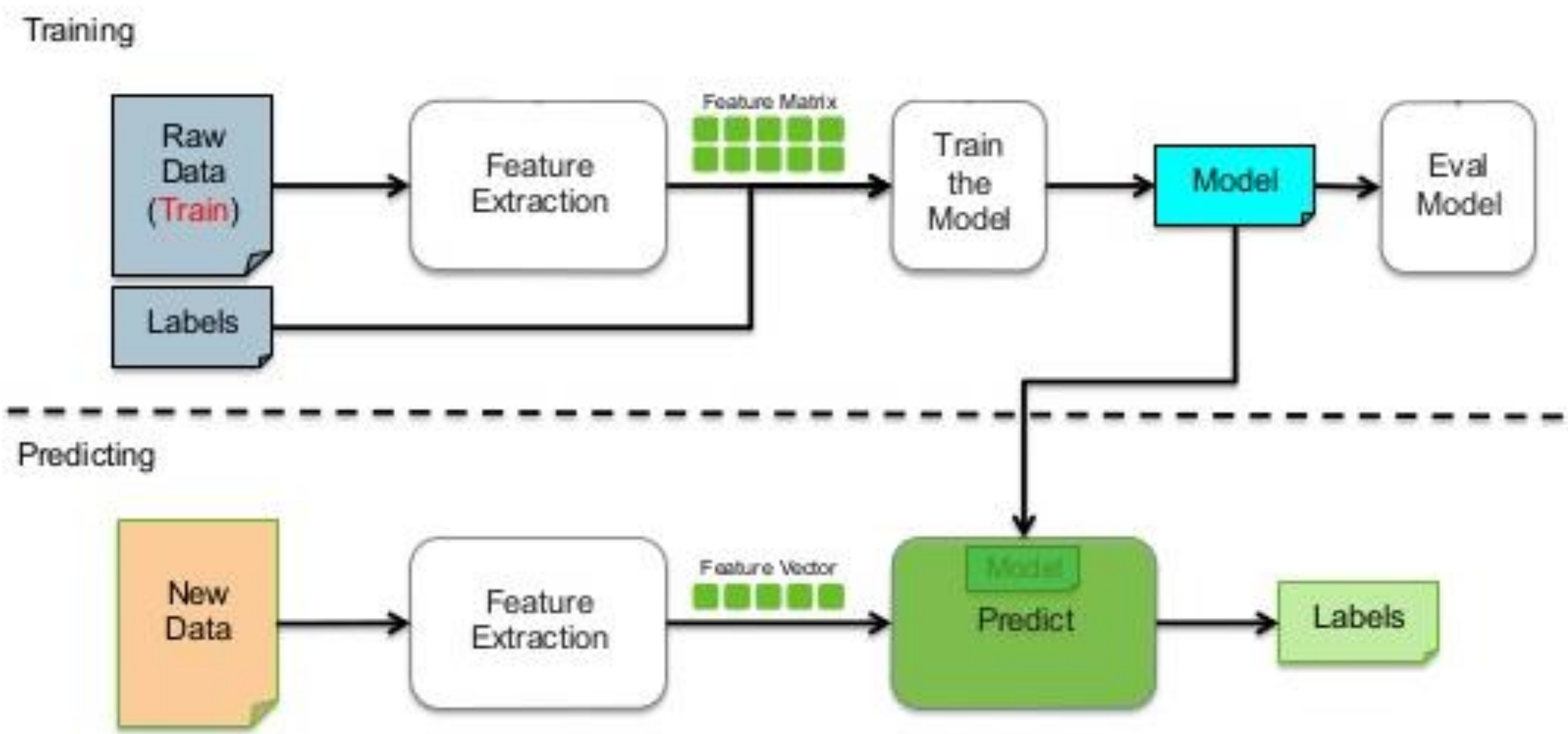
- Supervised Learning
  - Classification/Regression
  - Semi-supervised Learning/Weakly supervised Learning/…
- Unsupervised Learning
  - Clustering
  - Feature Learning
  - Generative Model Learning
- Reinforcement Learning
  - Deep Q-Learning
  - Policy Gradient Learning



Supervised learning



Supervised learning workflow



Supervised vs unsupervised

Supervised Learning

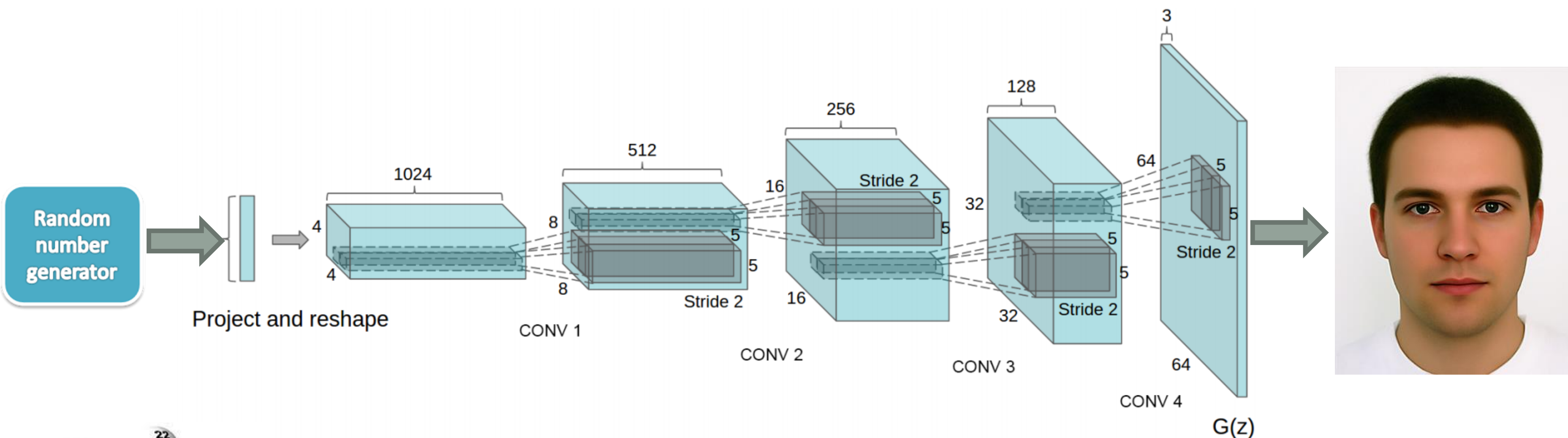
- **Data**: (x, y)
  - x is data, y is label
- **Goal**:
  - Learn a *function* to map x -> y
- **Examples**:
  - Classification, regression, object detection, semantic segmentation, image captioning, etc

Unsupervised Learning

- **Data**: x
  - Just data, no labels!
- **Goal**:
  - Learn some *structure* of the data
- **Examples**:
  - Clustering, dimensionality reduction, feature learning, generative models, etc

Unsupervised Learning

- Generative Model (Generative Adversarial Network)

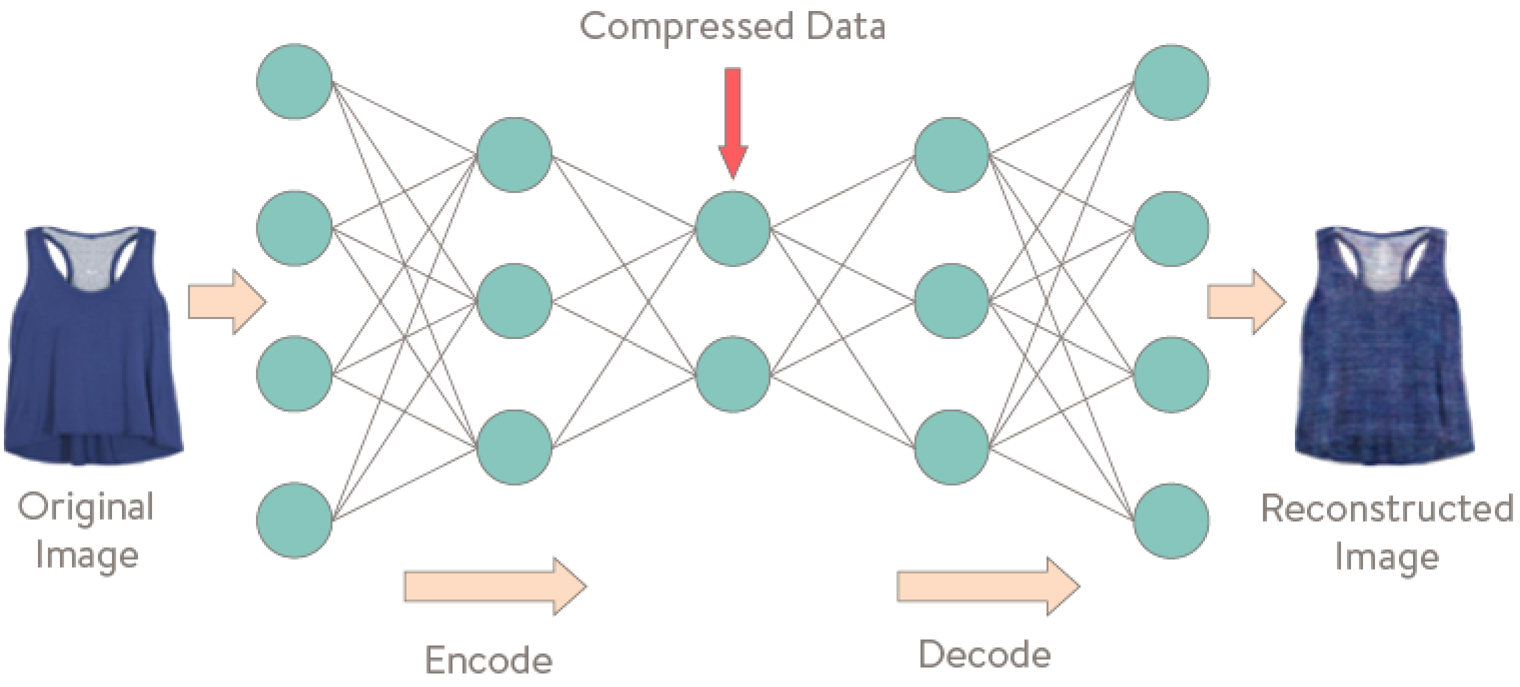


Unsupervised Learning

- Generative Model (Generative Adversarial Network)

Unsupervised Learning

- Dimension Reduction/Feature Learning (Auto-Encoder)



Supervised vs Reinforcement

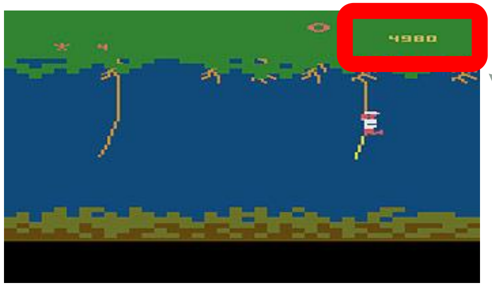
Supervised Learning

- **Data**: (x, y)
  - x is data, y is label
- **Goal**:
  - Learn a *function* to map x -> y
- **Examples**:

$(x_i, y_i) = ($  ,  $)$

Game state   Joystick control

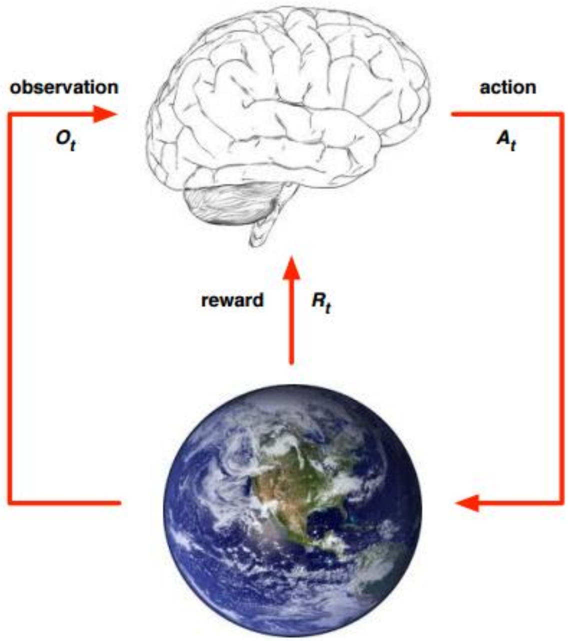
Reinforcement Learning



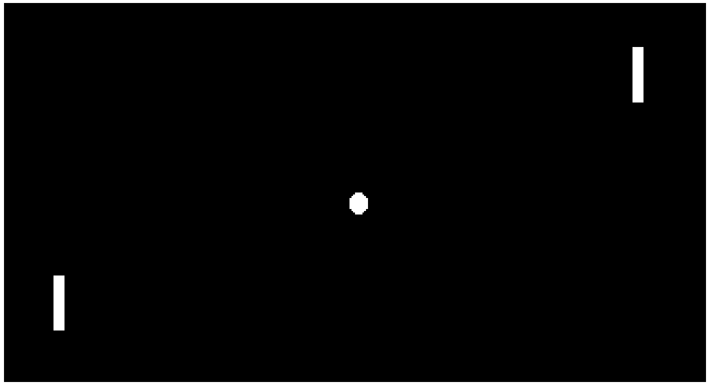
reward

Reinforcement Learning

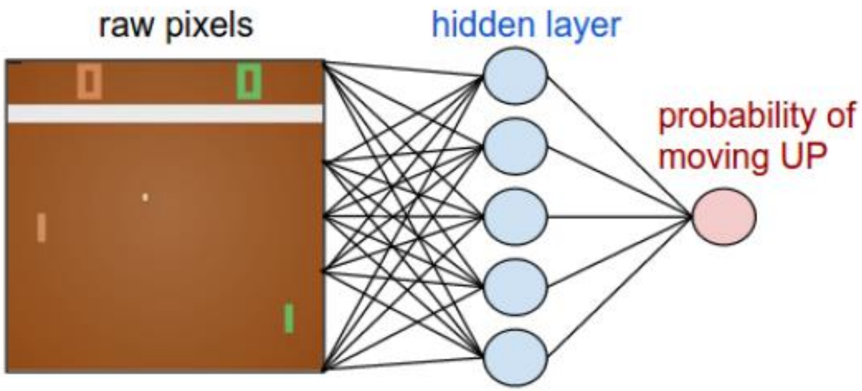
- Reinforcement learning is an area of machine learning concerned with how software **agents** ought to take actions in an environment so as to maximize some notion of cumulative reward.



Reinforcement Learning



Policy Network:



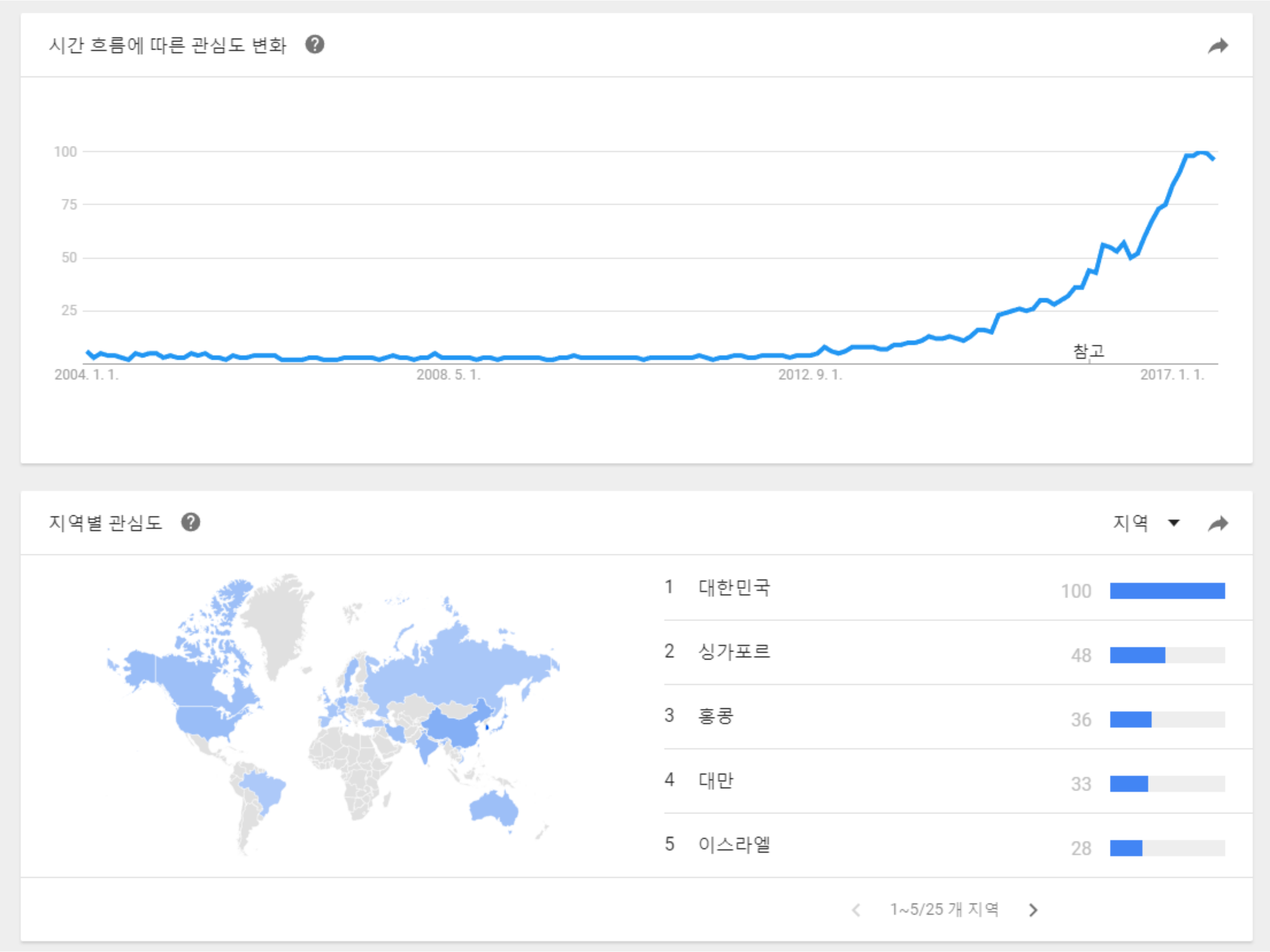
- 80x80 image (difference image)
- 2 actions: up or down
- 200,000 Pong games

This is a step towards general purpose artificial intelligence!

DEEP LEARNING

<table>
<tr><td colspan="5" rowspan="4"></td><th colspan="6">Tasks</th></tr>
<tr><th colspan="5">ADAS</th><th></th></tr>
<tr><th colspan="3">Self Driving</th><th></th><th></th><th></th></tr>
<tr><th>Localization</th><th>Perception</th><th>Planning/ Control</th><th>Driver state</th><th>Vehicle Diagnosis</th><th>Smart factory</th></tr>
<tr><td rowspan="7">Methods</td><td rowspan="2">Traditional</td><td colspan="3">Non-machine Learning</td><td>GPS, SLAM</td><td></td><td>Optimal control</td><td></td><td></td><td></td></tr>
<tr><td rowspan="6">Machine-Learning based method</td><td rowspan="4">Supervised</td><td>SVM
MLP</td><td></td><td>Pedestrian detection (HOG+SVM)</td><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="5">Deep-Learning based</td><td>CNN</td><td></td><td>Detection/ Segmentation/Classification</td><td>End-to-end Learning</td><td></td><td></td><td></td></tr>
<tr><td>RNN (LSTM)</td><td></td><td>Dry/wet road classification</td><td>End-to-end Learning</td><td>Behavior Prediction/ Driver identification</td><td></td><td>*</td></tr>
<tr><td>DNN</td><td></td><td></td><td></td><td></td><td>*</td><td>*</td></tr>
<tr><td colspan="2">Reinforcement</td><td></td><td></td><td>*</td><td></td><td></td><td></td></tr>
<tr><td colspan="2">Unsupervised</td><td></td><td></td><td></td><td></td><td></td><td>*</td></tr>
</table>

구글 트렌드: 딥러닝

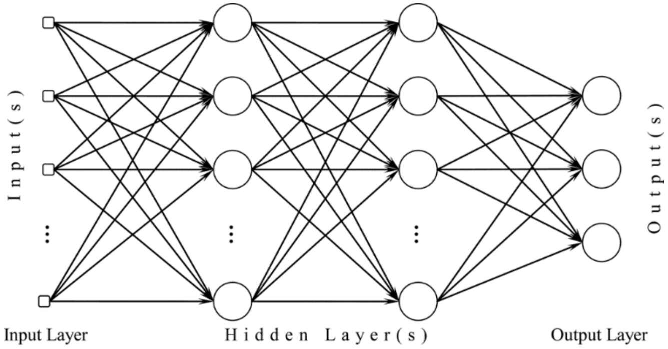


시간 흐름에 따른 관심도 변화
100
75
50
25
참고
2004. 1. 1.
2008. 5. 1.
2012. 9. 1.
2017. 1. 1.
지역별 관심도
지역
1 대한민국 100
2 싱가포르 48
3 홍콩 36
4 대만 33
5 이스라엘 28
1~5/25 개 지역

Neural network
Back propagation,
Nature

1986

- 장점
  - 일반적인 문제에 적용할 수 있는 학습법
  - Biological 시스템과 관련이 깊음
- 문제점
  - Training 이 쉽지 않음
  - 현실적인 문제에 잘 동작하지 않음



AI Winters

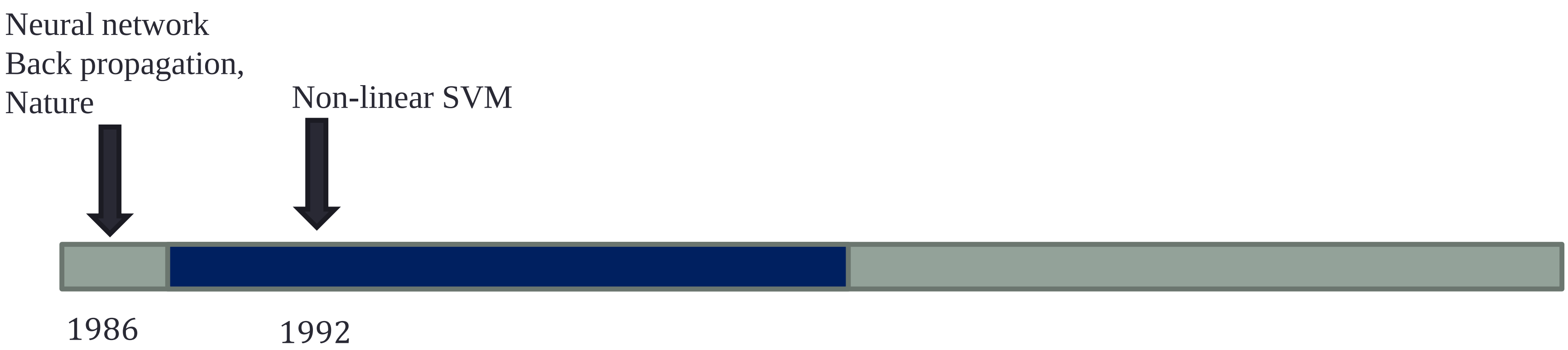
Two major episodes:

- 1974-80
- 1987-93

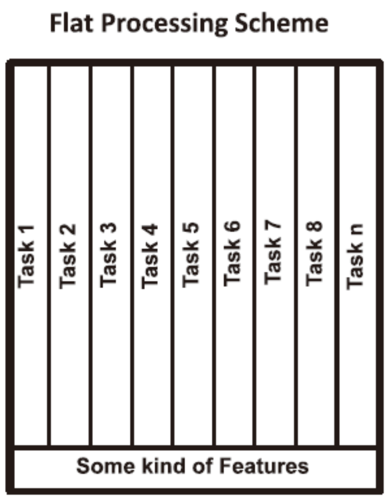
Smaller episodes:

- 1966: the failure of machine translation
- 1970: the abandonment of connectionism
- 1971-75: DARPA's frustration with the Speech Understanding Research program
- 1973: the large decrease in AI research in the UK in response to the Lighthill report.
- 1973–74: DARPA's cutbacks to academic AI research in general
- 1987: the collapse of the Lisp machine market
- 1988: the cancellation of new spending on AI by the Strategic Computing Initiative
- 1993: expert systems slowly reaching the bottom
- 1990s: the quiet disappearance of the fifth-generation computer project's original goals.

"In no part of the field have discoveries made so far produced the major impact that was then promised."



- 다양한 시도들
  - Flat structure
    - SVM, Boosting, …
  - Biological 시스템과 거리가 생김
  - 특정한 문제를 해결하는 특정한 방법 (SIFT, LBP, HOG, GMM-HMM)

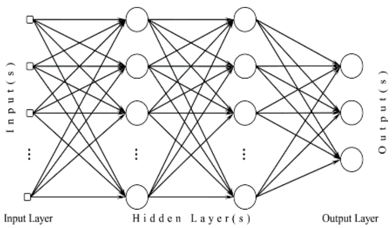


Neural network
Back propagation,
Nature

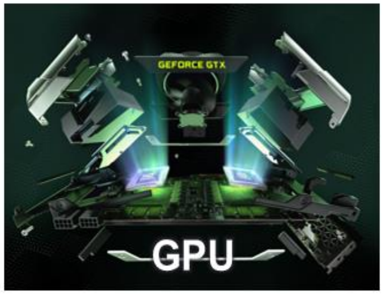
Deep belief network,
Science

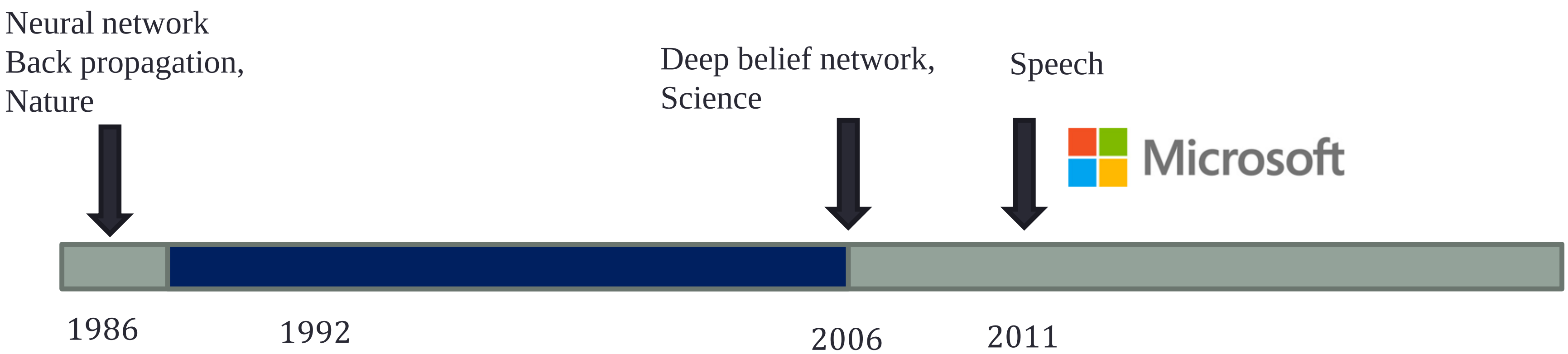
1986 1992 2006

- 비지도 학습을 이용한 pre-training
- Training 방법 향상
  - Dropout, RectLinear, Normalization, …
- 컴퓨터 구조의 발달
  - GPU
  - Multi-core computer 시스템
- 빅데이터



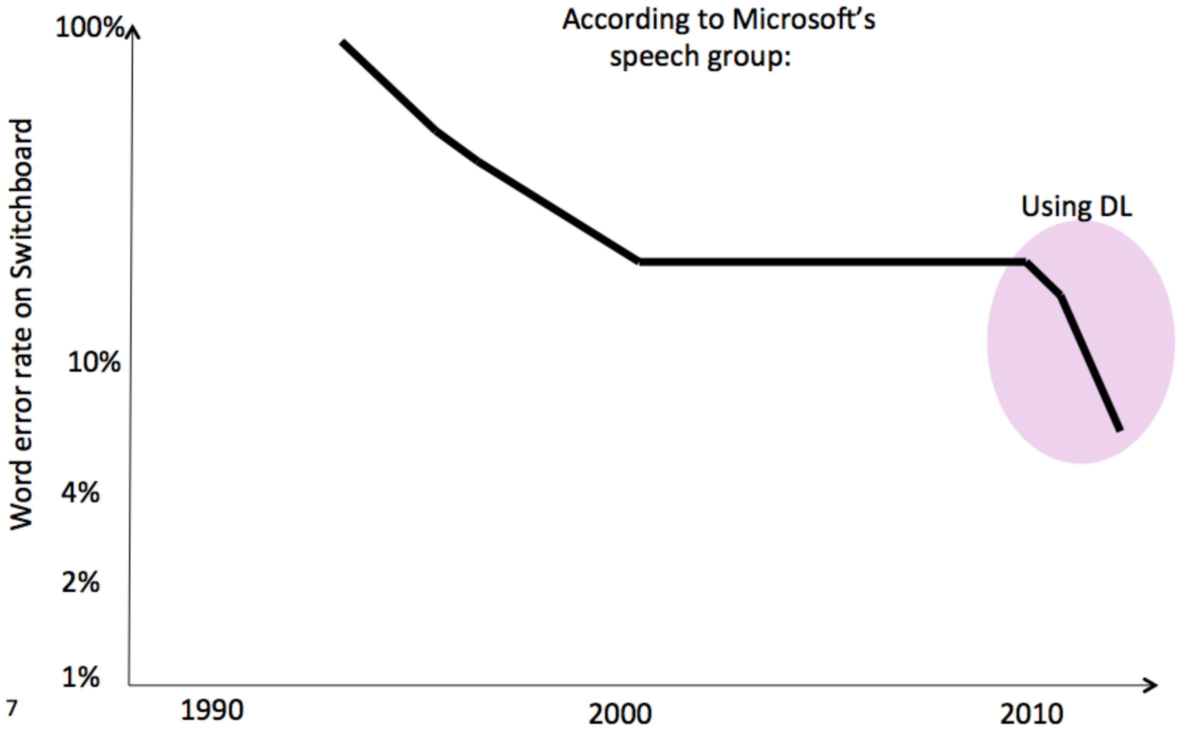
+





task	hours of training data	DNN-HMM	GMM-HMM with same data	GMM-HMM with more data
Switchboard (test set 1)	309	18.5	27.4	18.6 (2000 hrs)
Switchboard (test set 2)	309	16.1	23.6	17.1 (2000 hrs)
English Broadcast News	50	17.5	18.8	
Bing Voice Search (Sentence error rates)	24	30.4	36.2	
Google Voice Input	5,870	12.3		16.0 (>>5,870hrs)
Youtube	1,400	47.6	52.3	

음성 인식 성능


According to Microsoft's
speech group:
Using DL
Word error rate on Switchboard
100%
10%
4%
2%
1%
1990
2000
2010

HOME PAGE | TODAY'S PAPER | VIDEO | MOST POPULAR | U.S. Edition

SUBSCRIBE NOW | Log In | Register Now | Help

The New York Times

Science

Search All NYTimes.com

Go

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION | ARTS | STYLE | TRAVEL | JOBS | REAL ESTATE | AUTOS

ENVIRONMENT | SPACE & COSMOS

Scientists See Promise in Deep-Learning Programs

Hao Zhang/The New York Times

A voice recognition program translated a speech given by Richard F. Rashid, Microsoft's top scientist, into Mandarin Chinese.

By JOHN MARKOFF

Published: November 23, 2012

FACEBOOK

TWITTER

GOOGLE+

SAVE

Using an artificial intelligence technique inspired by theories about how the brain recognizes patterns, technology companies are reporting startling gains in fields as diverse as computer vision, speech recognition and the identification of promising new molecules for designing drugs.

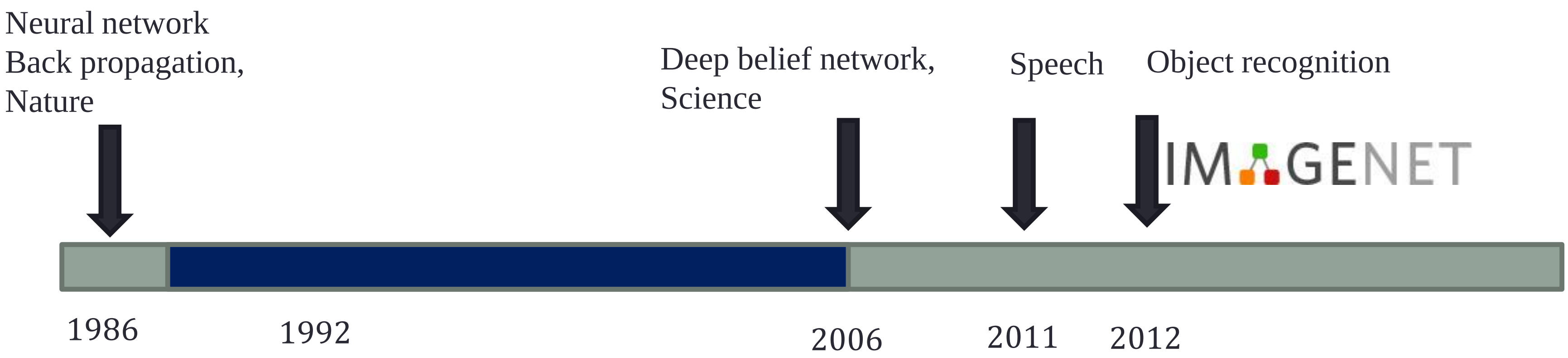
MOST EMAILED | MOST VIEWED

1. THOMAS L. FRIEDMAN
How to Get a Job at Google, Part 2

2. OP-ED CONTRIBUTOR
The Public Health Crisis Hiding in Our Food

3. 50 Years Into the War on Poverty, Hardship Hits Back

4. Taking On Adam Smith (and Karl Marx)



Submission	Method	Error rate
Supervision	Deep CNN	0.16422
ISI	FV: SIFT, LBP, GIST, CSIFT	0.26172
XRCE/INRIA	FV: SIFT and color 1M-dim features	0.27058
OXFORD_VGG	FV: SIFT and color 270K-dim features	0.27302

ImageNet Large Scale Visual Recognition Competition (ILSVRC)

Steel drum

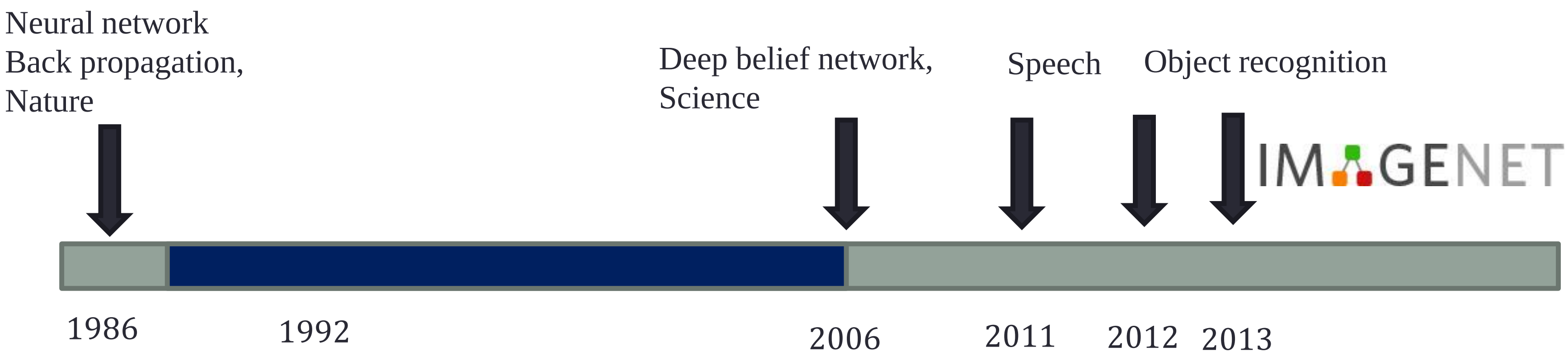
Output:
Scale
T-shirt
<u>Steel drum</u>
Drumstick
Mud turtle

✔

Output:
Scale
T-shirt
Giant panda
Drumstick
Mud turtle

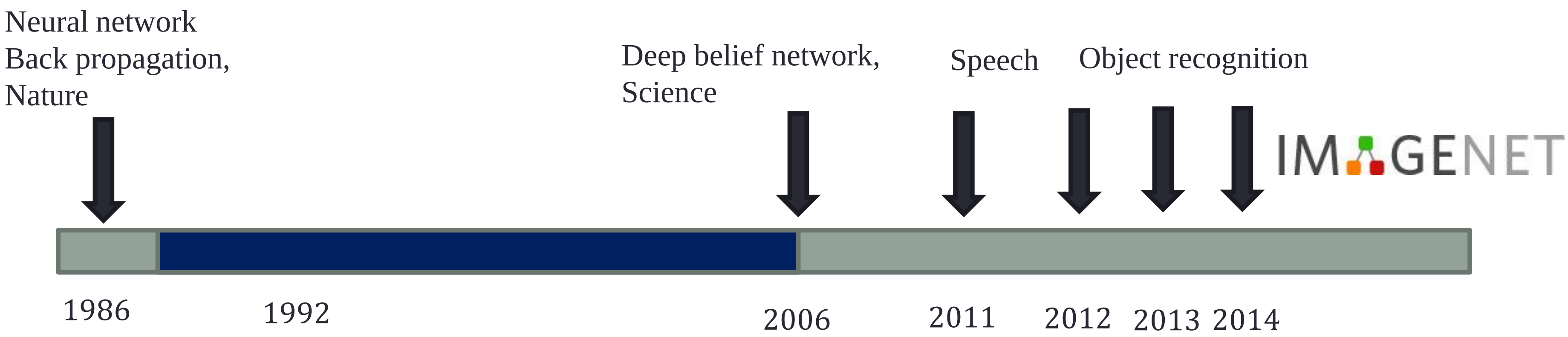
✗

http://www.image-net.org/challenges/LSVRC/



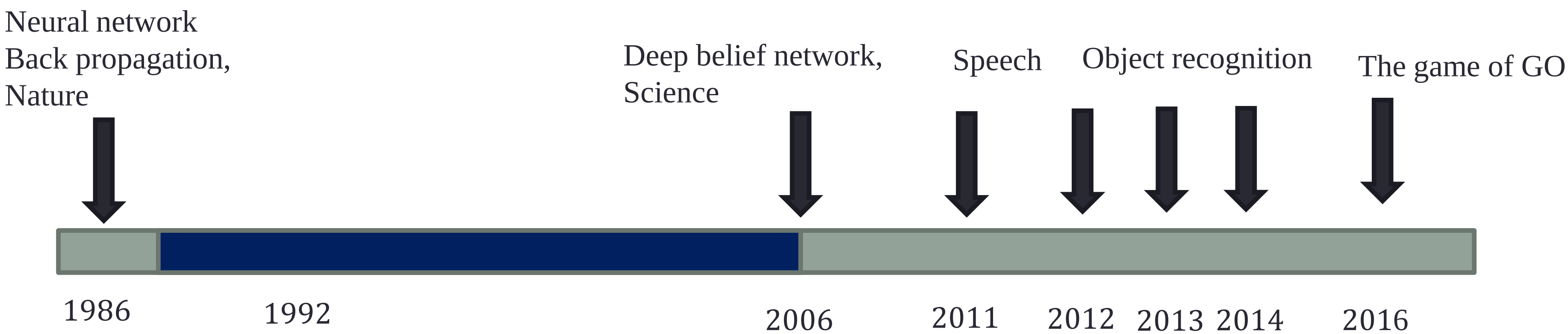
- IMAGENET 2013: 영상 인식

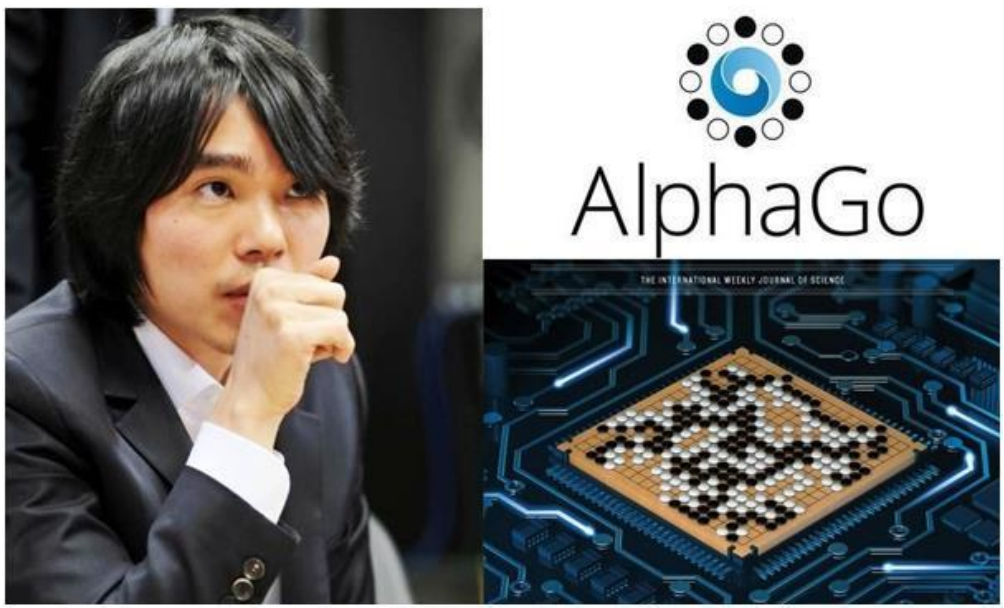
RANK	Name	Error rate	Description
1	NYU	0.11197	Deep Learning
2	NUS	0.12535	Deep Learning
3	OXFORD	0.13555	Deep Learning



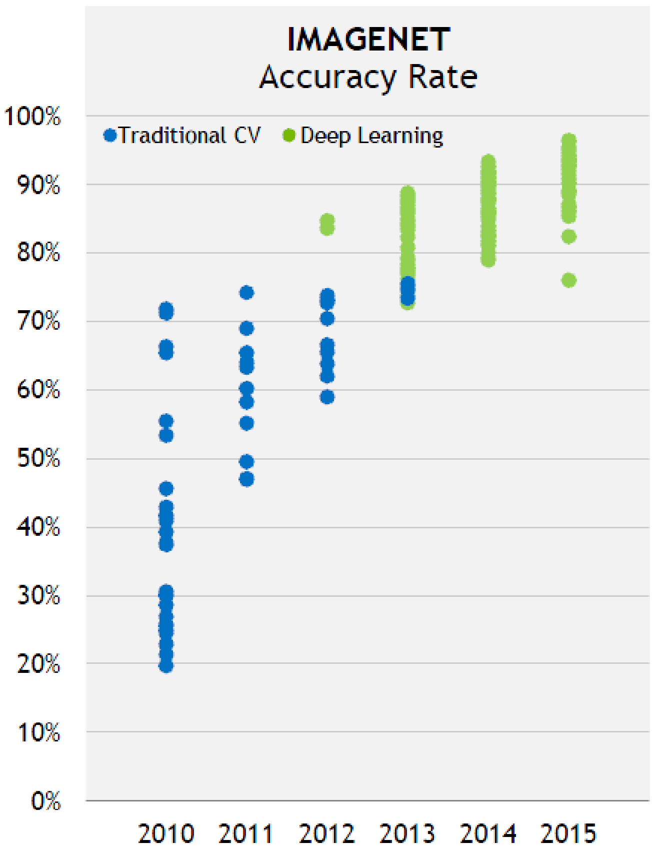
- IMAGENET 2013: 영상 인식

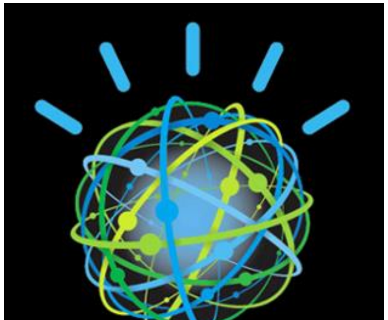
RANK	Name	Error rate	Description
1	Google	0.06656	Deep Learning
2	Oxford	0.07325	Deep Learning
3	MSRA	0.08062	Deep Learning


Neural network
Back propagation,
Nature
Deep belief network,
Science
Speech
Object recognition
The game of GO
1986
1992
2006
2011
2012
2013
2014
2016


AlphaGo

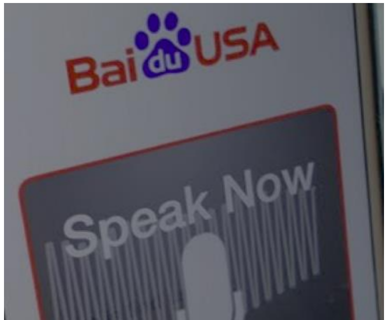
The AI race is on





IBM Watson Achieves Breakthrough
In Natural Language Processing

Facebook
Launches Big Sur



Baidu Deep Speech 2
Beats Humans

TensorFlow

Google
Launches TensorFlow

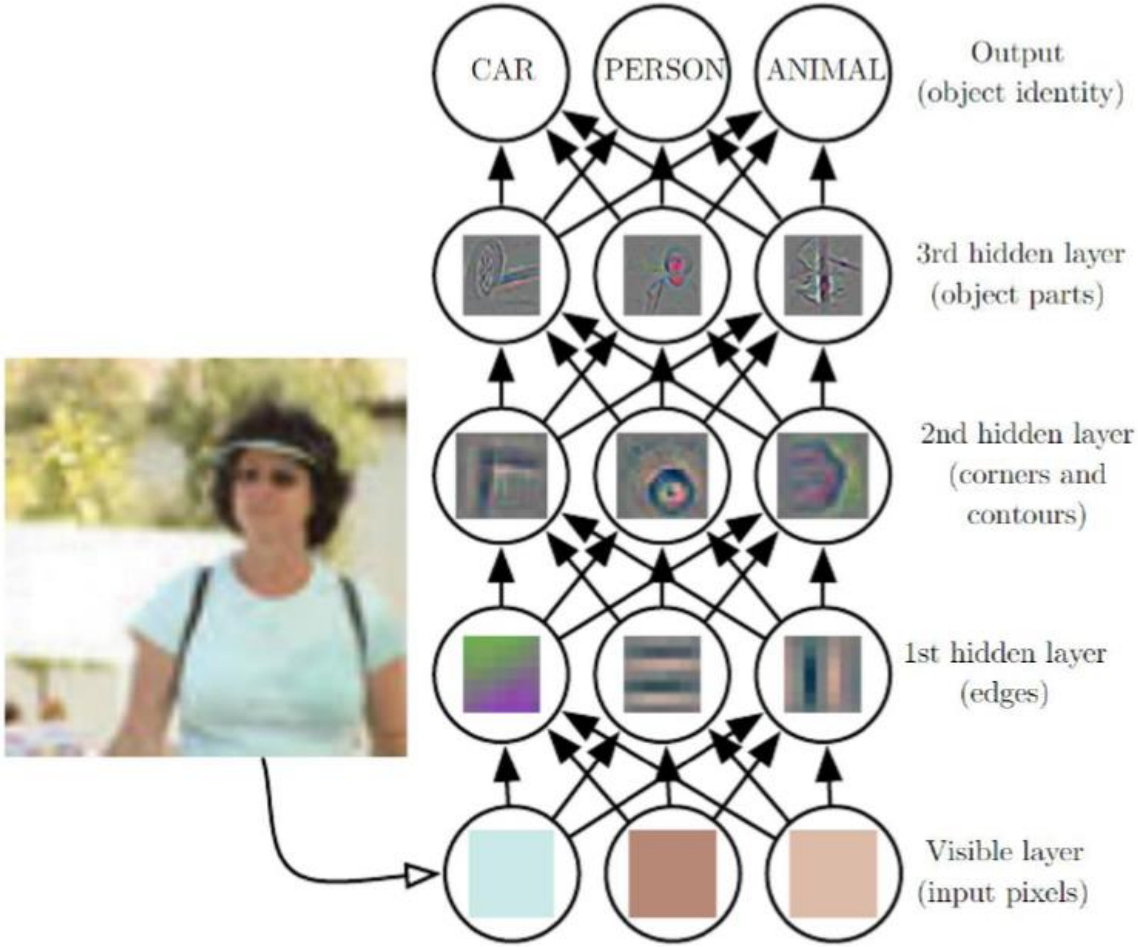
TOYOTA
RESEARCH INSTITUTE

Toyota Invests $1B
In AI Labs

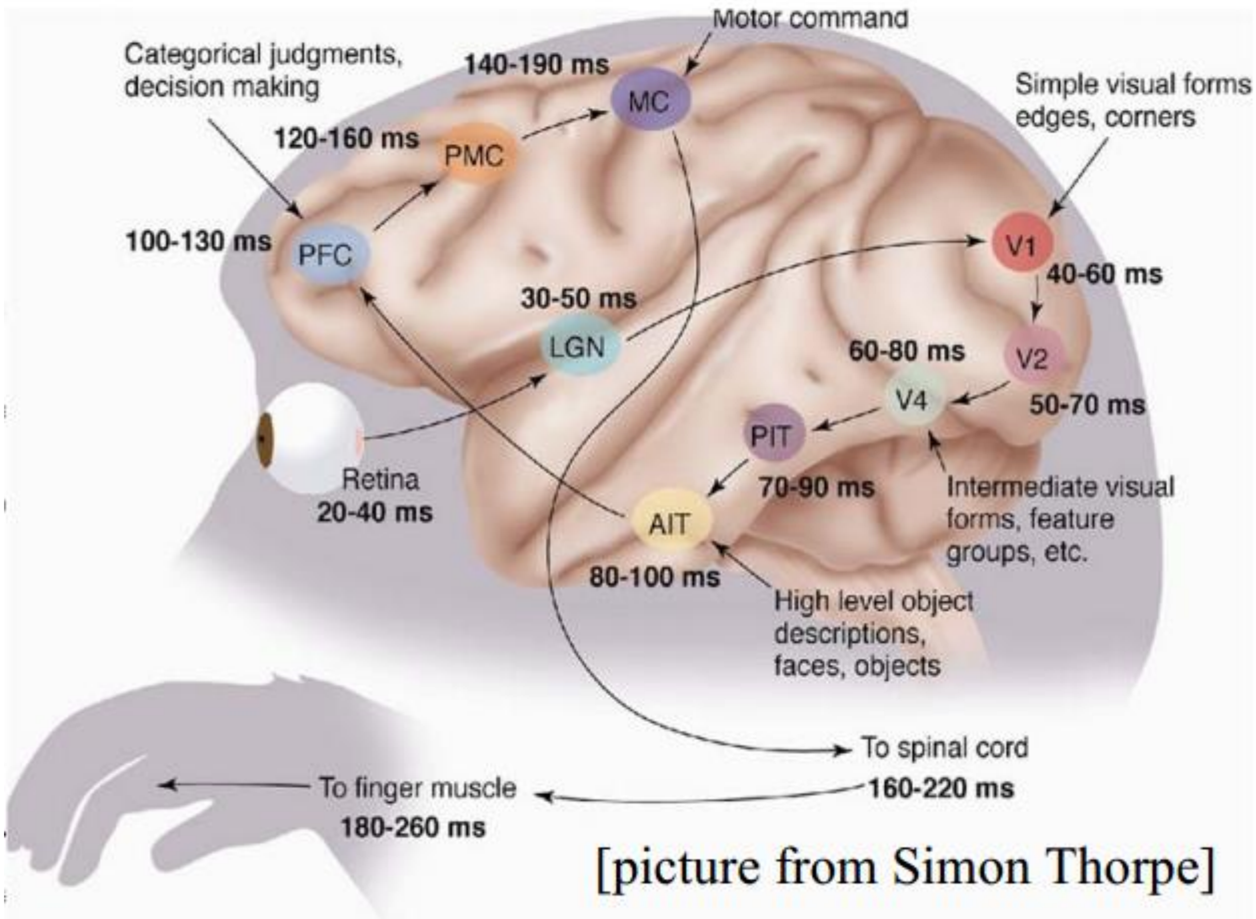
Microsoft

Microsoft & U.Science & Tech, China
Beat Humans on IQ

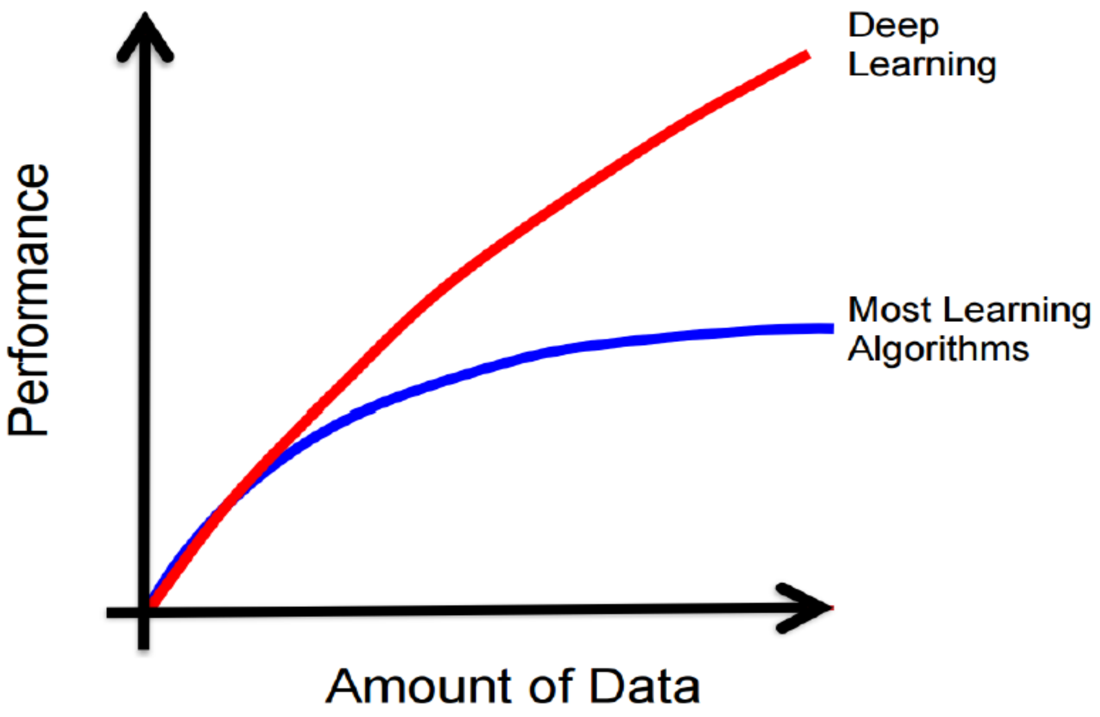
Deep Learning: Representation Learning



The Mammalian Visual Cortex is Hierarchical

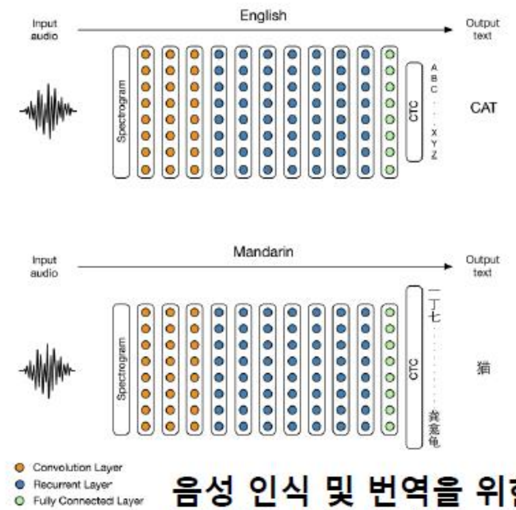


Deep Learning: Scalable Machine Learning

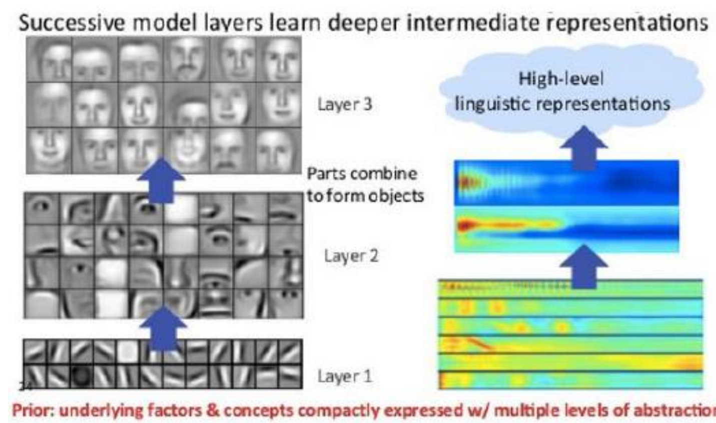


딥러닝 모델의 특징

- 다층구조 (multi layer)
  - 신경망의 구조 모사
  - 상위 층으로 갈 수록 추상화된 정보가 학습과정에서 자동으로 생성
- 문제 해결과정자동화
  - End-to-end learning
  - 사람의 개입을 배제하고 오직 raw input과 output 사이에 모든 과정을 데이터에서 학습하는 방향 추구
- 분산 표현
  - Distributed representation
  - 여러 뉴런이 협력하여 정보 저장/처리



음성 인식 및 번역을 위한 딥러닝 모델 사례(Baidu)



DRAWBACKS

Current Drawbacks

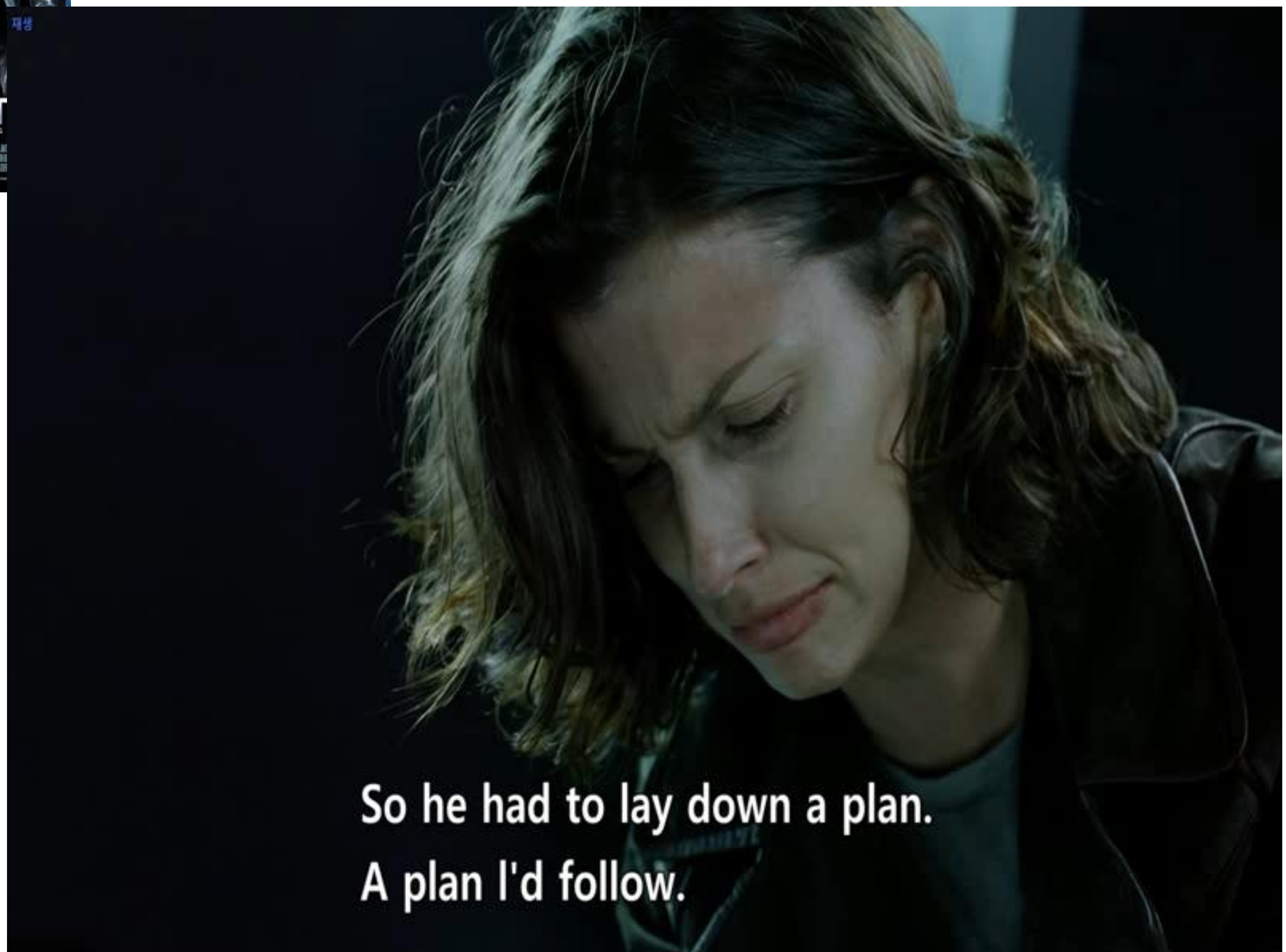
- Big data: inefficient at learning from data
- Supervised data: costly to annotate real-world data
- Need to manually select network structure
- Need to hyper-parameter tuning
  - Learning rate
  - Loss function
  - Mini-batch size
  - Number of training iterations
  - Momentum
  - Optimizer selection
- Defining a good reward function is difficult

Faulty Reward Functions in the Wild


WILL SMITH
I,ROBOT
ONE MAN SAW IT COMING


Unless the killer's still in here.

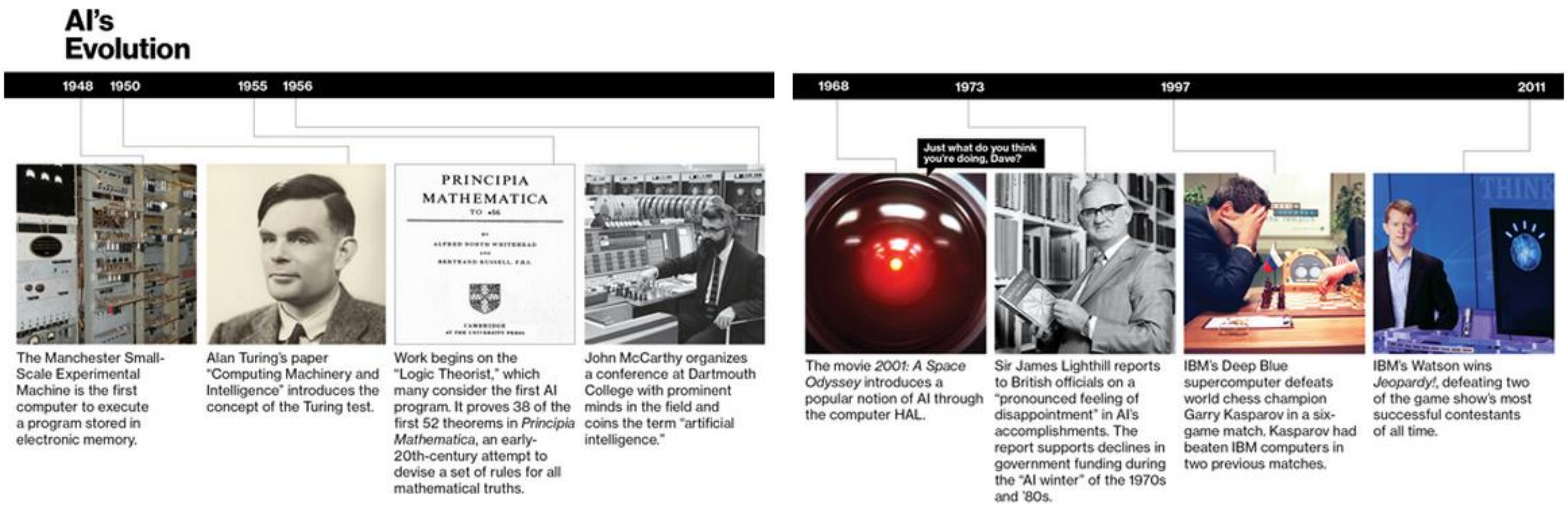

WILL SMITH
I,ROBOT
ONE MAN SAW IT COMING


So he had to lay down a plan.
A plan I'd follow.

AI 역사

AI 역사

MIT Technology Review

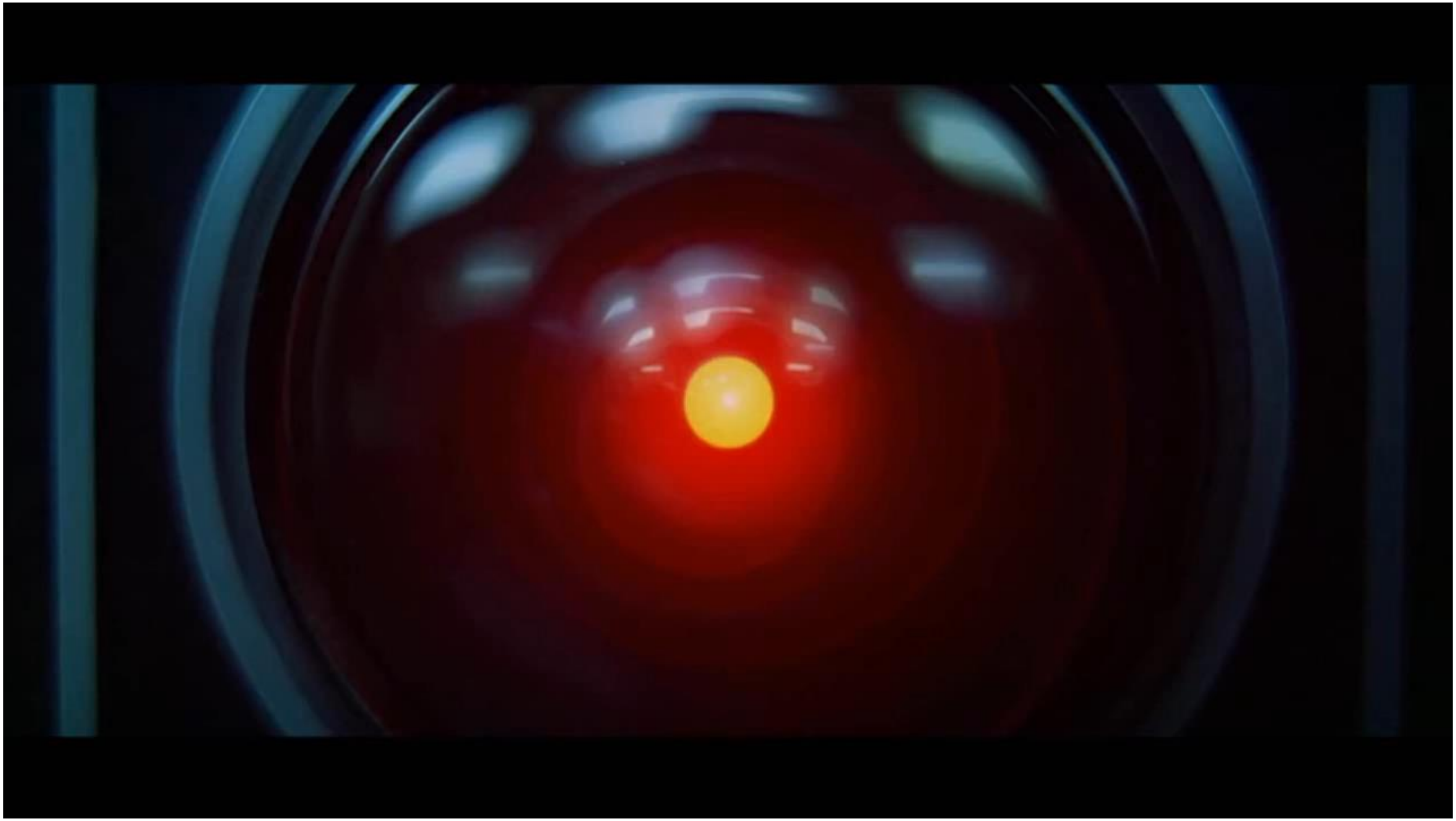



An epic drama of
adventure and exploration
Man's colony on the Moon…a whole new generation has been born and is living here…a quarter-million miles from Earth.
MGM PRESENTS A STANLEY KUBRICK PRODUCTION
2001: a space odyssey
Super Panavision and Metrocolor


WRITTEN AND DIRECTED BY
ALEX GARLAND THE WRITER OF
28 DAYS LATER
EX_MACHINA
THERE IS NOTHING MORE HUMAN THAN THE WILL TO SURVIVE
COMING SOON

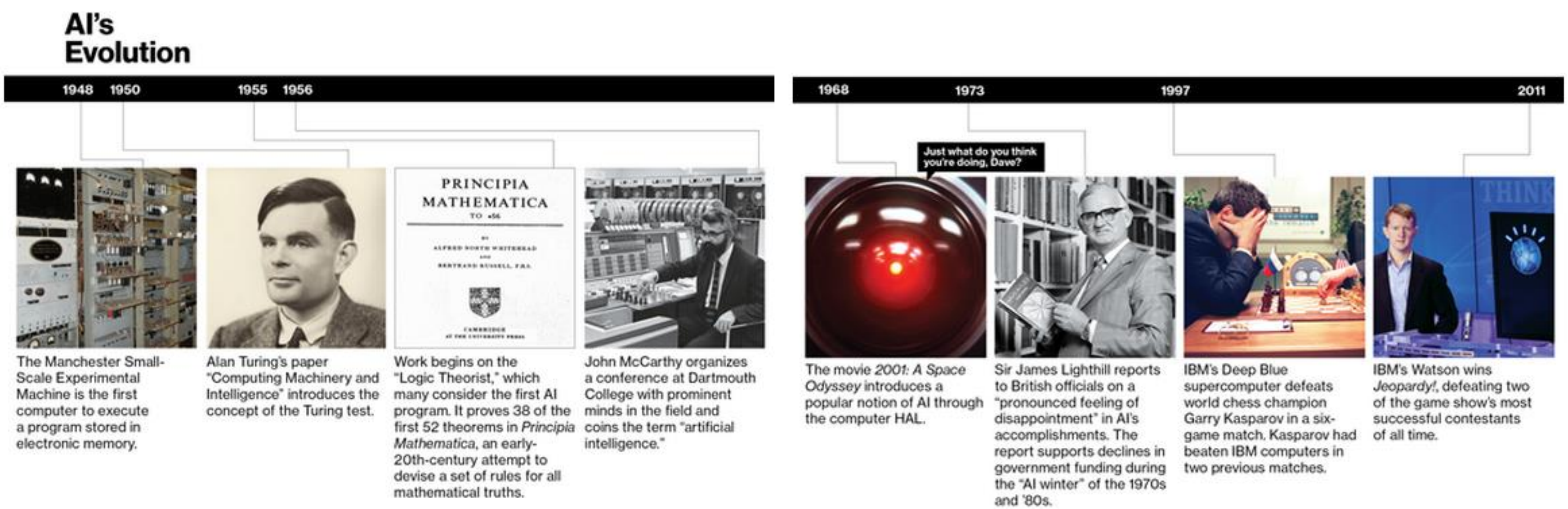

MOMMY
DADDY
CHAPPIE
CHAPPiE

I'm sorry Dave.



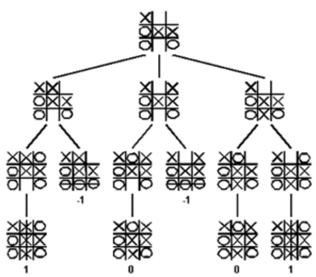
AI 역사

MIT Technology Review



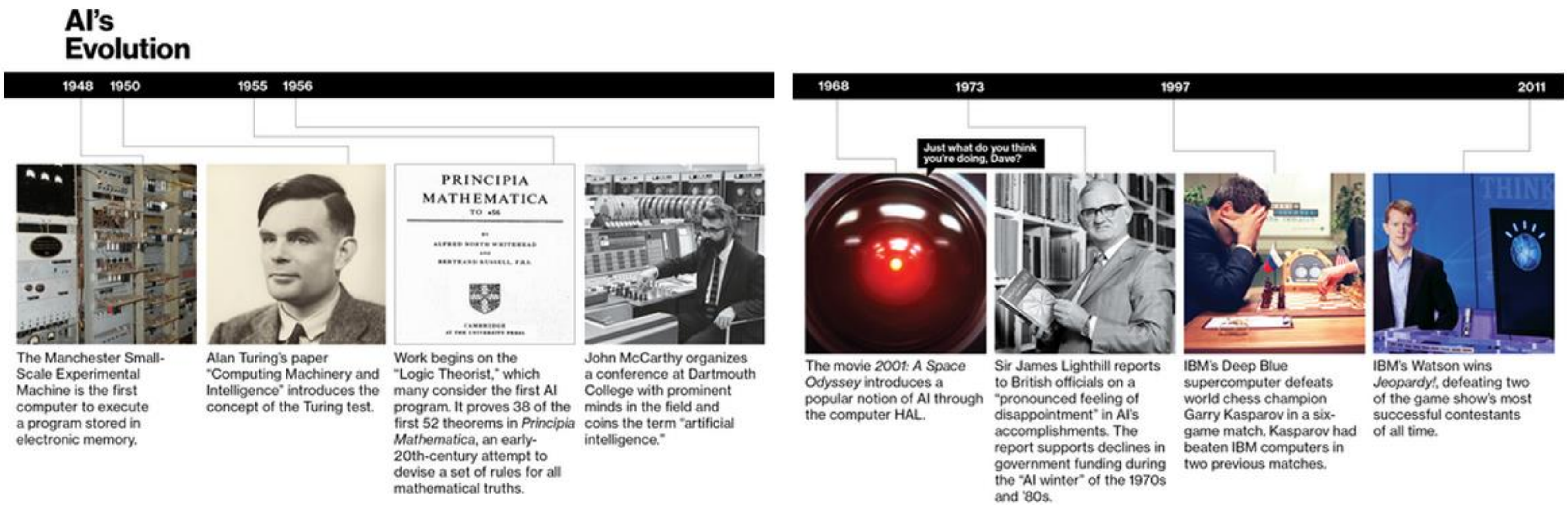
딥블루

- **딥블루** vs **게리 카스파로프**, 1997
  - Deep Blue vs Kasparov
    - $3\frac{1}{2}$ vs $2\frac{1}{2}$
  - Brute-force search power
    - 6~8 **수를 내다봄**



AI 역사

MIT Technology Review

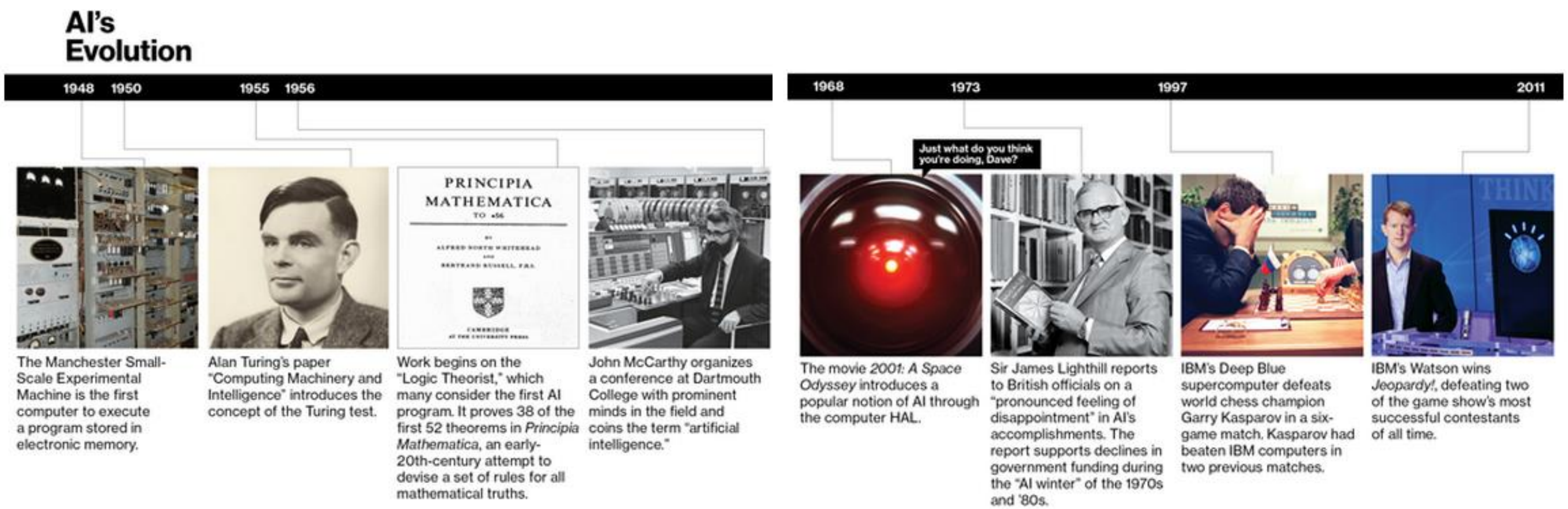


IBM Watson 슈퍼컴퓨터

- **질문예시)**
  - **Kathleen kenyon's excavation of this city mentioned in joshua showed the wall had been repaired 17 times**
    - WHAT is "Jericho"
  - **This child star got his first on-screen kiss in "MY GIRL"**
    - WHO is "Macaulay Culkin"

AI 역사

MIT Technology Review



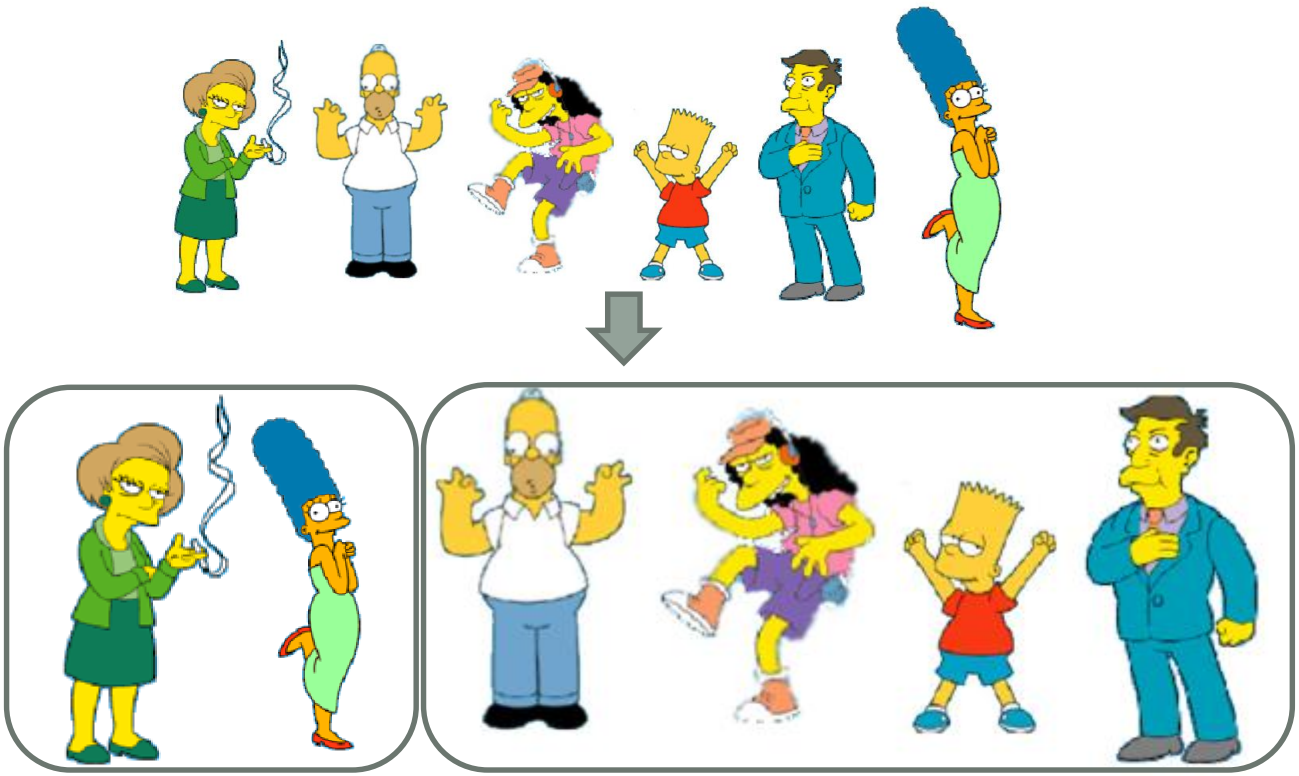
2016

BACKUPS

CLUSTERING

Unsupervised Learning

- Clustering
  - To group the object having same feature

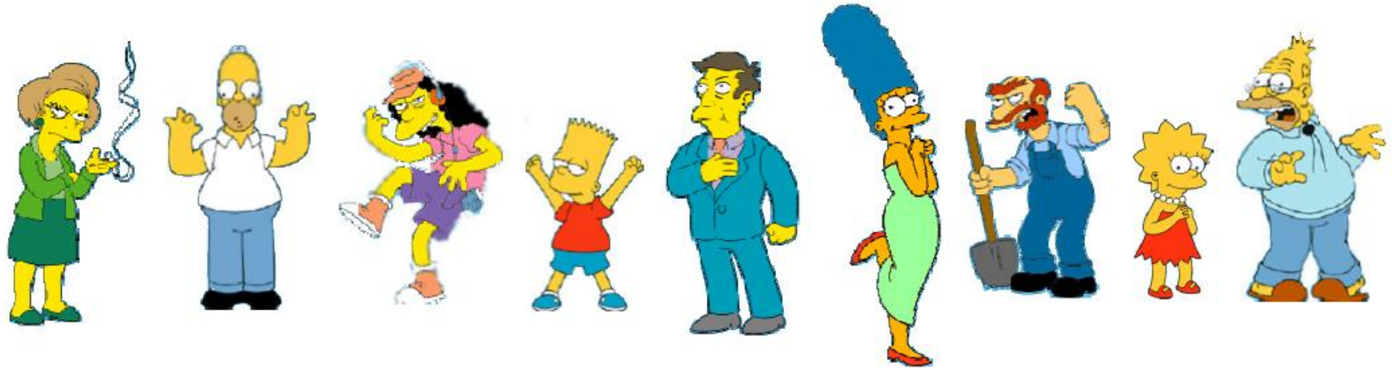


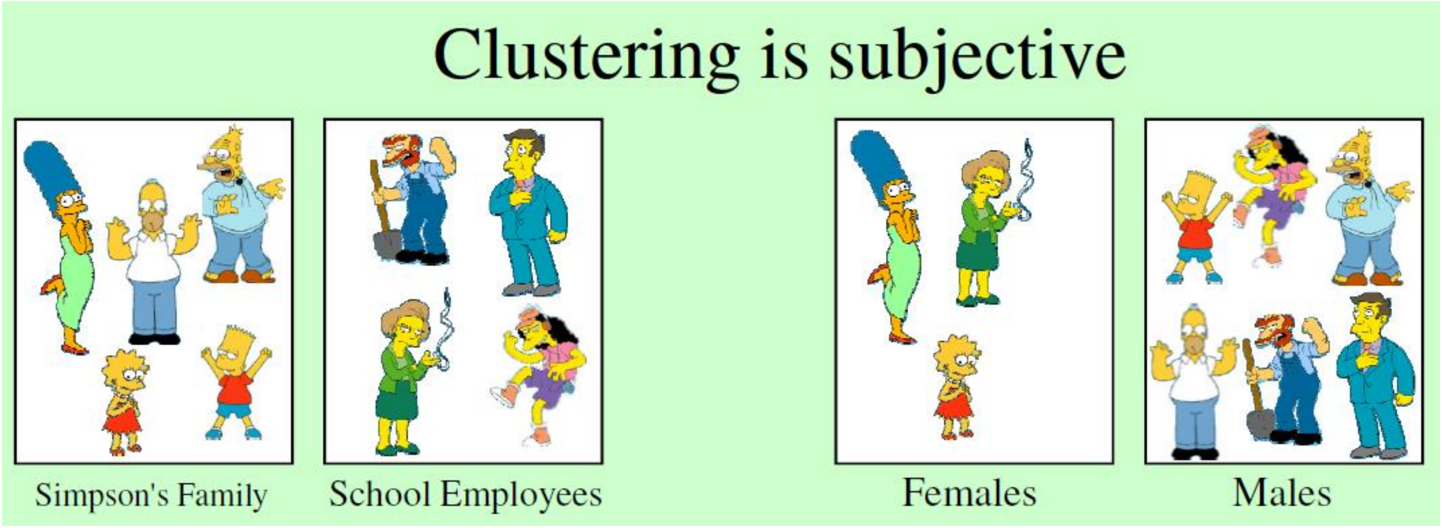
Female group

male group

Unsupervised Learning

- Clustering





Clustering Example