

ICT

머신 러닝

Machine Learning

ICT 또는 IT

Information

Machine Learning (ML)



컴퓨터에게

Information

뭔가를 가르쳐서
잘하게 한다.

ML로 더 잘하게 된다는 것 (최적화)
=> 성장

Daum 루빅스



루빅스는 계속 진화한다. 축적된 데이터를 스스로 학습한다. 루빅스가 모바일 앱에 적용된 이후 뉴스 이용량이 늘고 제공되는 뉴스가 다양해졌다. 다음뉴스의 주간 이용자 수는 2015년 11월 1910만 명에서 2016년 11월 2710만명으로 늘었다. 다음 첫 화면의 뉴스 개수는 평균 3.5배 늘었다. IT 과학

https://biz.chosun.com/site/data/html_dir/2017/07/31/2017073100012.html

“알수없는 알고리즘이
나를 여기로..!!”, 유튜브



<https://www.zdnet.co.kr/view/?no=20180511150948>

인공지능 유용성

□ 감지,
추리,
행동,
적응하는 프로그램.



인텔의 정의

ARTIFICIAL INTELLIGENCE

A program that can sense, reason, act, and adapt

MACHINE LEARNING

Algorithms whose performance improve as they are exposed to more data over time

DEEP LEARNING

Subset of machine learning in which multilayered neural networks learn from vast amounts of data



인텔의 정의

ARTIFICIAL INTELLIGENCE

A program that can sense, reason,
act, and adapt

AI에는 학습기능이 없는 프로그램도 존재
(스마트폰 펌웨어도 다양한 AI가 제공되지만,
정작 Update는 프로그래머가 해줘야 함)

Sense (Recognition)

감지 필요!

□잃어버린 우리 강아지를 찾아주세요!



Reasoning

Action for Safety



인텔의 정의

ARTIFICIAL INTELLIGENCE

A program that can sense, reason,
act, and adapt

Sense (Recognition)

감지 필요!

□ 잃어버린 우리 강아지를 찾아주세요!



Reasoning

Action for Safety



인텔의 정의

ARTIFICIAL INTELLIGENCE

A program that can sense, reason, act, and adapt

Sense (Recognition)

감지 필요!

□ 잃어버린 우리 강아지를 찾아주세요!



Reasoning

추리: 문맥 정보의 포함

□ 아버지 가방에 들어가신다.

Papago

한국어 -



영어 -

아버지 가방에 들어가신다.

Father is entering the room.

Google

번역

도움말 | 약관 | 개인정보 | 서비스 | 회사 | 제휴 | 기타

검색결과 약 1,790,000,000개 (0.35초)

한국어 -

영어 -

아버지 가방에 들어가신다.

Enter your father's bag.



Action for Safety



인텔의 정의

ARTIFICIAL INTELLIGENCE

A program that can sense, reason, act, and adapt

Sense (Recognition) 감지 필요!

□ 잃어버린 우리 강아지를 찾아주세요!



Reasoning

추리: 문맥 정보의 포함

□ 아버지 가방에 들어가신다.

Papago

한국어 -

아버지 가방에 들어가신다.



영어 -

Father is entering the room.

Google

번역

한국어 -

아버지 가방에 들어가신다.

×

Enter your father's bag.



Action for Safety

불완전 기술 개선되도록

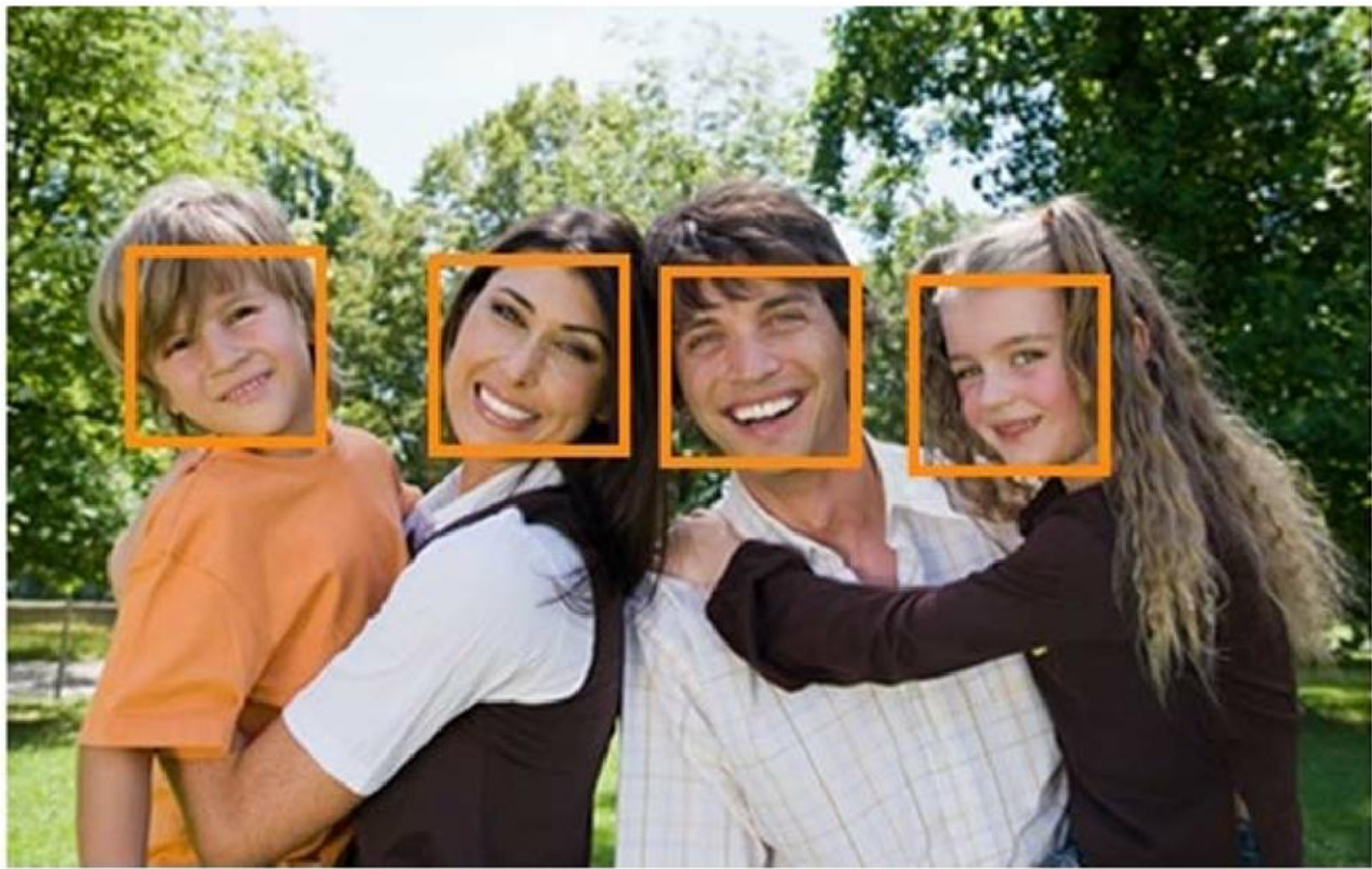
□ 검증 전에 이미 도입된 기술들

러닝(학습) 기능이 없는 AI라면
자동 주차, 자율 주행 등은 도입 불가



당시 '자동 주차 기능'을 시연하는 중이었습니다





인공지능 비서 챗봇 (예: 비행예약 봇)



사람 → 기계

실물 외관(HardWare)의 기계가 아닌
내재된(SoftWare)의 기계

‘ 사람처럼 학습능력을 가진 기계 ’

산업혁명과 주목된 능력

산업혁명 이전: 힘과 가내 수공 능력

1차, 증기발명: 단순 공업 기술

2차, 전기발명: 고도 공업 기술

3차, 컴퓨터발명 IT, 금융, 서비스: 고도 정보 기술
+ 심리 이해 능력(마케팅)

4차, 스마트폰+머신러닝 고도화 : 어떤 능력일까요?

4차 산업혁명

= 머신러닝 + 빅데이터 + 산업별 융합

= 데이터 과학

머신러닝

=

사람의 지적업무를 대신함

1811년부터 1817년

Luddite(**러다이트**)는 영국에서 1차 산업혁명이 초래할 실업 위험에 반대해 기계를 파괴하는 등 폭동을 일으킨 직공단원

노동자들은 실업과 생활고의 원인을 기계의 탓으로 돌리고 기계 파괴운동을 일으켰다.

노팅엄 직물공장에서 시작, 확산
1812년 기계 파괴자 = 사형
법안 통과의 계기.

source: 교양영어사전1

<https://terms.naver.com/entry.nhn?docId=1922117&cid=41810&categoryId=41811>



러다이트 운동

산업혁명

인간의노동 ➡ 증기기관

머신러닝

인간의지식 → 알고리즘

Information

머신 러닝

Machine Learning

빅 데이터

Big Data

데이터 마이닝

Data Mining

데이터 과학

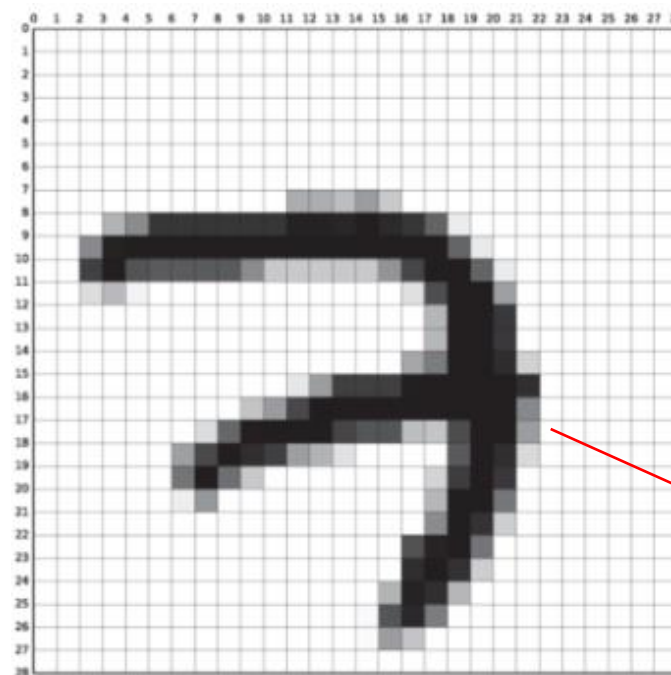
Data Science

Key concepts: Machine Learning

데이터로부터 '패턴을 학습'시키는 기법
[?] 알고리즘과 모델을 만드는 행위 자체

- 데이터의 양이 중요하지만,
적다고 불가능한 것은 아님

ex) 내 손 글씨를 보고,
컴퓨터가 어떤 숫자인지
맞추도록 가르치기



(a) MNIST sample belonging to the digit '7'.



(b) 100 samples from the MNIST training set.

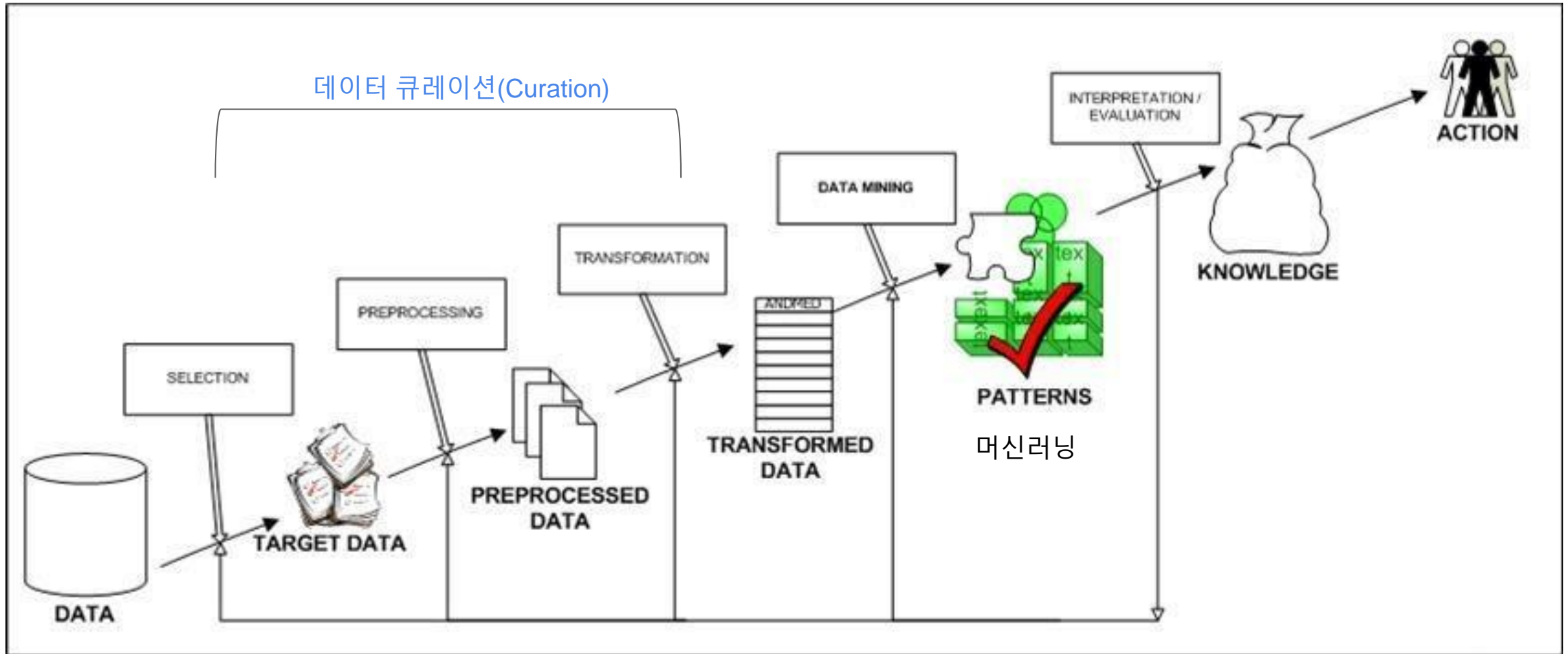
Key concepts: Data Mining

- 데이터 광산에서 의미있는 보석 캐기
- 데이터 분석을 인공지능이 돕는 것
- 적용 사례: 비즈니스 마케팅
(유튜브 알고리즘)

Key concepts: Big Data

- 정해진 시간내에 한 대의 머신으로는 처리가 어려운 양의 데이터
- 고성능 머신 또는 다수의 머신을 사용해 빅데이터를 처리할 수 있게 됨.
- ex) 실시간 대용량 Database 처리 Backend (DB)

Key concepts: 용어 (소통 단어) 익숙하기

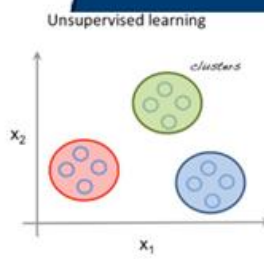
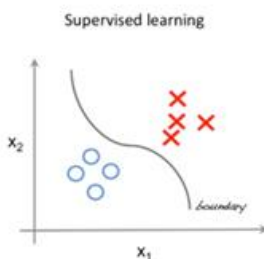


Source - <http://dataminingwarehousing.blogspot.kr/2008/10/data-mining-steps-of-data-mining.html>

머신러닝

□ 경험을 누적하며
학습하는 알고리즘

□ 정답 지도 여부



머신러닝의 구분 - 지도학습(Supervised Learning)과 비지도학습 (비지도 학습 : 비지도학습)

지도학습 vs 비지도학습

Supervised vs Unsupervised

+

입력과 결과 뿐 아니라

과정 '강화' 학습

Reinforcement



인텔의 정의

ARTIFICIAL INTELLIGENCE

A program that can sense, reason,
act, and adapt

MACHINE LEARNING

Algorithms whose performance improve
as they are exposed to more data over time

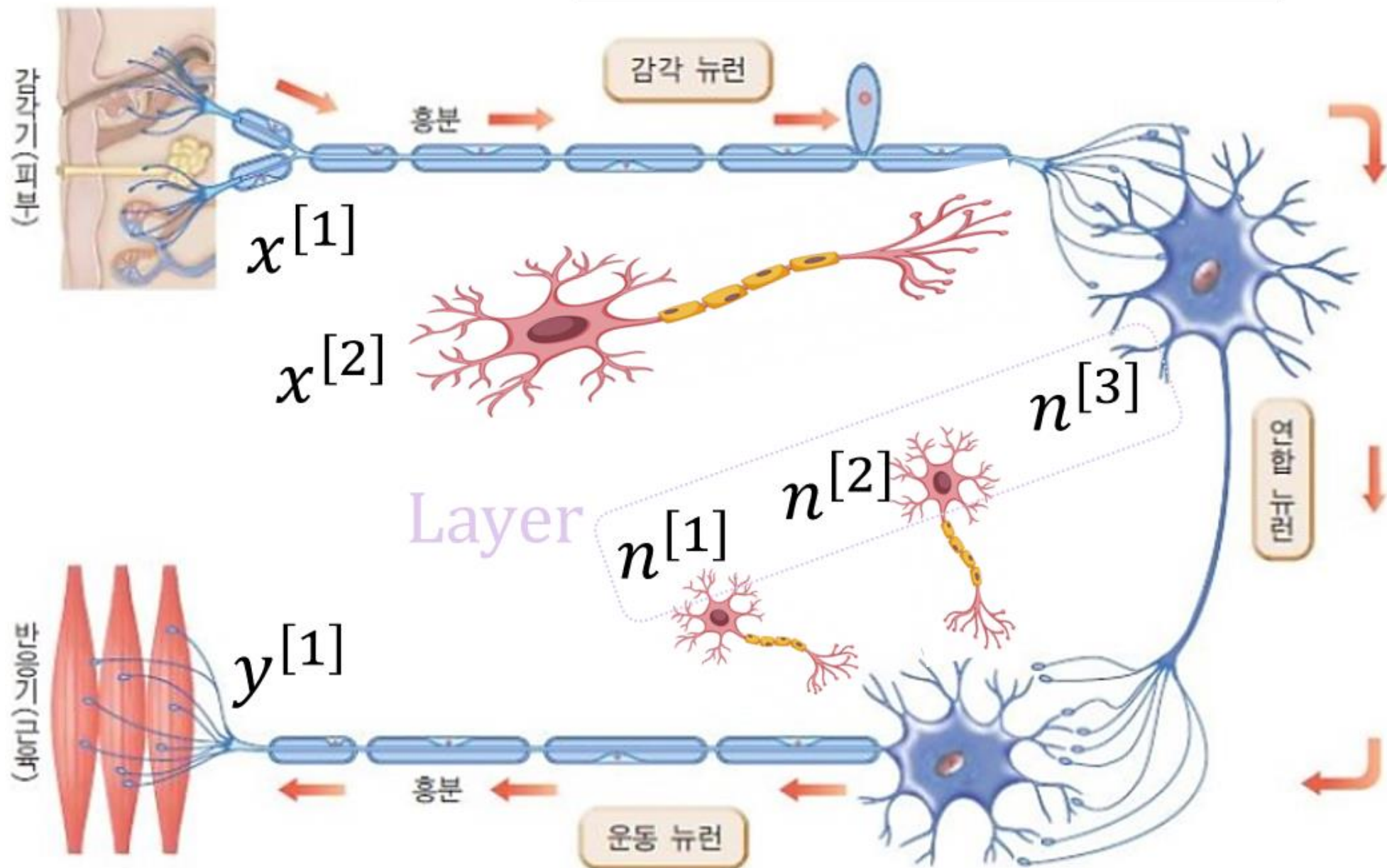
DEEP

Here are the five AI job titles with the highest salaries,
according to Indeed:

1. Machine learning engineer. Average **salary**: \$142,859.
2. Data scientist. Average **salary**: \$126,927.
3. Computer vision engineer. Average **salary**: \$126,400.
4. Data warehouse architect. Average **salary**: \$126,008.
5. Algorithm engineer. Average **salary**: \$109,313.



Neural Network



딥 러닝

Deep Learning

Key concepts: DeepLearning

- 얕았던(Shallow) Neural Net의 층을 깊게 !
 - ☐ 과거 계산량이 많아서 어려운 일이었으나,
하드웨어와 알고리즘의 개선으로
높은 성능 달성, 각광
- 사진내 객체 인식, 번역 등 탁월한 성능

ImageNet Challenge

IMAGENET

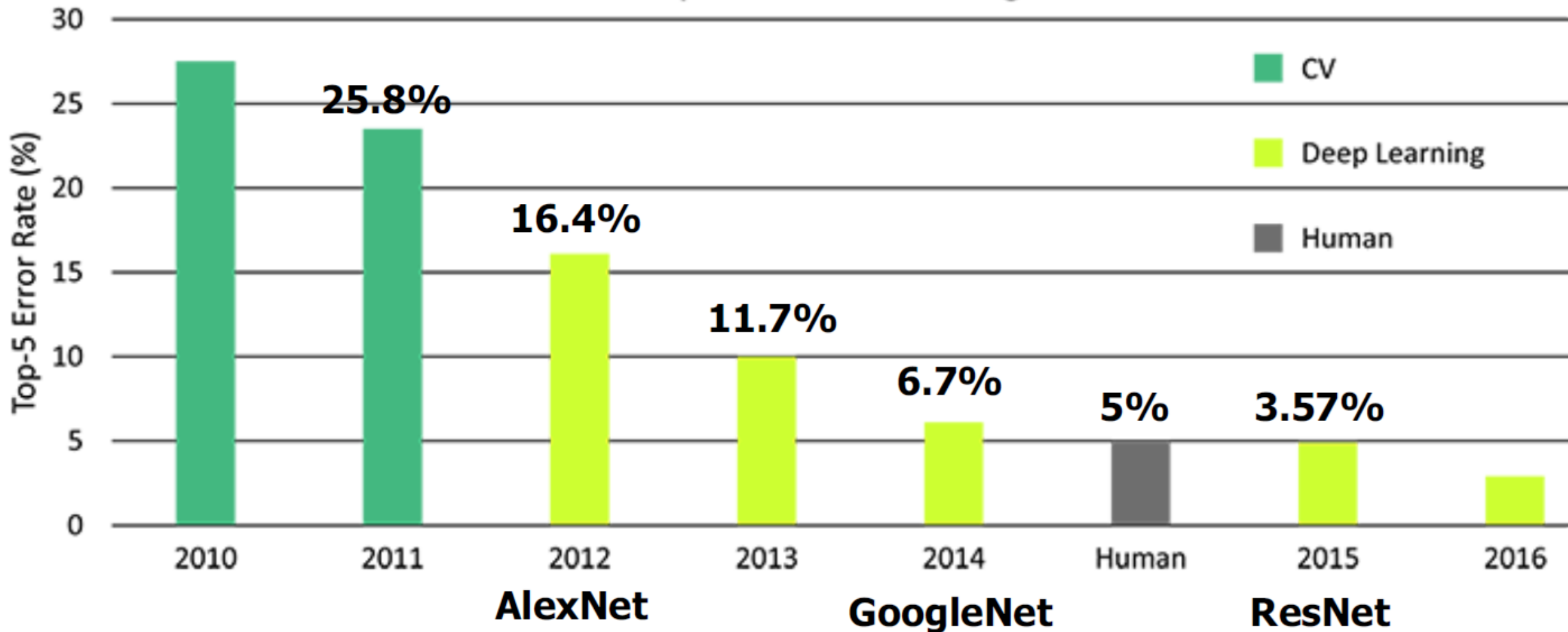
- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.



인간보다 낮아진 이미지 오인식률

(이미지 인식 경진)

ILSVRC Top 5 Error on ImageNet



Source - <https://goo.gl/UrBKWa>

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



1950's

1960's

1970's

1980's

1990's

2000's

2010's

Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

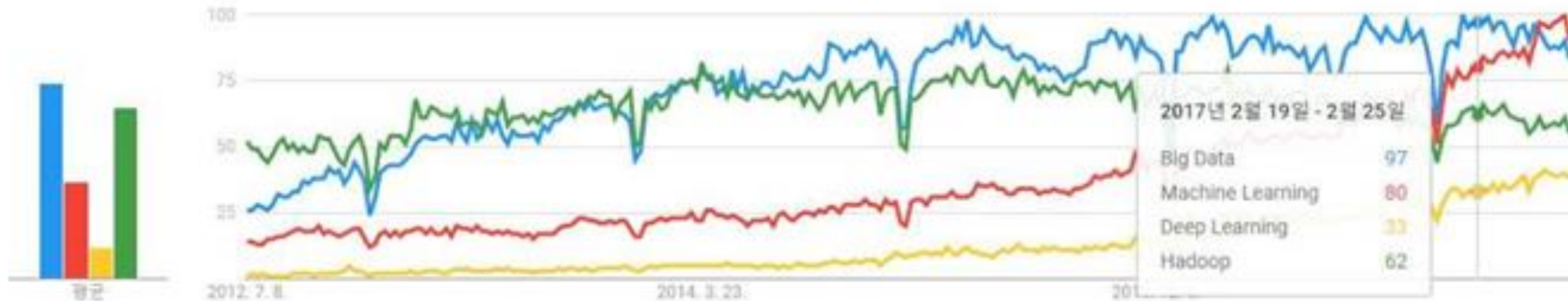
Source: <https://goo.gl/gz7EZT>

Big Data 검색어 Machine Learning 검색어 Deep Learning 검색어 Hadoop 검색어 +

전 세계 지난 5년 모든 카테고리 웹 검색

Google Trends

시간 흐름에 따른 관심도 변화 ?



빅데이터 ➡ 머신러닝 ➡ 딥러닝

딥러닝은 만능이 아닙니다

**머신러닝 부터
차근 차근**

Data Science

빅 데이터 처리

분산 처리
(병렬처리)

Database

Hadoop / Spark / NoSQL

MapReduce

HDFS

분산 파일저장

대용량 데이터 처리
Data Engineering

대형 데이터 처리

인공지능

Data Mining

Computer
Vision

NLP

자연어 처리

가능하게 함

Machine
Learning

Pattern
Recognition

Neural Net

SVM

Decision
Tree

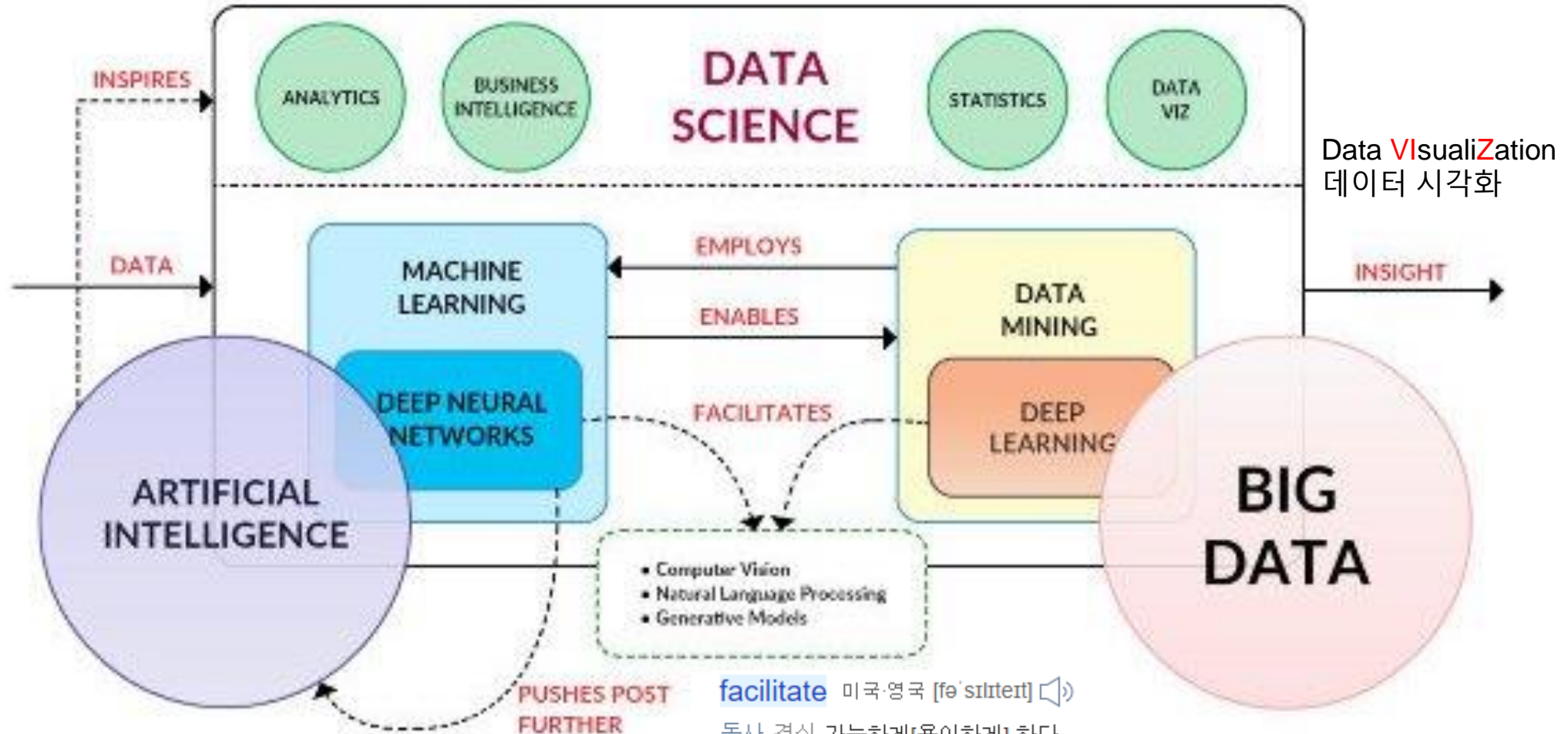
Statics
Probability

Linear
Algebra

Information
Theory

데이터 분석
Data Analysis

DATA SCIENCE 정의는 조금씩 다릅니다



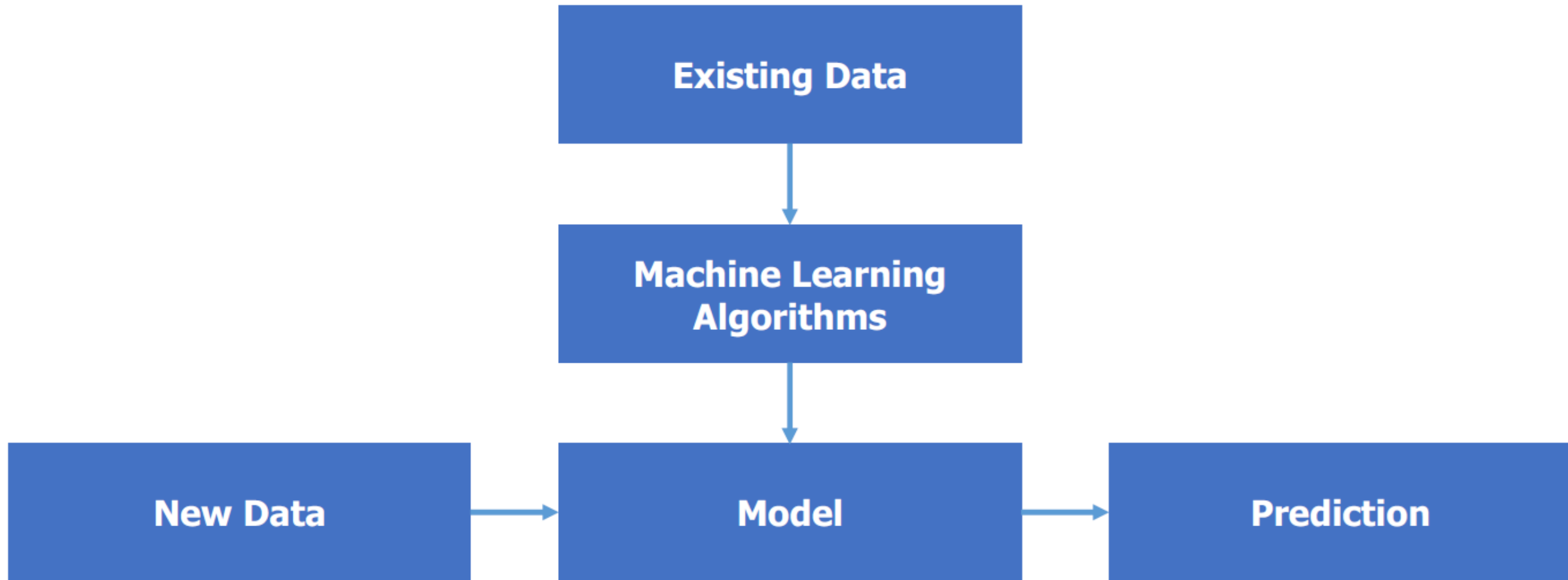
Data **Vi**suali**Z**ation
데이터 시각화

facilitate 미국·영국 [fə'sɪlɪteɪt]

동사 격식 가능하게[용이하게] 하다

옥스퍼드 영한사전

Machine Learning Process



Key concepts

Model - 예측을 위한 수학 공식, 함수

1차 방정식, 확률분포, condition rule

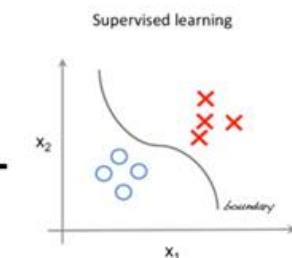
Algorithms - 어떠한 문제를 풀기 위한 과정

Model을 생성하기 위한 (훈련) 과정

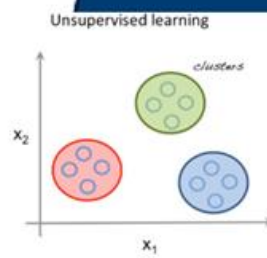
머신러닝

□ 경험을 누적하며
학습하는 알고리즘

□ 정답 지도 여부



출처: 머신러닝의 기본 - 지도학습(Supervised Learning)과 비지도학습(Unsupervised Learning) (가인 홍지, 정보리치스쿨)



지도학습 vs 비지도학습

Supervised vs Unsupervised

+

입력과 결과 뿐 아니라

과정 '강화' 학습

Reinforcement



인텔의 정의

ARTIFICIAL INTELLIGENCE

A program that can sense, reason,
act, and adapt

MACHINE LEARNING

Algorithms whose performance improve
as they are exposed to more data over time

DEEP

Here are the five AI job titles with the highest salaries,
according to Indeed:

1. Machine learning engineer. Average **salary**: \$142,859.
2. Data scientist. Average **salary**: \$126,927.
3. Computer vision engineer. Average **salary**: \$126,400.
4. Data warehouse architect. Average **salary**: \$126,008.
5. Algorithm engineer. Average **salary**: \$109,313.



감지, 추리, 행동, 적응하기 위한

Classification vs Regression

집단 분류 구조물 vs 수치 추정 구조물

1반 vs 2반

회 돌아올 회 歸 돌아갈 귀

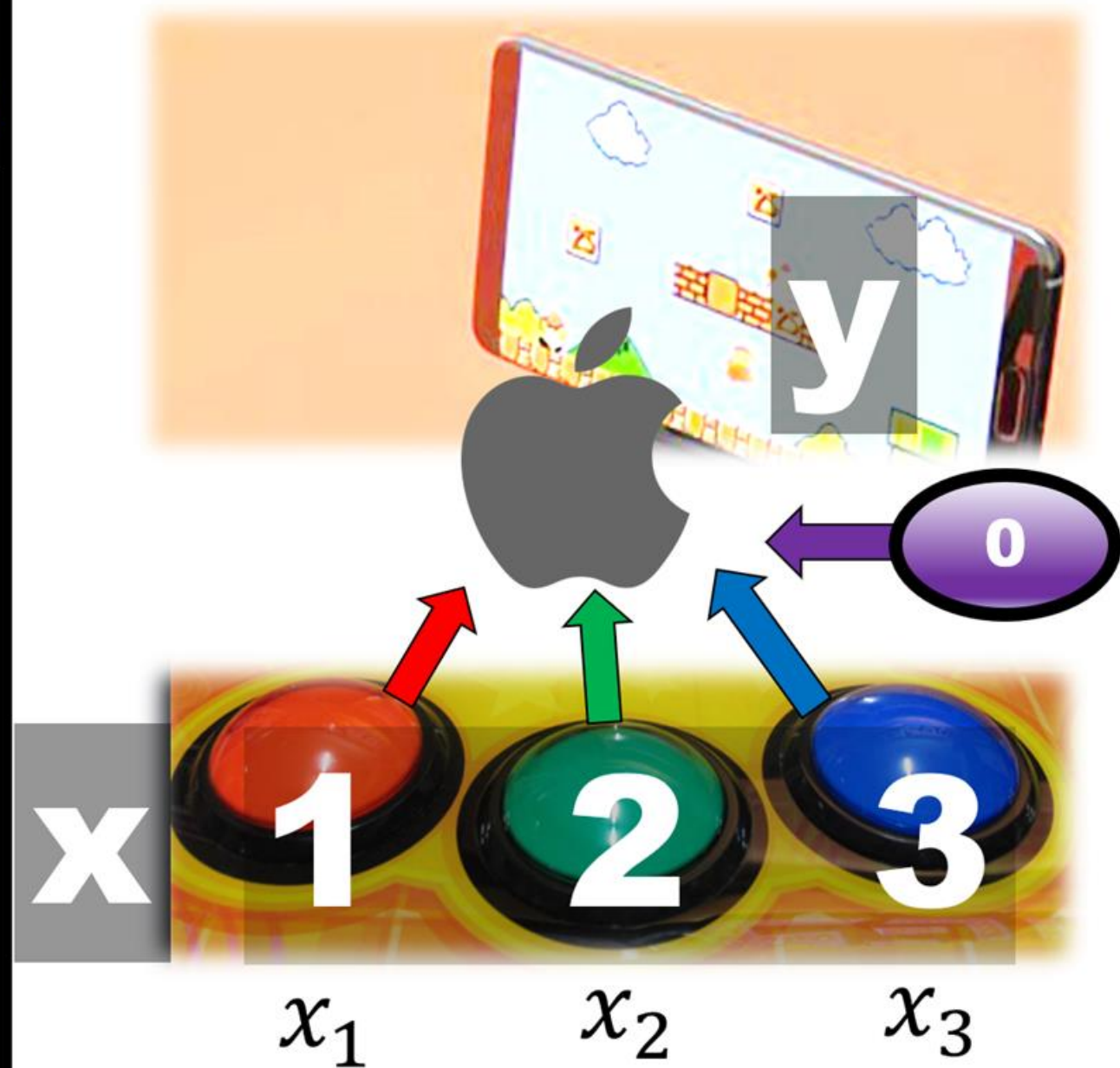


Regression

Regression 왜 회귀라 불리는지,



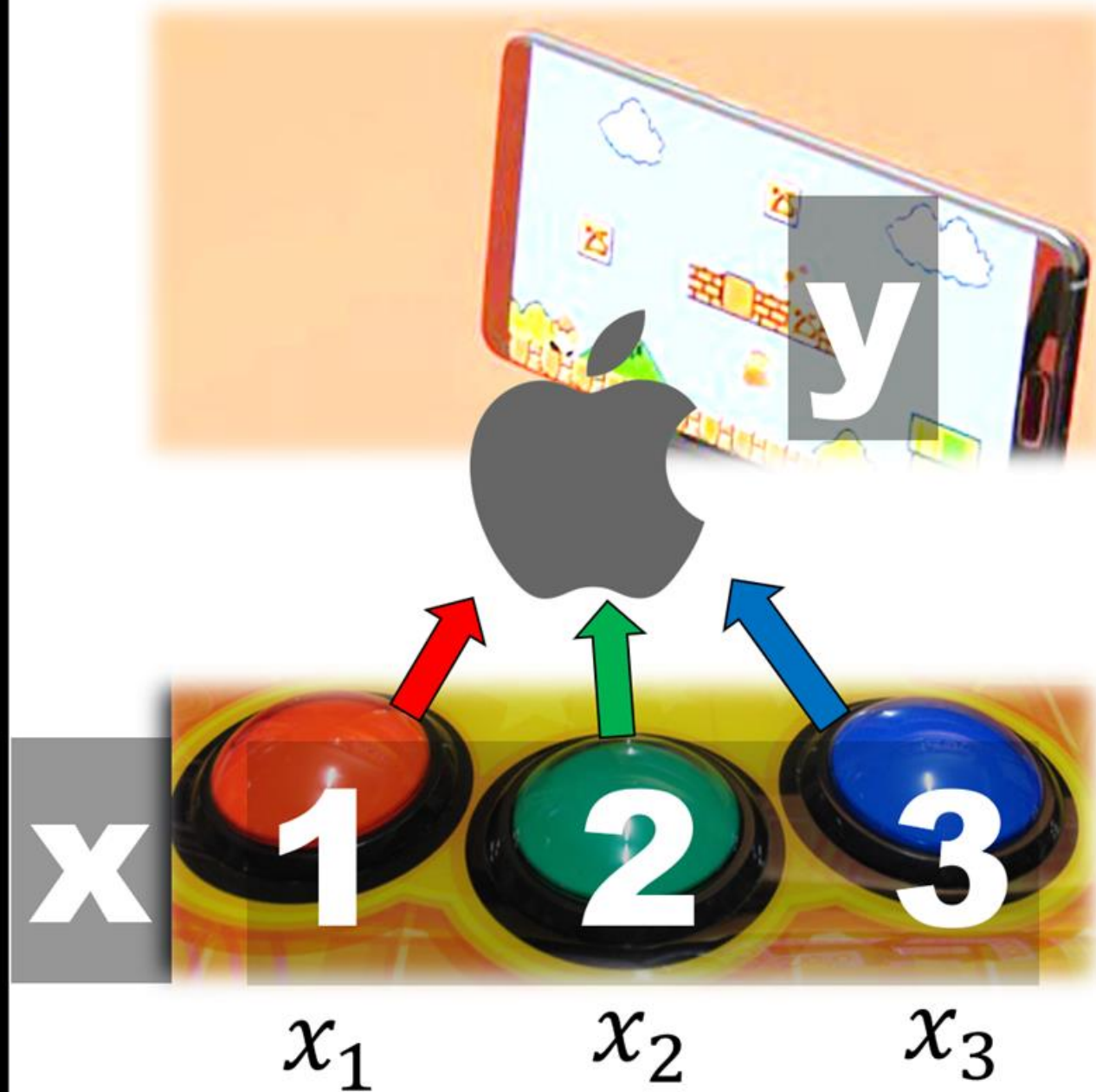
목표 : 오락기 구조를 아는 것
이유 : 새버튼 '0'의 결과 예측!



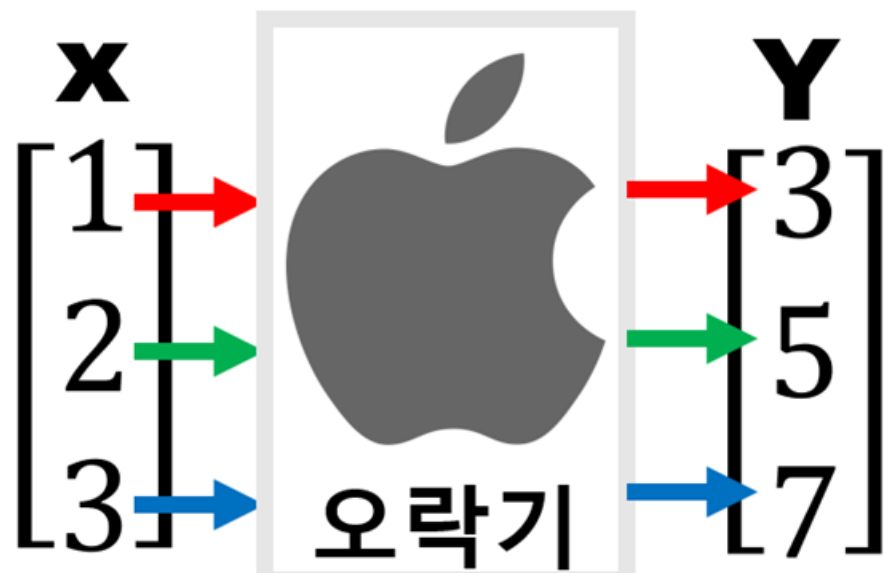
목표 : 오락기 구조를 아는 것
이유 : 새버튼 '0'의 결과 예측!



실제, $y_i = @x_i + \$$



목표 : 오락기 구조를 아는 것
이유 : 새버튼 '0'의 결과 예측!



실제, $y_i = @x_i + \$$

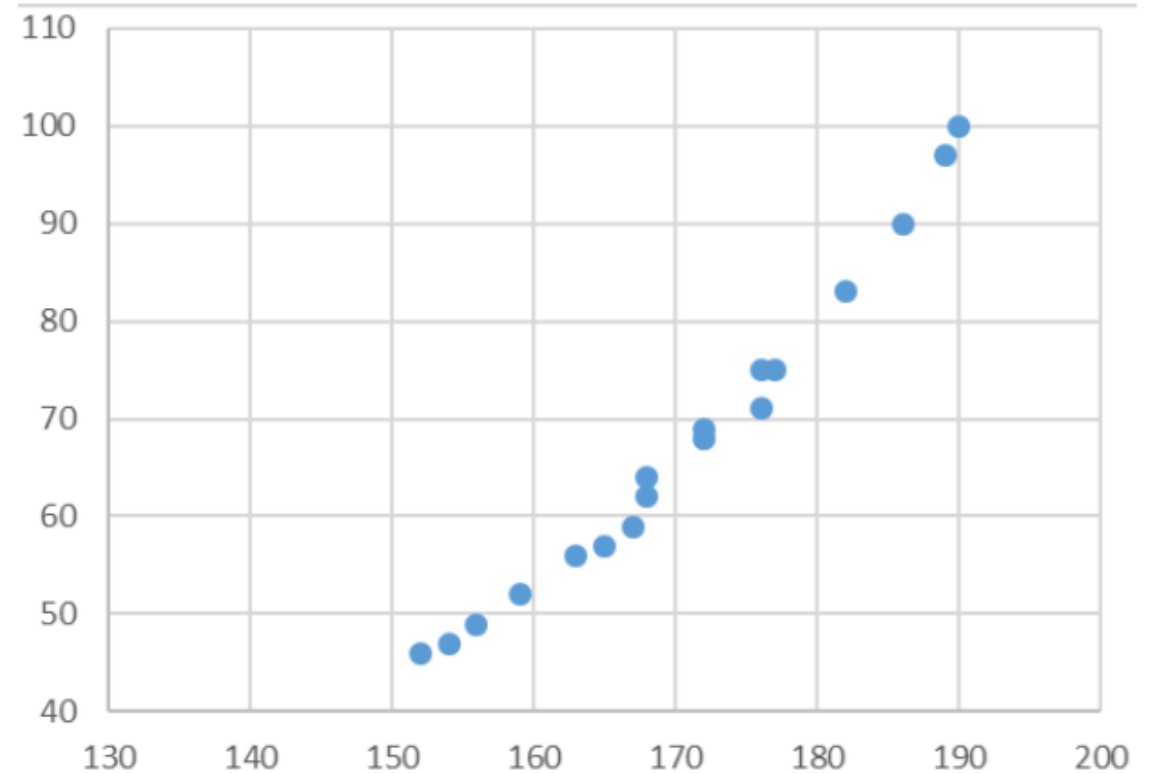
가설, $H(w, b) = Wx_i + b$

$i = 1, 2, 3$ 까지는 경험
 $i = 4$ 는 새로운 입력

가설, $Hypothesis(W, b)$
 $= Weight \cdot x_i + bias$

키를 가지고 몸무게를 예측할 수 있을까요?

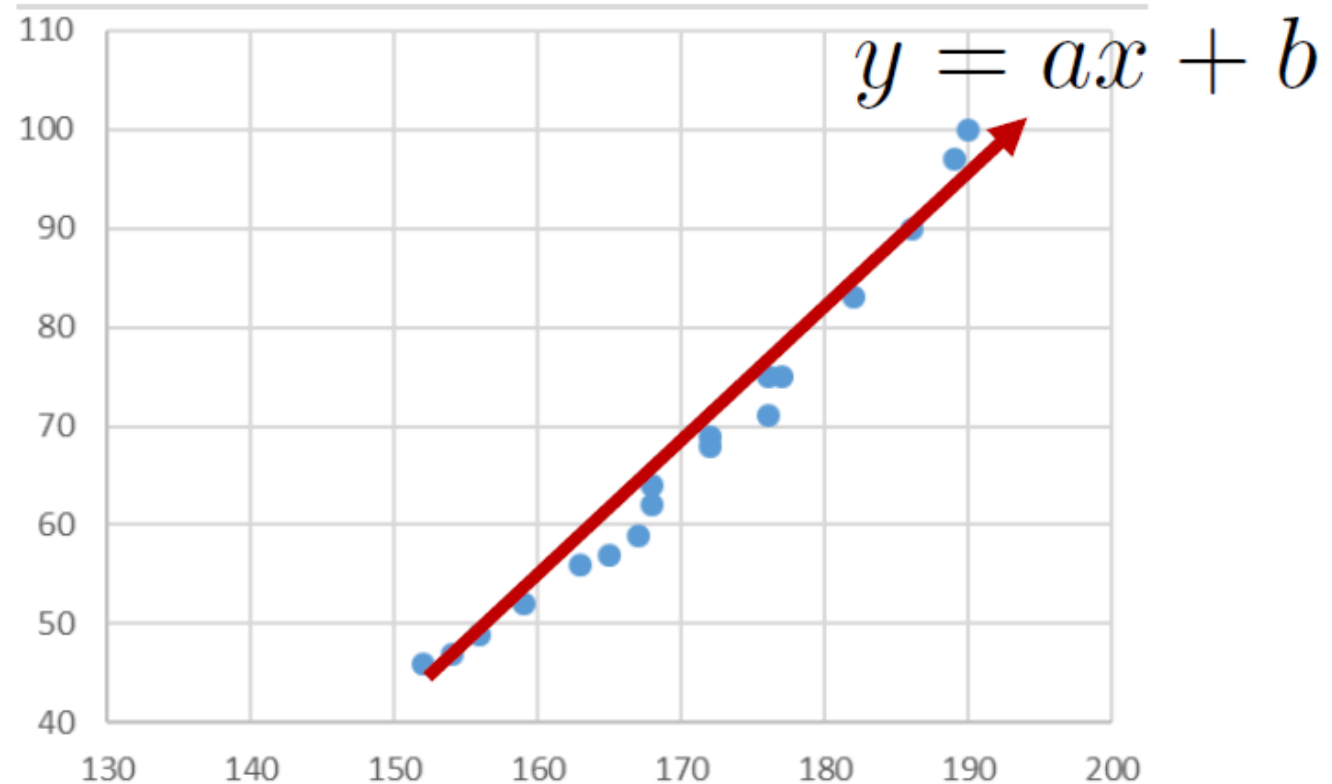
Height	Weight	Height	Weight
152	46	172	69
154	47	172	68
156	49	176	71
159	52	176	75
163	56	177	75
165	57	182	83
167	59	186	90
168	64	189	97
168	62	190	100



Data source: <http://goo.gl/gDscUQ>

키를 가지고 몸무게를 예측할 수 있을까요?

Height	Weight	Height	Weight
152	46	172	69
154	47	172	68
156	49	176	71
159	52	176	75
163	56	177	75
165	57	182	83
167	59	186	90
168	64	189	97
168	62	190	100



키를 안다면 추정된 몸무게 값은 이 **모델**로 돌아오는 경향이 있
다!라는 뜻의 **회귀**

Data source: <http://goo.gl/gDscUQ>

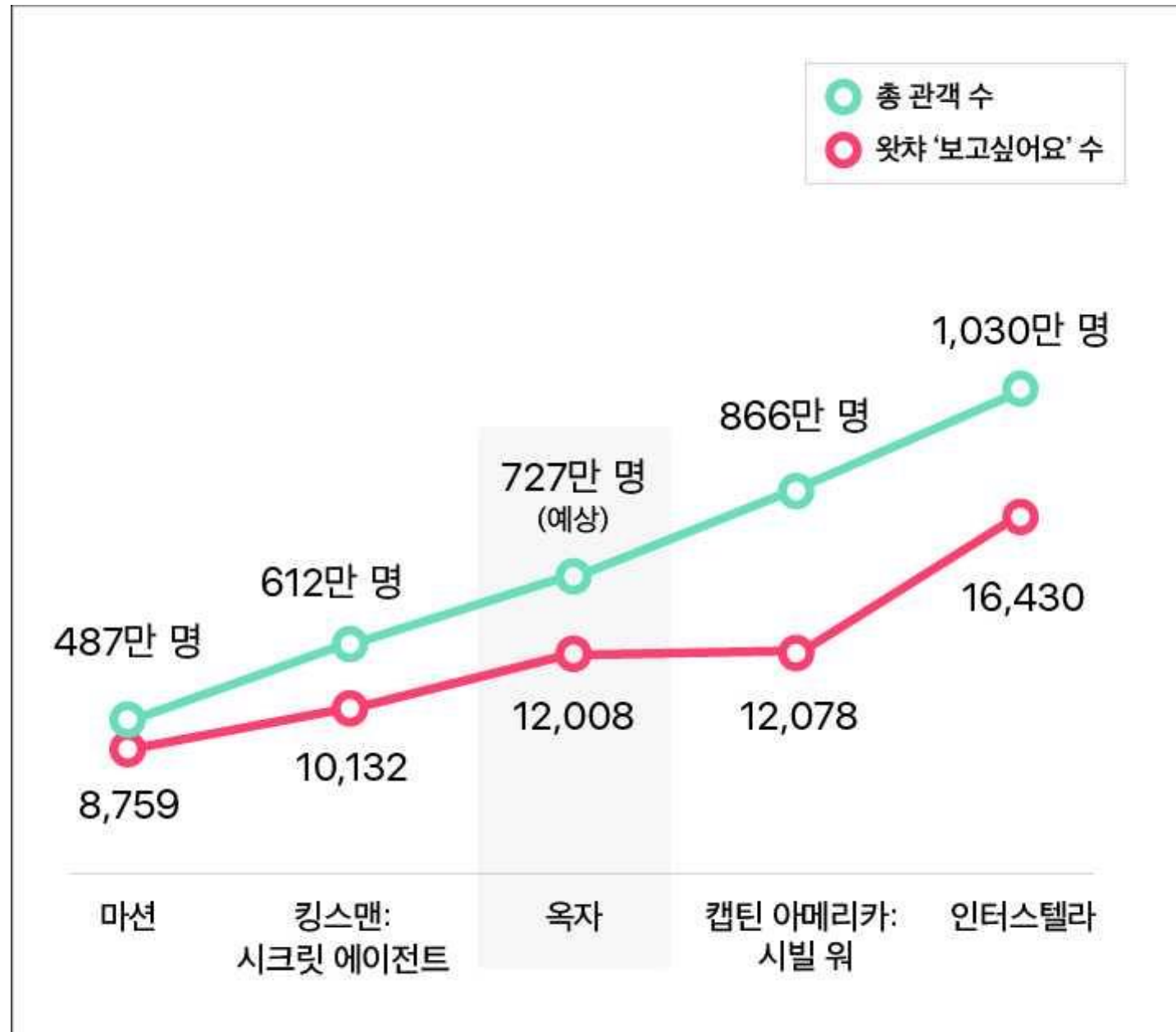
회귀의 사례

추정하고 싶은 수치 (Estimate)

“영화관에서 개봉은 안한 옥자가
만약 극장에서 개봉했다면 예상 관객수는?”

추정에 사용된 정보의 특징 (Feature)

= 왓차라는 영상앱에서 ‘보고싶어요’ 숫자

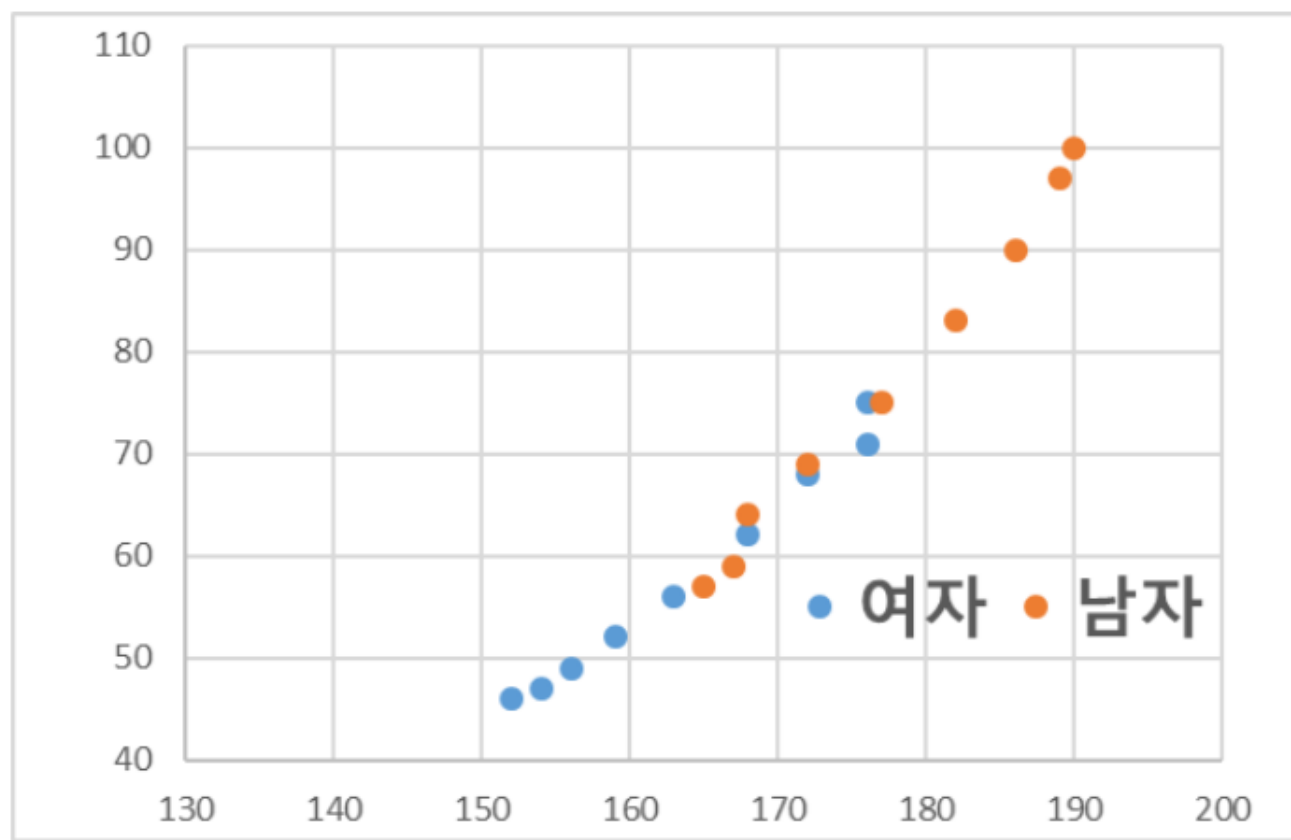


Classification

키와 몸무게를 가지고 남/녀를 분류할 수 있을까요?

새 데이터: H 186, W 82 남자? 여자?

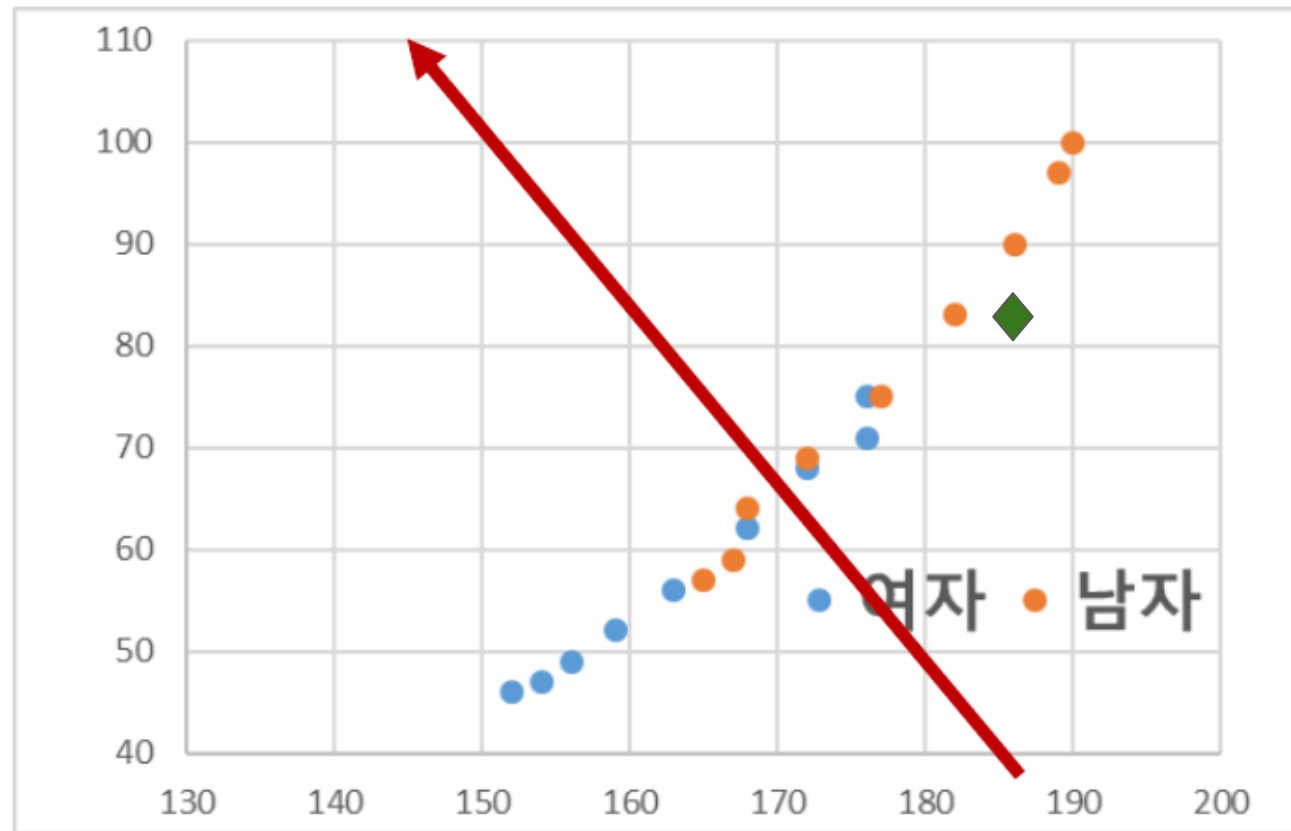
Sex	Height	Weight	Sex	Height	Weight
여자	152	46	남자	172	69
여자	154	47	여자	172	68
여자	156	49	여자	176	71
여자	159	52	여자	176	75
여자	163	56	남자	177	75
남자	165	57	남자	182	83
남자	167	59	남자	186	90
남자	168	64	남자	189	97
여자	168	62	남자	190	100



Classification – 분류, 데이터의 유형

기존 데이터를 바탕으로 데이터 유형을 나눠 보기

Sex	Height	Weight	Sex	Height	Weight
여자	152	46	남자	172	69
여자	154	47	여자	172	68
여자	156	49	여자	176	71
여자	159	52	여자	176	75
여자	163	56	남자	177	75
남자	165	57	남자	182	83
남자	167	59	남자	186	90
남자	168	64	남자	189	97
여자	168	62	남자	190	100

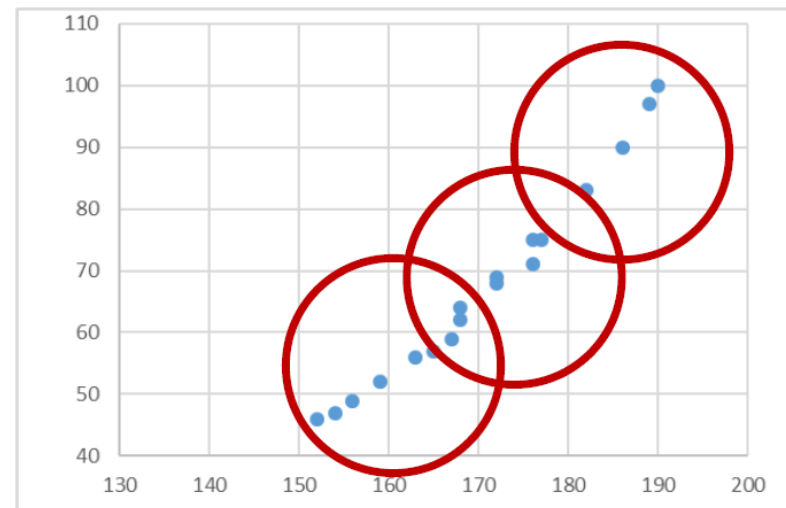


<https://goo.gl/z5rEu6>

추가 정보

Clustering 뭉치기 - 데이터를 그룹으로 뭉치는 행위

해당 데이터를 특정 기준에 따라
더 가까워서 속하게 된 '그룹으로 뭉치는 행위'
머신러닝이 아닌 단순 규칙, 뭉치기도 존재.



‘머신러닝으로 뭉치는 과정을 다 거쳐서 최종 결과인 경우’
= ‘ML Classification (분류)’로 이름 붙기도 함.

또다른 정의들로도 용어가 혼용

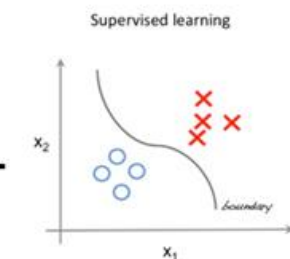
(문맥으로 파악, 예: ‘비지도학습은 Clustering을 수행’)

정리

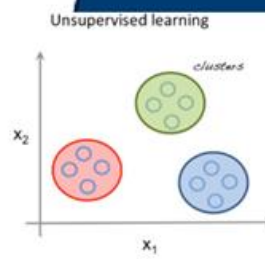
머신러닝

□ 경험을 누적하며
학습하는 알고리즘

□ 정답 지도 여부



머신러닝의 구분 - 지도학습(Supervised Learning)과 비지도학습 (비지도 학습 : 비지도학습)



지도학습 vs 비지도학습

Supervised vs Unsupervised

+

입력과 결과 뿐 아니라

과정 '강화' 학습

Reinforcement



인텔의 정의

ARTIFICIAL INTELLIGENCE

A program that can sense, reason,
act, and adapt

MACHINE LEARNING

Algorithms whose performance improve
as they are exposed to more data over time

DEEP

Here are the five AI job titles with the highest salaries,
according to Indeed:

1. Machine learning engineer. Average **salary**: \$142,859.
2. Data scientist. Average **salary**: \$126,927.
3. Computer vision engineer. Average **salary**: \$126,400.
4. Data warehouse architect. Average **salary**: \$126,008.
5. Algorithm engineer. Average **salary**: \$109,313.



다음 시간 Next Lecture

쿠키를 상자 속에서 안보고 집었을 때

바닐라 맛일까 초콜렛 맛일까

그 확률 모델을 만들어봅시다.

Naive Bayes 미리 살펴보기를 추천합니다.