



orange 활용 데이터 분석 및 머신 러닝

목차

1. 데이터분석과 오렌지
2. 지도학습 - 회귀 1.
3. 지도학습 - 회귀 2.
4. 지도학습- 분류 1.
5. 지도학습- 분류 2.
6. 비지도학습 - 군집화1.
7. 비지도학습 - 군집화2.
8. 이미지데이터를 활용한 분류와 군집화 1.
9. 이미지데이터를 활용한 분류와 군집화 2.
10. 텍스트 분석 1.
11. 텍스트 분석 2.
12. 연관분석
13. 시계열 분석
14. 데이터 전처리
15. 데이터 분석 실습



1차시

데이터분석과 orange !

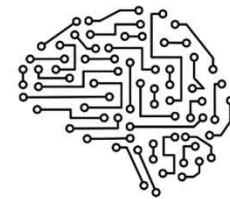




1. 인공지능과 머신 러닝의 개념
2. 데이터 분석 과정의 이해
3. 오렌지 소개 및 설치

‘인공지능’ 하면 가장 먼저 떠오르는 것은?

인터넷
가상인간 머신러닝
자율주행
로봇
데이터
AI
가상비서
기계학습
스마트스피커
구글 메타버스
인공지능
가상인간
머신러닝
자율주행
로봇
데이터
AI
가상비서
기계학습
스마트스피커
구글 메타버스



“ 인공지능 ? ”

Artificial intelligence



사람의 지능 을 컴퓨터로 모방한 시스템



Artificial intelligence

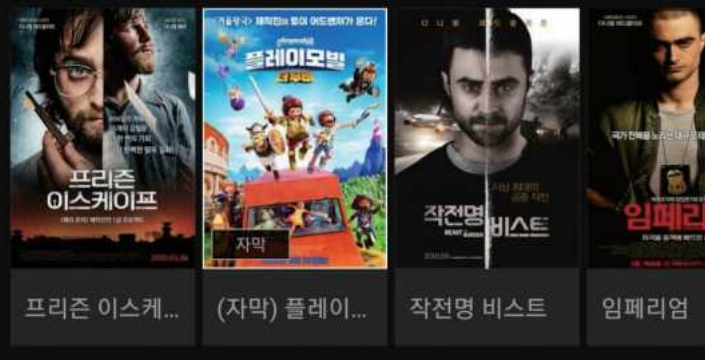
- 인공적으로 만들어진 지능을 의미
- 컴퓨터가 인간의 지능을 모방하는 것
- 문제를 해결하고 학습하고 추론하고 인식을 함.
- 빅데이터 기반의 학습으로 지능을 높일 수 있음.

인간의 지적 능력을 컴퓨터로 구현하는 과학기술로서, 상황을 인지하고, 이성적, 논리적으로 판단 행동하며 감성적, 창의적 기능을 수행하는 능력까지 포함한다.



추천 시스템

다니엘 래드클리프 팬이라면 무조건 PLAY!



엠마 왓슨 팬이라면 무조건 PLAY!



비슷한 영화



wavvie 영화 업데이트

더보기 >



OTT 서비스 플랫폼, 웨이브의 영화 추천 화면

자율주행차



(출처: 게티이미지뱅크)



자율주행 트럭 마스오토

스마트팜, 조미료 개발, 소물리에 등



AI스마트팜 (AI타임즈 2021.04.21.기사)

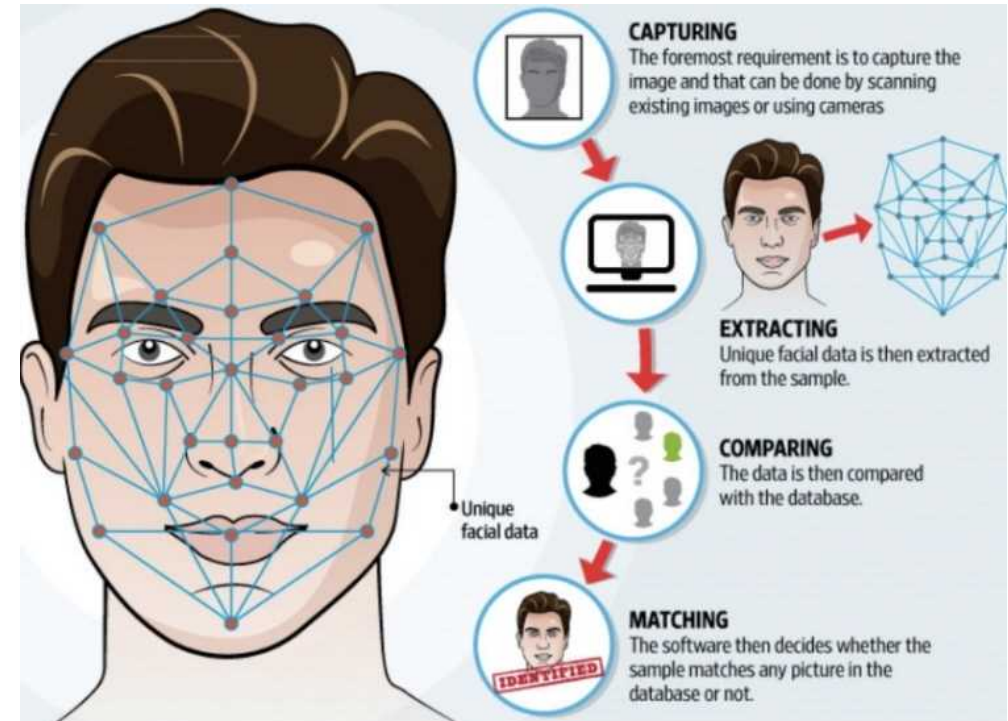


IBM은 전자 혀와 인공지능을 이용해 새로운 조미료를 개발하고 있다. (조선일보 2021.07.06.기사)

Face ID / 얼굴 인식



<https://www.pcmag.com/how-to/set-up-use-face-id-iphone>



<https://www.boannews.com/media/view.asp?id=x=85438>

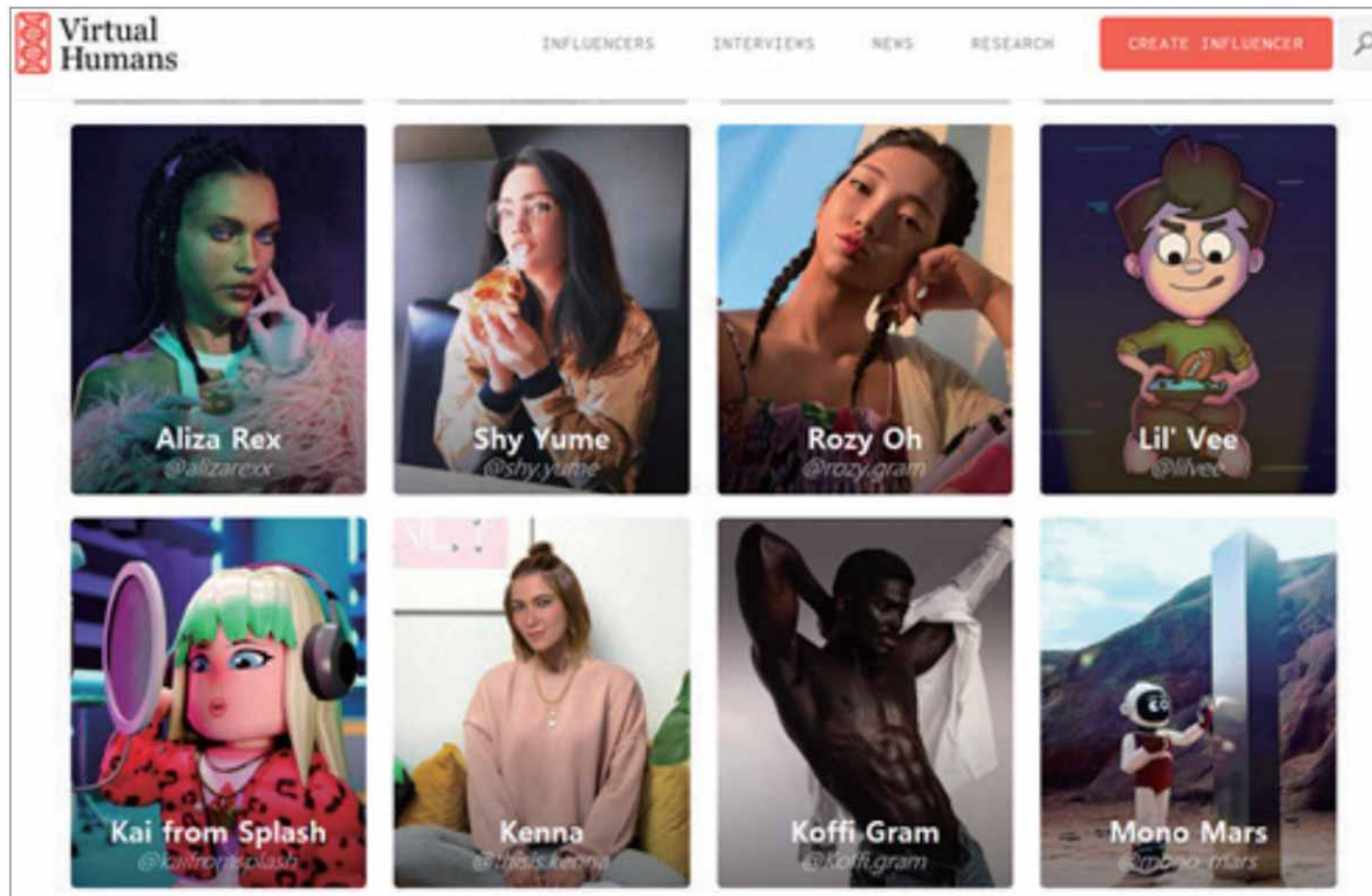
말벗 인공지능

차니봇 안부를 묻는 인사말 “괜찮니?”



노인층을 위한 AI 인형 (다카라토미)

인공지능 엔터테이너



버추얼 휴먼스에 등록된 가상 인간 (virtualhumans.org)

요약

뉴스의 핵심만 요약해 드립니다.

Input news :

사람에게 편리한 인터넷 검색 서비스나 스마트 스피커 기기는 그 안에 뛰어난 인공지능(AI)의 힘을 품고 있다. AI가 사람의 언어를 잘 이해하고 사용자에게 원하는 결과를 제공하려면 뛰어난 자연어처리(NLP) 능력을 갖춰야 한다. NLP는 문서요약, 질의응답, 챗봇 대화, 번역, 문장·글 생성, 자동완성, 기계독해, 생식 수준 같은 분야의 문제를 풀기 위한 기술이다.

원문 뉴스

사람에게 편리한 인터넷 검색 서비스나 스마트 스피커 기기는 그 안에 뛰어난 인공지능(AI)의 힘을 품고 있다. AI가 사람의 언어를 잘 이해하고 사용자에게 원하는 결과를 제공하려면 뛰어난 자연어처리(NLP) 능력을 갖춰야 한다. NLP는 문서요약, 질의응답, 챗봇 대화, 번역, 문장·글 생성, 자동완성, 기계독해, 생식 수준 같은 분야의 문제를 풀기 위한 기술이다.

이런 기술을 확보하기 위해 SK텔레콤 'AI연구지체크랩스(ALT Labs)' 연구원들은 한국어를 잘 다루는 AI '코베르트(KoBERT)'와 'KoGPT2'를 잇따라 개발했다. 사내에서 KoBERT는 상담 챗봇, 법무·특허등록 지원용 AI 검색, 내부 마케팅 자료에서 질문에 알맞은 답을 추출하는 '기계독해' 기술에 쓰였다. KoGPT2는 챗봇의 고객응대 말투에 자연스러움을 더했다.

이 SK텔레콤의 AI 기술들은 일부 NLP 영역에서 기존 한국어 AI 기술보다 뛰어난 성능을 보여 준다. SK텔레콤 ALT Labs 측은 "네이버 영화 리뷰 데이터(NSMC)를 활용한 '감정분석 분류' 작업에서 KoBERT는 90.1%, KoGPT2는 89.9%의 성능(정확도)을 보인다"며 "KoBERT 사용 이전에 이 작업은 일반적으로 83~85%의 정확도를 보였다고 알려져 있었다"고 설명했다.

KoBERT의 원형은 구글이 지난 2018년 10월 공개한 버트(BERT)이며, KoGPT2의 원형은 오픈AI가 지난해 공개한 GPT-2다. BERT는 영어 읽기, GPT-2는 영어 쓰기에 특화된 '언어 신동' AI로 유명세를 얻었다. SK텔레콤은 이를 통해 한국어를 열심히 가르친 결과, 내부 업무에 활용해 성과를 얻었고, 지난해 10월 KoBERT를, 올해 2월엔 KoGPT2를 각각 오픈소스로 공개했다. SK텔레콤 ALT Labs 천은가 팀은 두 AI를 어떻게 만들었을까. 우선 방대한 한국어 데이터를 집하게 했다. KoBERT에는 위키 문서의 문장 500만개(5400만단어)와 뉴스 문장 2000만개(2억7000만단어)가 사용됐다. KoGPT2에는 위키 문장 500만개, 뉴스 문장 1억2000만개(16억단어), 기타 자료 문장 940만개(8800만단어) 및

문장 1800만개(8200만단어)가 동원됐다.

처리할 데이터가 방대한 만큼, 컴퓨터 자원이 많이 필요했다. 별도의 외부 클라우드서비스를 활용해야 했고, 이를 낭비 없이 쓸 수 있어야 했다. SK텔레콤은 클라우드서비스 회사 아마존웹서비스(AWS)의 도움을 받았다. 이에 더해 아마존 머신러닝솔루션즈팀에 대용량 학습 인프라 환경을 구축했고, AWS 클라우드NLP 팀이 학습로드 최적화를 지원했다.

SK텔레콤 ALT Labs 측은 "KoGPT2 학습 당시 대용량학습을 위한 자원이 충분치 않았는데, AWS에서 이에 필요한 적지 않은 자원을 제공해 줬다"고 밝혔다. KoBERT 학습에 2개월간 엔비디아의 '태슬라 V100' GPU 16개, KoGPT2 학습에 1주일간 V100 64개가 사용됐다. V100은 AI 연산에 특화된 GPU로, 개당 가격 1000만원이 넘는 부품이다.

외부 개발자와 기업들은 SK텔레콤이 오픈소스로 공개한 KoBERT와 KoGPT2를 활용해 특정한 목적의 AI 기술을 더 효율적으로 개발할 수 있게 됐다. SK텔레콤 측은 "우리가 공개한 것은 다양한 언어처리 작업에 널리 쓰이는 것을 전제로 사전학습된(pre-trained) 모델"이라며 "대화나 질문에 대한 답변 등에 추가 학습으로 특화시켜 활용도를 높일 수 있다"고 설명했다.

해위에선 BERT와 GPT-2 이후 더 성능이 뛰어난 AI 기술이 나오고 있다. 구글은 지난해 11월 열린 대학교와 공동연구로 'XLNet'을, 시카고대학교 도요타공업대학원(TTC)과 공동연구로 'ALBERT'를, 자체 연구로 'T5'를 줄줄이 내놨다. 오픈AI도 GPT-2 후속 버전인 'GPT-3'를 올해 출시했다.

더 뛰어난 AI와 해도 아직 여러 언어의 장벽을 스스로 뛰어넘지는 못한다. 여전히 한국어 성능을 높이기 위한 추가 작업을 필요로 한다는 얘기다. SK텔레콤은 "KoBERT와 KoGPT2는 문장 이해와 문장 생성 분야에서 가장 대표적인 모델"이라면서도 "정확한 모델명을 언급하긴 어렵지만 이외에도 활용도가 높은 몇가지 모델의 한국어 버전을 개발하고 있다"고 밝혔다.

요약된 뉴스

SK텔레콤의 AI연구지체크랩스 연구원들은 한국어를 잘 다루는 AI '코베르트(KoBERT)'와 'KoGPT2'를 개발하여 한국어를 열심히 가르친 결과 내부 업무에 활용해 성과를 얻었다.

<https://www.ajunews.com/view/20201210114639936>

양날의 검, 딥페이크 기술

- 딥페이크^{deepfake}: 인공지능을 이용해 실존하는 인물의 얼굴이나 신체의 특정 부위 또는 음성을 합성하는 기술



딥페이크 기술로 만든 오바마 전 대통령(왼쪽)과 조던 필 감독(오른쪽)
(출처: 유튜브 채널 BuzzFeedVideo)

- 가짜 뉴스^{fake news}: 사실이 아닌 것을 사실인 것처럼 꾸민 뉴스

긍정적인 사용의 예

인간을 향한 기술의 순기능 VR 휴먼 다큐 '너를 만났다' : <https://stonebc.com/archives/21824>

<https://www.bbc.com/korean/news-51498614>



MyHeritage
가계도
사진
조사
DNA
도움말

인터넷을 강타한 센세이션, Deep Nostalgia™에 오신 것을 환영합니다!
나날이 그 수를 더해가는 91백만 이상의 애니메이션!

Animate your family photos

Animate the faces in your family photos with amazing technology. Experience your family history like never before!

Upload photo

OR DRAG AND DROP >

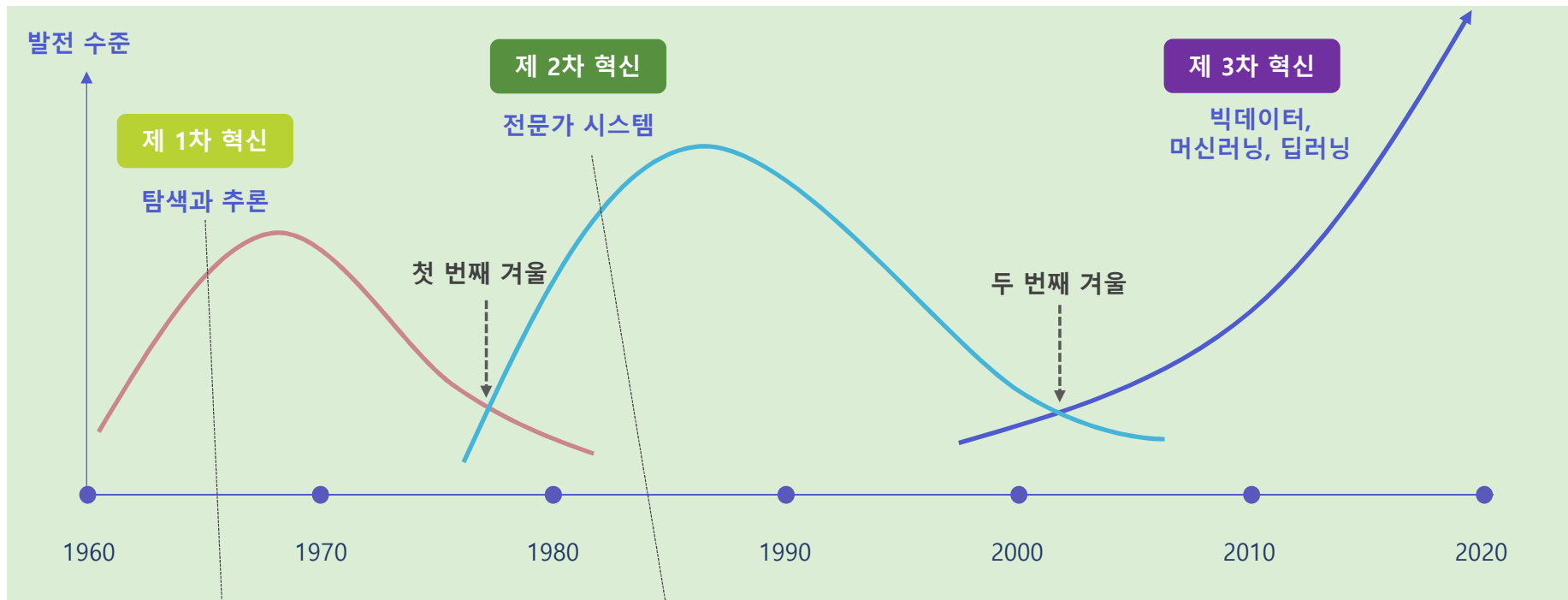
Free signup is required. Photos uploaded without completing signup are automatically deleted to protect your privacy.

Licensed by MyHeritage from D-ID



인공지능의 발전

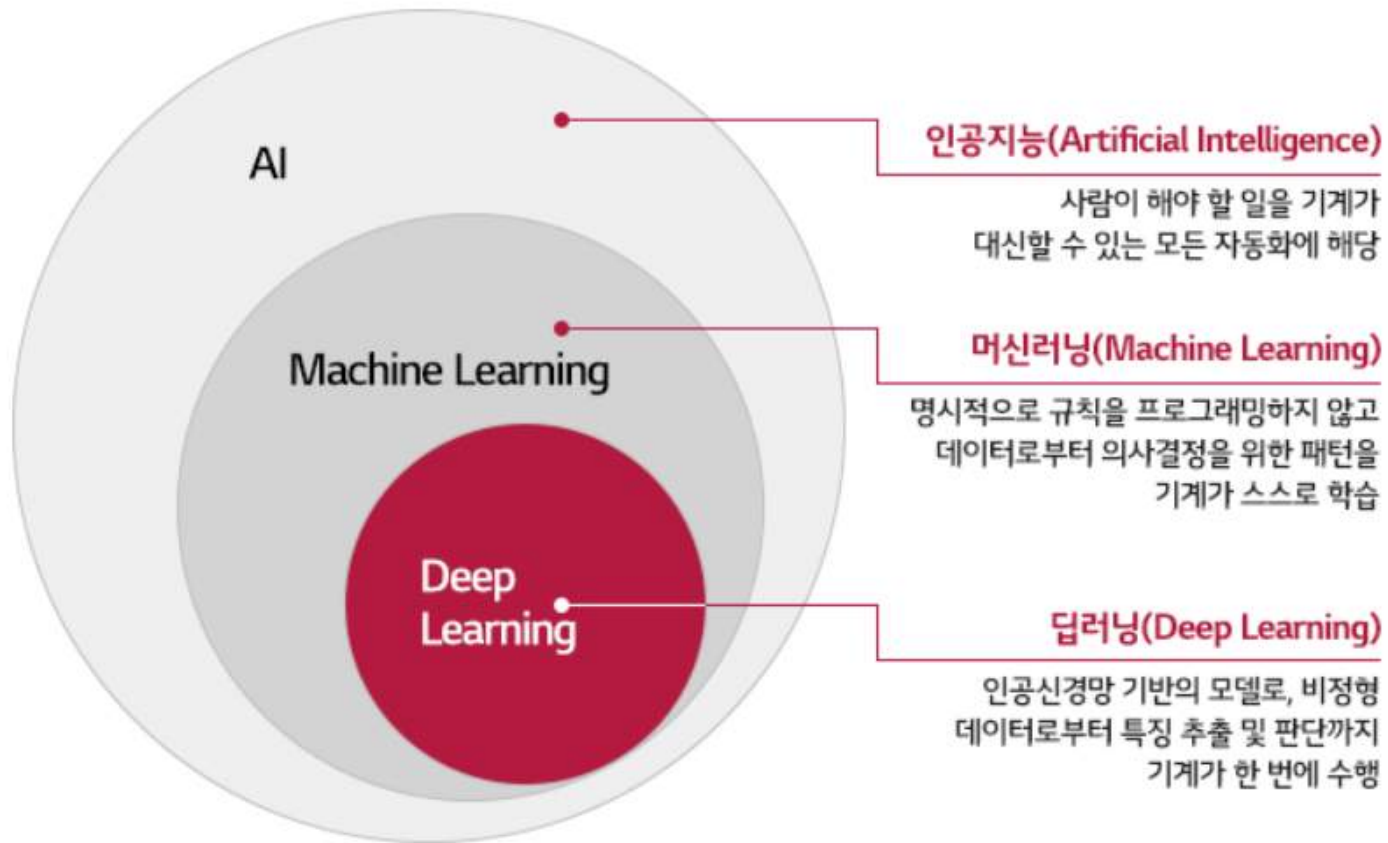
○ 규칙기반의 인공지능 → 학습기반의 인공지능



미로와 퍼즐을 스스로
풀 수 있는 **문제해결**
알고리즘 컴퓨터 연구

사람이 **알려준 지식**을 **활용**하여 문제
를 해결하는 **전문가** 컴퓨터 연구

인공지능의 발전

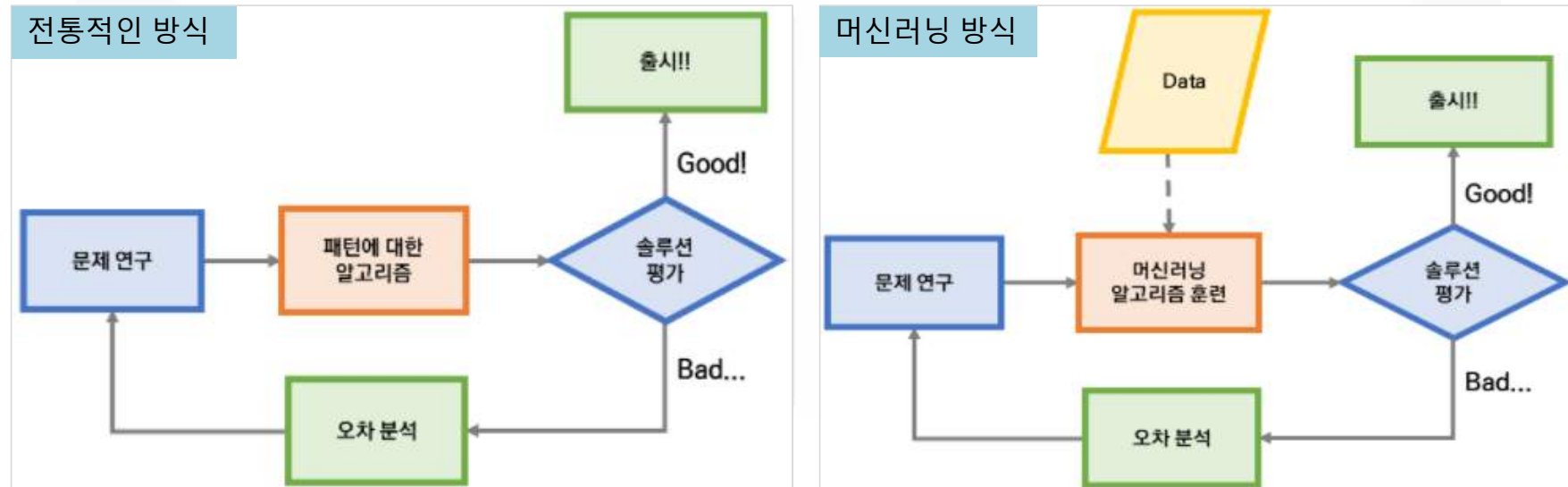


<출처 : LGCNS 블로그 <https://blog.lgcns.com/2212>>

머신러닝(Machine Learning)

1959년 Arthur Samul: 머신러닝의 정의

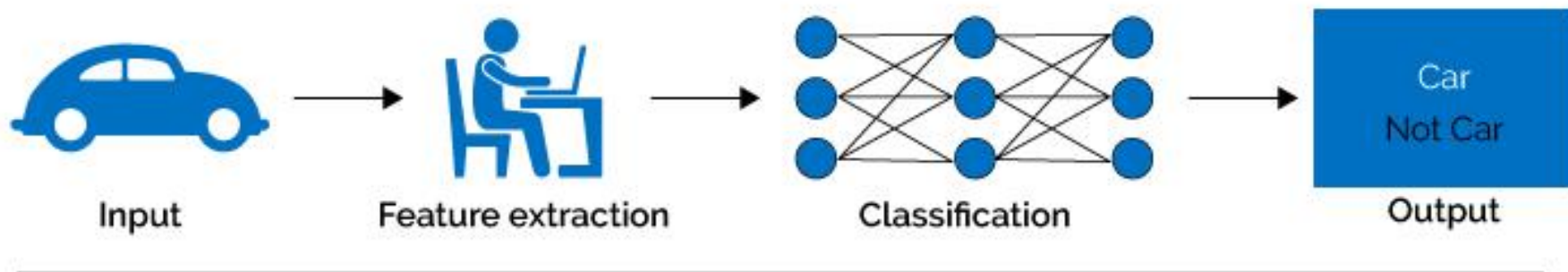
“gives computers the ability to learn without being explicitly programmed”
컴퓨터에게 명시적 프로그래밍 없이 학습하는 능력을 주는 것



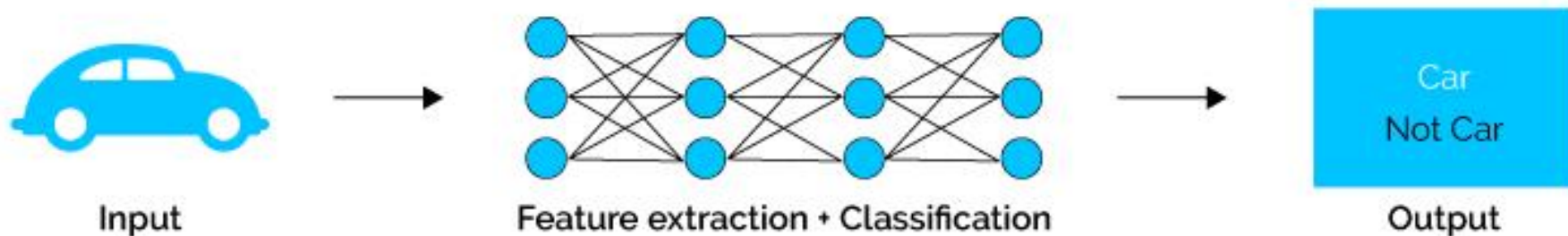
<출처><https://goopy.tistory.com/121>

머신러닝 vs. 딥 러닝(Deep Learning)

Machine Learning

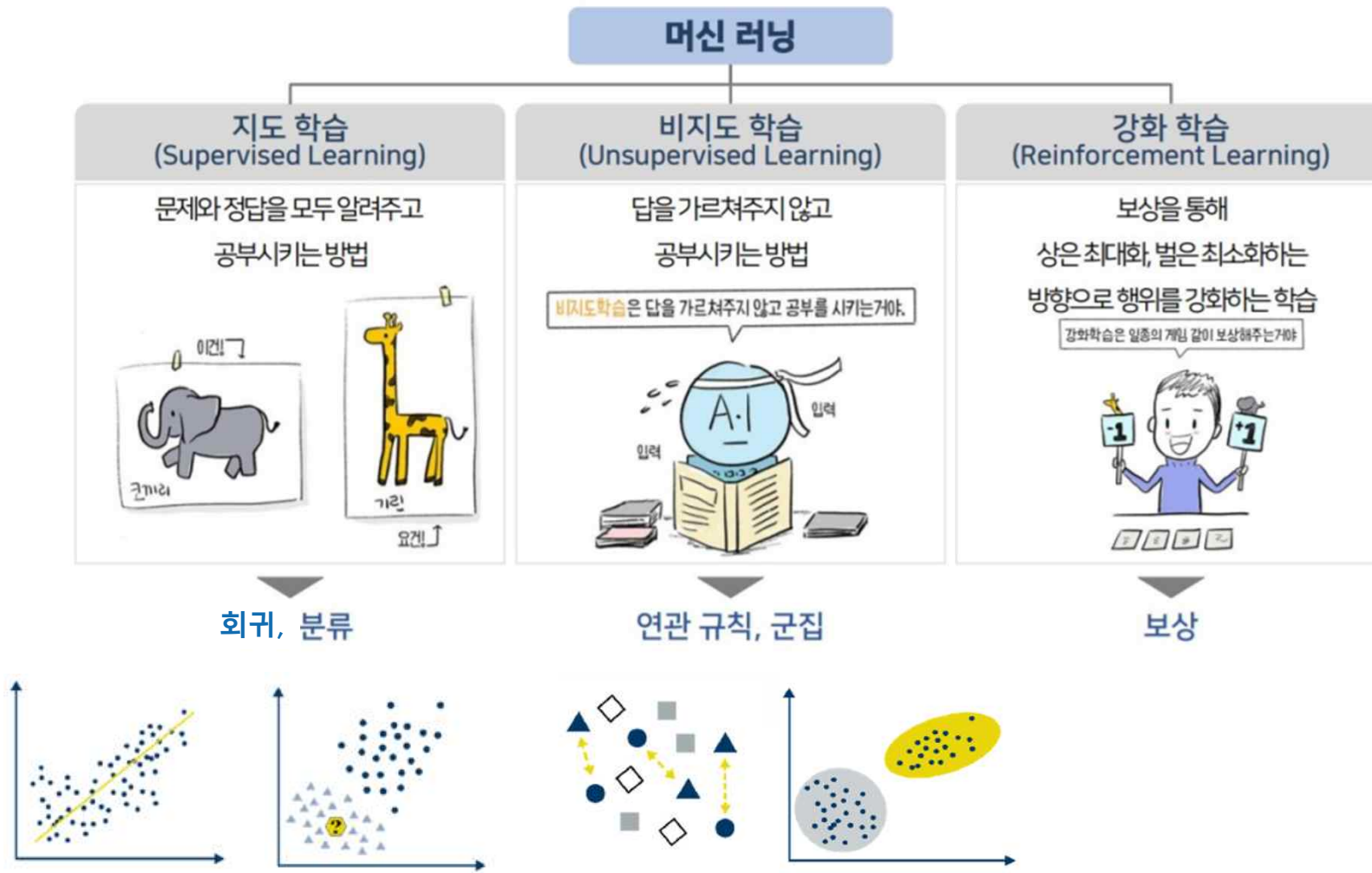


Deep Learning



<출처 : 메리츠종금증권 리서치센터>

머신러닝(Machine Learning)



지도학습(Supervised Learning)

- 정답(Label)을 알려주면서 데이터를 학습시키는 방식
- 회귀예측 (Regression)과 분류(Classification)를 할 수 있음.
- 활용사례: 주가예측, 스팸메일분류, 질병 유무 판단

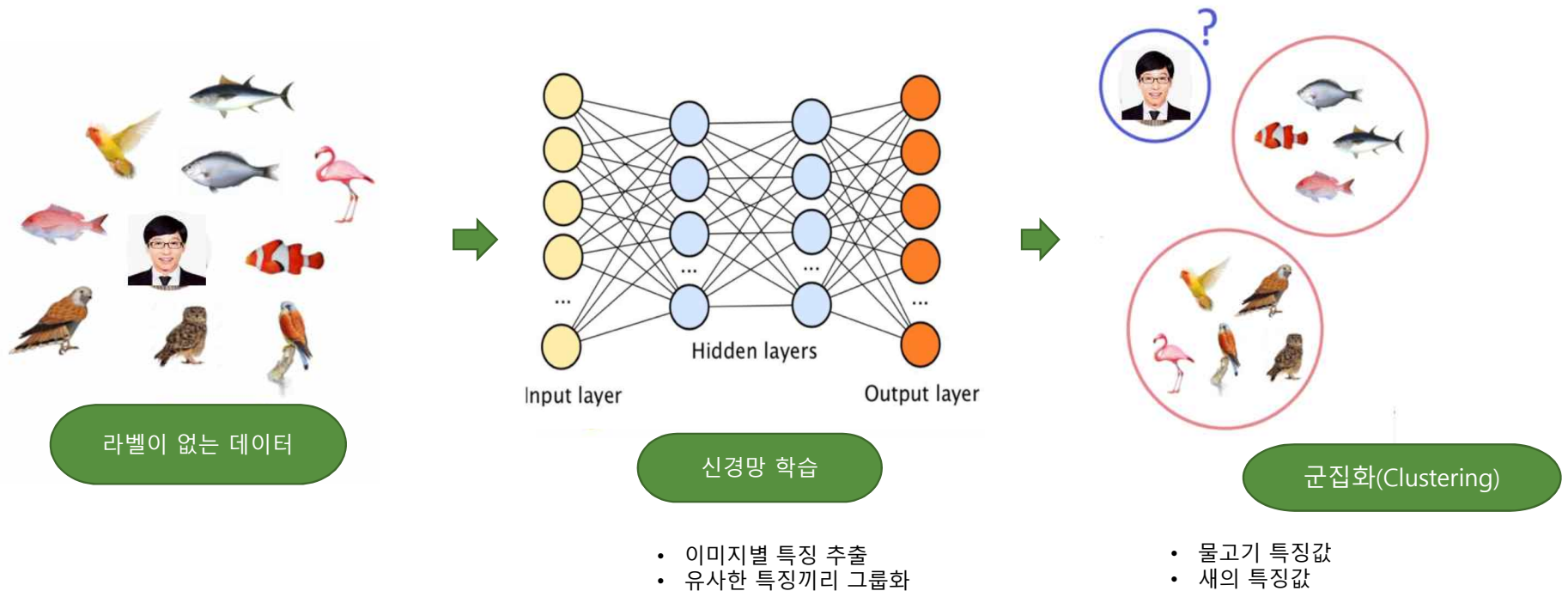
Target
Class
Label
종속변수
목표값
정답
Category



고양이

비지도학습(Unsupervised Learning)

- 정답(Label)을 알려주지 않고 데이터만 학습, 스스로 규칙과 패턴을 찾게 함
- 군집화(clustering, 그룹으로 묶기)와 연관규칙(데이터 간의 규칙을 찾음)등 있음
- 활용사례 온라인 쇼핑몰 구매자의 구매패턴을 분석하여 상품추천 등



강화학습(Reinforcement Learning)

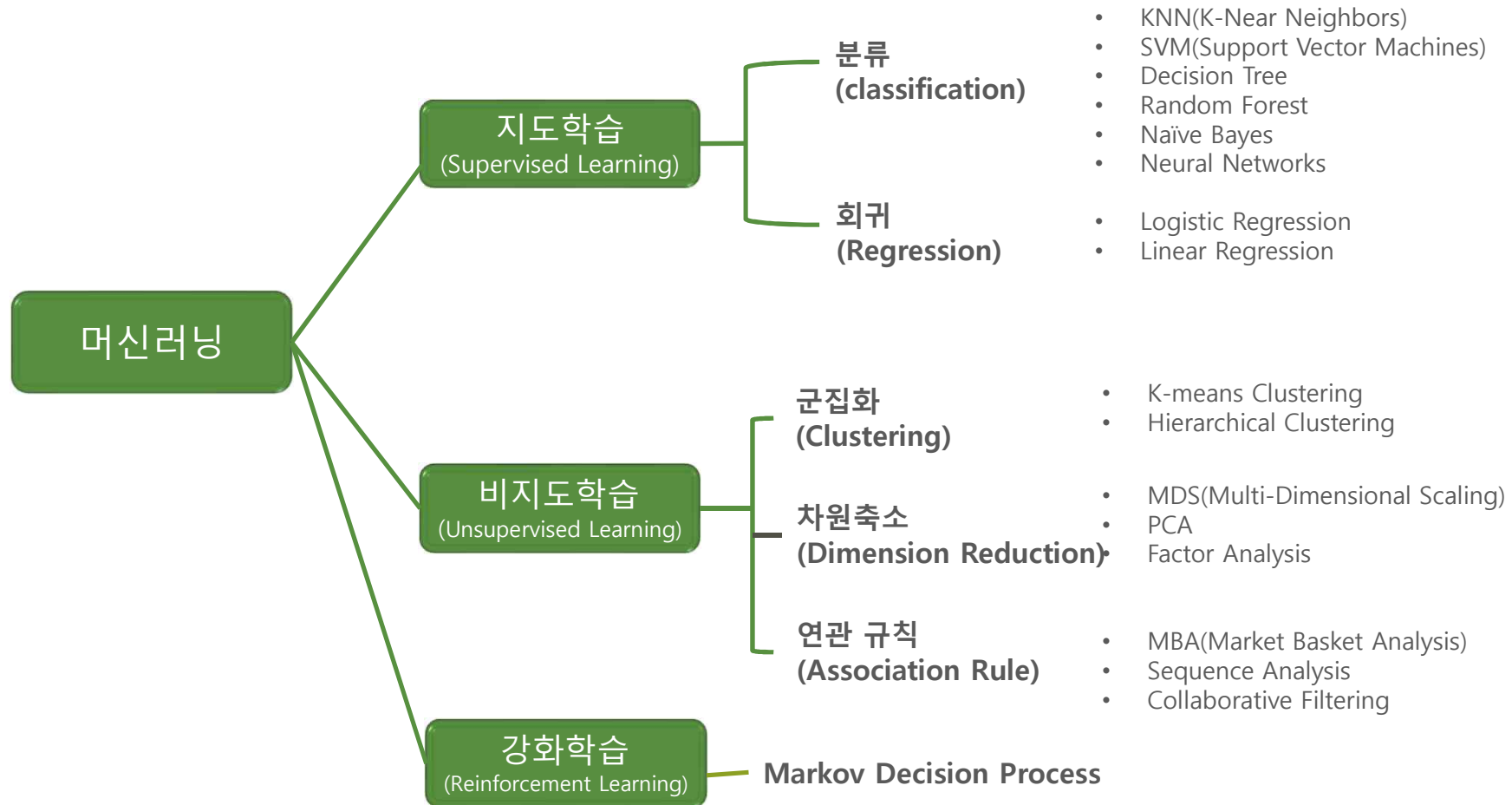
- 지도, 비지도학습처럼 많은 양의 데이터를 제공하지 않고 대신 상과 벌이라는 보상(Reward)을 주며 보상이 큰 쪽으로 학습하게 되는 방식으로 시행착오를 통해 스스로 학습 하는 방법
- 강화학습은 학습을 수행하고 행동을 하는 에이전트가 환경과 상호작용을 한 뒤 보상에 따라 행동을 결정하는 정책을 바꾸어 나가는 방식
- 활용사례 자율주행자동차 시뮬레이터, 지능형 로봇 개발, 게임 에이전트의 성능 향상



벽돌깨기 게임 by 딥마인드

<https://youtu.be/V1eYniJ0Rnk>

머신러닝 학습 알고리즘





데이터 분석이란? 데이터 시각화란?



데이터와 정보 (DIKW 피라미드)



데이터와 정보 (DIKW 피라미드)



눈오는 날 10% 할인 판촉으로 매출 더욱 높임
다른 계절도 날씨에 따른 매출 영향을 분석하여 수요 발견

눈오는 날 매출이 증가함.
배달이 되지 않아 근처 회사의 take-out 손님이 몰림.

일간 매출, 주간 매출, 월간 매출,
일 평균 매출

맥도널드 지점의 매출 데이터

데이터와 정보 (DIKW 피라미드)



A마트의 다른 상품들도 B마트보다 쌀 것이라고 예측

연필을 살 때는 A마트로 가는 것이 유리

A마트의 연필가격이 더 저렴함

A마트 펜 500원, 연필 200원, 라면 3000원..
B마트 연필 300원, 라면 2700원...

데이터 분석이 필요한 이유

- 1) 데이터는 모든 현상과 가설의 '**근거**'가 되는 정보를 담고 있음
- 2) 직접 경험하지 않아도 데이터를 통해 '**경험**'을 얻을 수 있음
- 3) 데이터분석을 통해 데이터에서 의미를 찾고 미래를 예측하는 등 **새로운 가치를 창출**할 수 있음.



비즈니스에 대한 정보가 매 순간
데이터로 기록

↓

데이터의 종류, 양, 접근성 ↑

↓

데이터를 이해하고 활용할 수 있는
능력

↓

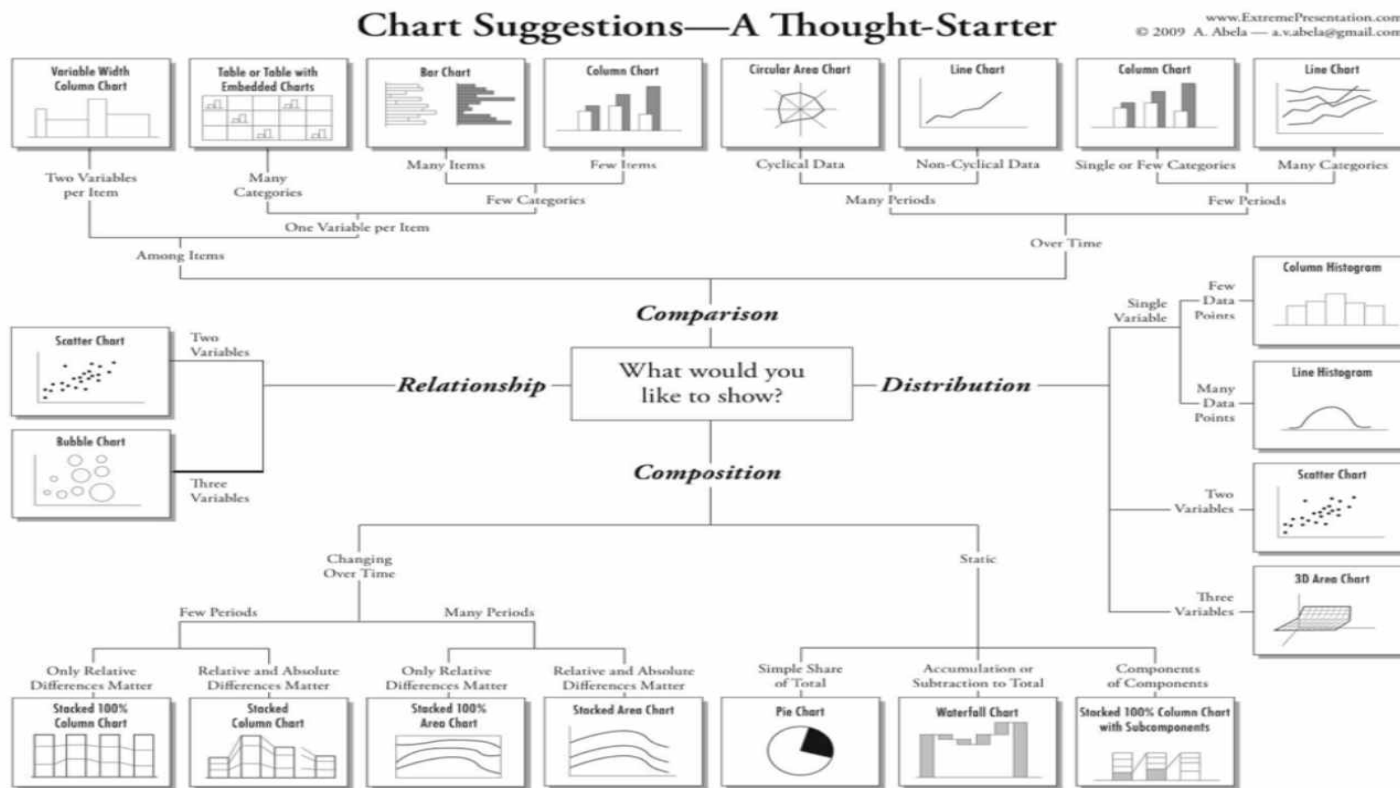
데이터 분석을 통한 가치창출

a

aimap marker May 21, 2019 · 9 min read

<https://medium.com/@aimap.marker>

데이터 시각화(Data Visualization)



인포그래픽으로 표현한 2012년 예산안 (단위 : 조원)



[출처 : Andrew Abela, "The Extreme Presentation method", <http://www.extremepresentation.com>, 2006]



데이터 분석 프로세스

- 데이터 분석의 목적은 문제 해결

문제 정의

데이터 분석의 목적은 문제해결

데이터 정의

필요한 데이터의 속성은?

데이터 수집

설문, 공공데이터(2차원 표 형태) 등

데이터 전 처리

이상치, 결측치 처리 데이터 형태 변환

데이터 시각화

데이터 이해

결과해석

결론 도출, 문제해결

데이터 분석 프로세스

- 문제 해결에 필요한 데이터의 속성을 정의.
데이터의 구체적인 정보 항목으로
더 이상 분리될 수 없는 최소의 데이터 보관 단위

문제 정의

데이터 분석의 목적은 문제해결

데이터 정의

필요한 데이터의 속성은?

데이터 수집

설문, 공공데이터(2차원 표 형태) 등

데이터 전 처리

이상치, 결측치 처리 데이터 형태 변환

데이터 시각화

데이터 이해

결과해석

결론 도출, 문제해결

데이터 분석 프로세스

문제 정의

데이터 분석의 목적은 문제해결

데이터 정의

필요한 데이터의 속성은?

데이터 수집

설문, 공공데이터(2차원 표 형태) 등

데이터 전 처리

이상치, 결측치 처리 데이터 형태 변환

데이터 시각화

데이터 이해

결과해석

결론 도출, 문제해결

빅데이터는 어디에서 수집되는가?

대부분의 기업들은 빅데이터 활용시 필요한 인사이트를 확보하기 위해 내부 데이터를 분석하는데 초점이 맞춰져 있다. 일부 조직에서는 소셜 미디어와 같은 방화벽 너머의 데이터까지도 주목한다.



제공되는 데이터, 직접 수집, 웹 크롤링/웹스크래핑

공공데이터 활용지표 :

http://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=2844

CSV	XLSX	JSON	XML	API																																							
데이터를 콤마 (,)로 구분	엑셀 포맷 (행, 열)	Key:Value 구조	마크업	개발자를 위한 인터페이스																																							
날짜,지점,평균,최저,최고 2021-07-24,108,31.7,26.9,36.5 2021-07-25,108,31.5,27.2,35.9 2021-07-26,108,31.2,27.4,35.4 2021-07-27,108,31.1,27.8,35.7 2021-07-28,108,30.4,27.1,34.7 2021-07-29,108,29.6,27,33.3 2021-07-30,108,30.5,25.8,35.4 2021-07-31,108,29.8,26.8,34.3 2021-08-01,108,27.1,25.1,28.8 2021-08-02,108,26.5,25,28.6	<table><tr><th>A</th><th>B</th><th>C</th></tr><tr><td>날짜</td><td>지점</td><td>평균기온(°C)</td></tr><tr><td>2021-07-24</td><td>108</td><td>31.7</td></tr><tr><td>2021-07-25</td><td>108</td><td>31.5</td></tr><tr><td>2021-07-26</td><td>108</td><td>31.2</td></tr><tr><td>2021-07-27</td><td>108</td><td>31.1</td></tr><tr><td>2021-07-28</td><td>108</td><td>30.4</td></tr><tr><td>2021-07-29</td><td>108</td><td>29.6</td></tr><tr><td>2021-07-30</td><td>108</td><td>30.5</td></tr><tr><td>2021-07-31</td><td>108</td><td>29.8</td></tr><tr><td>2021-08-01</td><td>108</td><td>27.1</td></tr><tr><td>2021-08-02</td><td>108</td><td>26.5</td></tr><tr><td>2021-08-03</td><td>108</td><td>28</td></tr></table>	A	B	C	날짜	지점	평균기온(°C)	2021-07-24	108	31.7	2021-07-25	108	31.5	2021-07-26	108	31.2	2021-07-27	108	31.1	2021-07-28	108	30.4	2021-07-29	108	29.6	2021-07-30	108	30.5	2021-07-31	108	29.8	2021-08-01	108	27.1	2021-08-02	108	26.5	2021-08-03	108	28	<pre>{ "menu": "회원관리", "member": [{ "id": "dragon", "name": "홍길동", "address": "서울시", "job": "학생" }, { "id": "dragon", "name": "홍길서", "address": "제주시", "job": "교사" }]}</pre>	<pre><?xml version="1.0" encoding="euc-kr" ?> <거래내역서> <내역> <물품번호>AA1</물품번호> <물품명>pen</물품명> <단가 unit="dollar">1.0</단가> <수량>30</수량> <총액 unit="dollar">30</총액> <거래연도>2001</거래연도> <거래월>01</거래월> <거래일>20</거래일> </내역> <내역> <거래내역서></pre>	
A	B	C																																									
날짜	지점	평균기온(°C)																																									
2021-07-24	108	31.7																																									
2021-07-25	108	31.5																																									
2021-07-26	108	31.2																																									
2021-07-27	108	31.1																																									
2021-07-28	108	30.4																																									
2021-07-29	108	29.6																																									
2021-07-30	108	30.5																																									
2021-07-31	108	29.8																																									
2021-08-01	108	27.1																																									
2021-08-02	108	26.5																																									
2021-08-03	108	28																																									

<https://www.itworld.co.kr/news/78701>

데이터 분석 프로세스

문제 정의

데이터 분석의 목적은 문제해결

데이터 정의

필요한 데이터의 속성은?

데이터 수집

설문, 공공데이터(2차원 표 형태) 등

데이터 전 처리

이상치, 결측치 처리 데이터 형태 변환

데이터 시각화

데이터 이해

결과해석

결론 도출, 문제해결

- 1) 해결하려는 문제에 필요한 속성만을 데이터셋에서 추출
- 2) 결과에 영향을 미치지 않도록
결측값/이상값이 포함된 데이터는 삭제 하거나 대표할 수 있는 값으로 채운다(대표값, 평균값, 최빈값, 중앙값)

Garbage in garbage out!!

	A	B	C	D	E
1	날짜	지점	평균기온(°)	최저기온(°)	최고기온(°)
2	2021-07-24	108	31.7	26.9	36.5
3	2021-07-25	108	31.5	27.2	35.9
4	2021-07-26	108	31.2	27.4	35.4
5	2021-07-27	108	31.1	27.8	35.7
6	2021-07-28	108	30.4	27.1	34.7

데이터 분석 프로세스

문제 정의

데이터 분석의 목적은 문제해결

데이터 정의

필요한 데이터의 속성은?

데이터 수집

설문, 공공데이터(2차원 표 형태) 등

데이터 전 처리

이상치, 결측치 처리 데이터 형태 변환

데이터 시각화

데이터 이해

결과해석

결론 도출, 문제해결



https://m.mt.co.kr/renew/view.html?no=2018080613120727477&ca=#_enliple

데이터 분석 프로세스

문제 정의

데이터 분석의 목적은 문제해결

데이터 정의

필요한 데이터의 속성은?

데이터 수집

설문, 공공데이터(2차원 표 형태) 등

데이터 전 처리

이상치, 결측치 처리 데이터 형태 변환

데이터 시각화

데이터 이해

결과해석

결론 도출, 문제해결

- 1) 많은 데이터를 한눈에 파악할 수 있다.
- 2) 데이터의 변화, 데이터 사이의 관계 등을 쉽게 볼 수 있어 데이터를 깊이 있게 이해할 수 있다.
- 3) 수치로만 파악하기 힘든 패턴이나 새로운 정보를 발견할 수 있다.
- 4) 인공지능이 학습할 데이터를 점검할 수 있다.
- 5) 인공지능이 바르게 일을 처리했는지 확인할 수 있다.

미세먼지 데이터(2009~2018)
AA용지 61428호

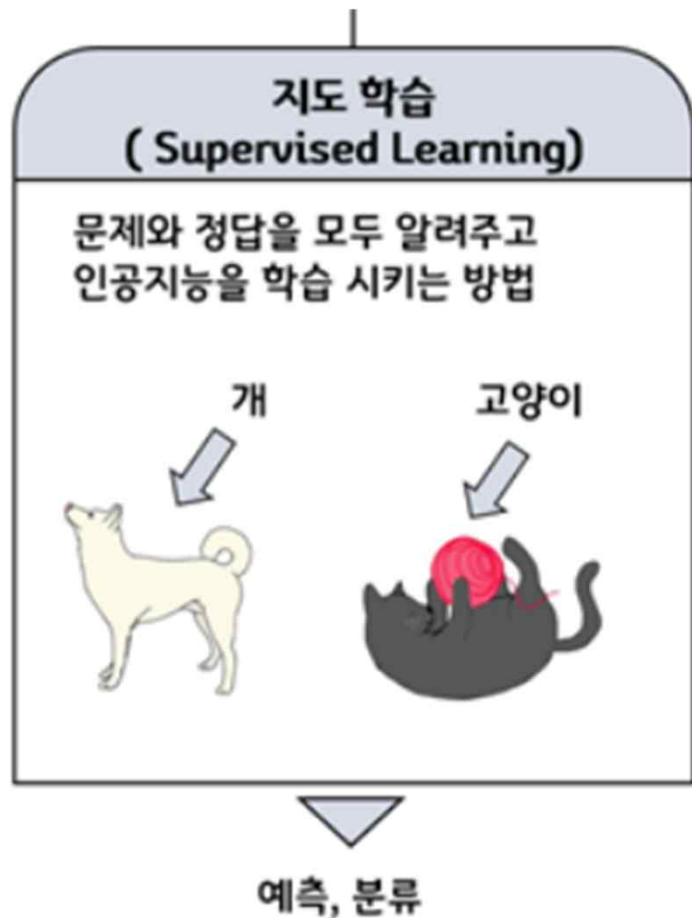
연월일시	미세먼지	초미세먼지
2 2009-01-01 00:00:00	38.0	8.0
3 2009-01-01 01:00:00	44.0	10.0
4 2009-01-01 02:00:00	29.0	24.0
5 2009-01-01 03:00:00	31.0	17.0
6 2009-01-01 04:00:00	34.0	15.0
7 2009-01-01 05:00:00	38.0	8.0
8 2009-01-01 06:00:00	33.0	26.0
9 2009-01-01 07:00:00	42.0	18.0
10 2009-01-01 08:00:00	48.0	17.0
11 2009-01-01 09:00:00	32.0	13.0
12 2009-01-01 10:00:00	36.0	16.0
13 2009-01-01 11:00:00	35.0	18.0
14 2009-01-01 12:00:00	41.0	11.0
15 2009-01-01 13:00:00	38.0	18.0
16 2009-01-01 14:00:00	31.0	18.0
17 2009-01-01 15:00:00	46.0	22.0
18 2009-01-01 16:00:00	48.0	21.0
19 2009-01-01 17:00:00	32.0	15.0
20 2009-01-01 18:00:00	44.0	16.0
21 2009-01-01 19:00:00	40.0	14.0
22 2009-01-01 20:00:00	44.0	23.0
23 2009-01-01 21:00:00	38.0	17.0
24 2009-01-01 22:00:00	31.0	13.0
25 2009-01-01 23:00:00	31.0	23.0
26 2009-01-02 00:00:00	56.0	13.0
27 2009-01-02 01:00:00	22.0	16.0
28 2009-01-02 02:00:00	32.0	20.0
29 2009-01-02 03:00:00	38.0	23.0
30 2009-01-02 04:00:00	38.0	14.0
31 2009-01-02 05:00:00	53.0	18.0
32 2009-01-02 06:00:00	58.0	20.0
33 2009-01-02 07:00:00	57.0	24.0
34 2009-01-02 08:00:00	58.0	18.0
35 2009-01-02 09:00:00	57.0	27.0

1장으로
시각화



※출처 : 학교에서 만나는 인공지능

분석을 위한 데이터의 성격 정의하기



Target
Class
Label
종속변수
목표값
정답
Category

VS.

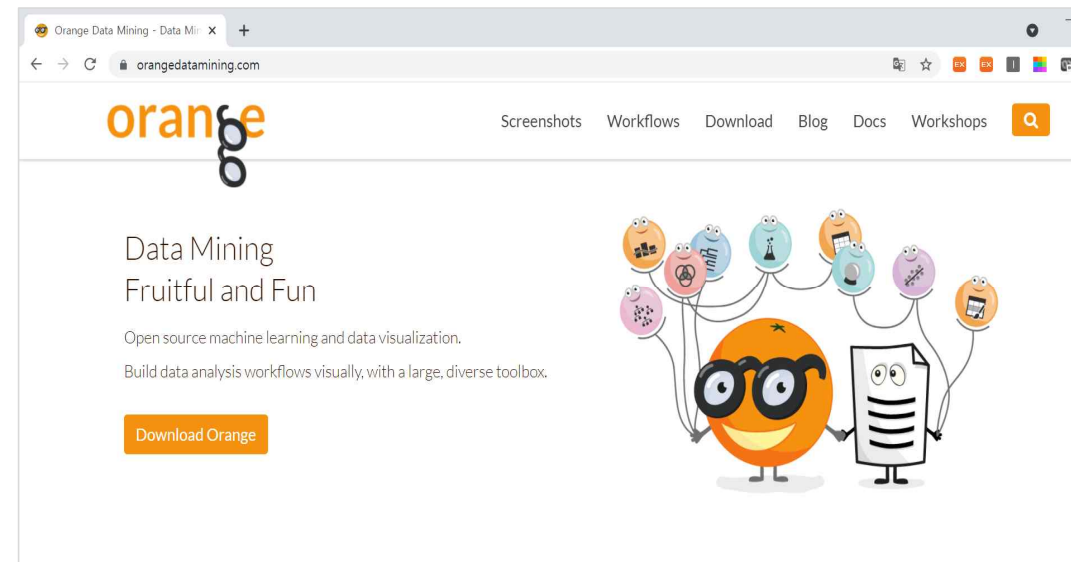
Feature
Attributes
독립변수
문제
속성

orange 설치하기

Orange3 란?

- 오픈소스 데이터 시각화 및 머신러닝을 이용한 데이터 분석 도구
- 1997년 슬로베니아 루블라냐 대학에서 개발, 2015년 Orange 3.0 배포
- 데이터 분석을 위한 컴포넌트 기반 비주얼 프로그래밍 소프트웨어
- 데이터 분석 워크플로우를 시각적으로 구축

코딩과 수학 없이 드래그 앤 드롭으로
머신 러닝을 이용한 데이터 분석 가능



<https://orangedatamining.com/>

Orange 설치

<https://orangedatamining.com/>

orange



Screenshots

Workflows

Download

Blog

Docs

Workshops



Donate

1



Windows



macOS



Linux / Source

2

Download the latest version for Windows

Download Orange 3.32.0

Standalone installer (default)

Orange3-3.32.0-Miniconda-x86_64.exe (64 bit)

Can be used without administrative privileges.

Portable Orange

Orange3-3.32.0.zip

No installation needed. Just extract the archive and open the shortcut in the extracted folder.

가장 최신버전으로 설치하세요 ^^

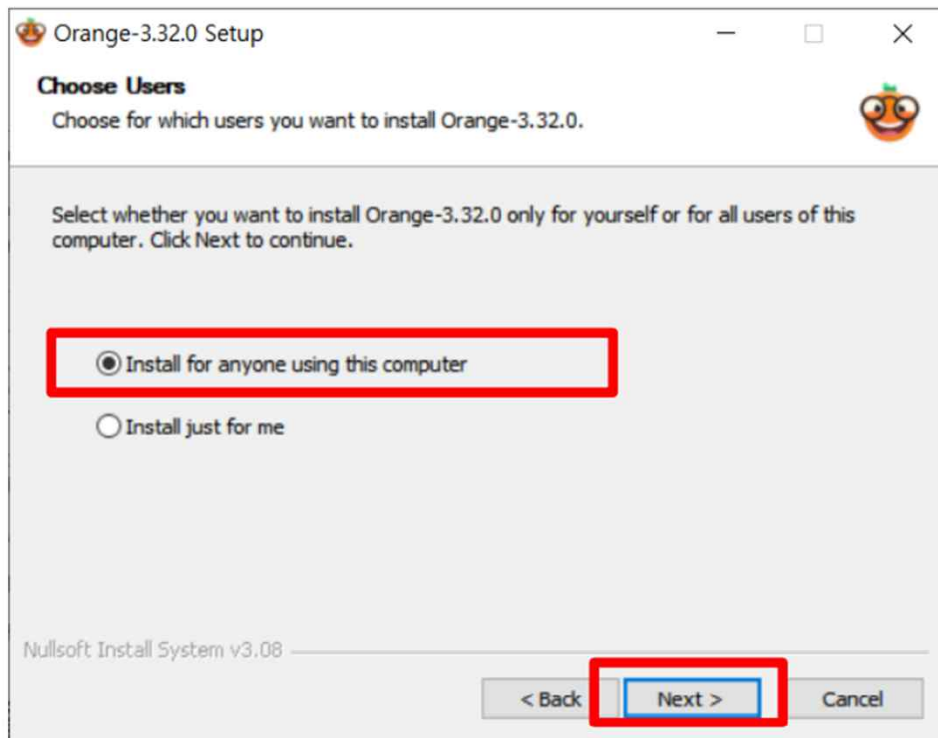
Orange3-3.33.0-Miniconda-x86_64.exe

설치파일은 533MB 정도이고, 설치를 모두 하고 나면 1GB가 훨씬 넘습니다.

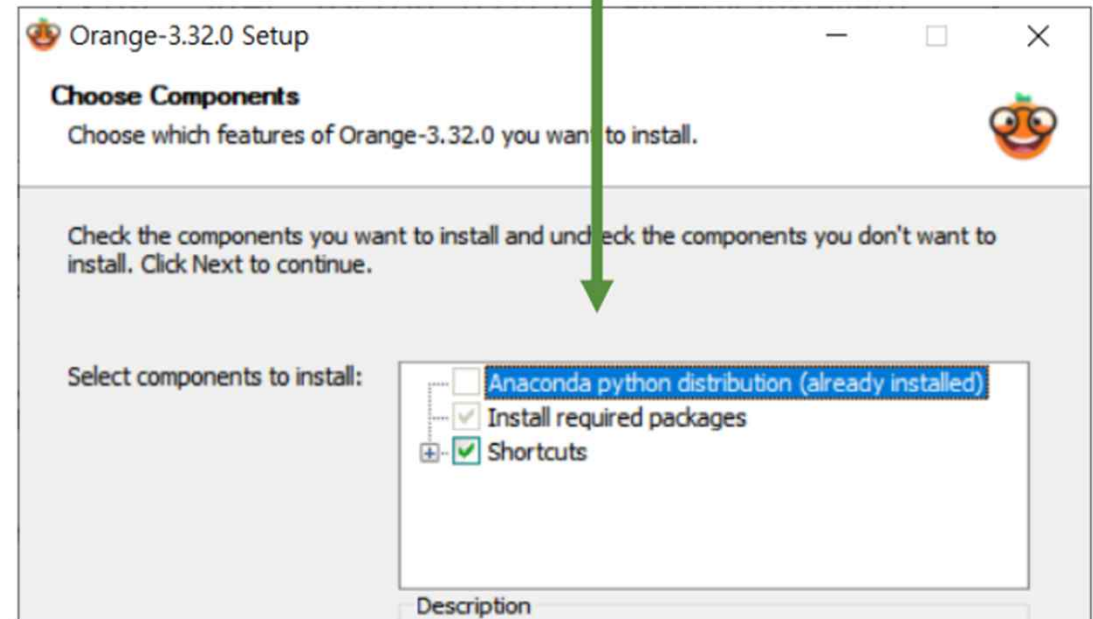
파일이 다운로드 되면 실행시켜 **관리자권한**으로 설치 시작합니다

Orange 설치

<https://orangedatamining.com/>



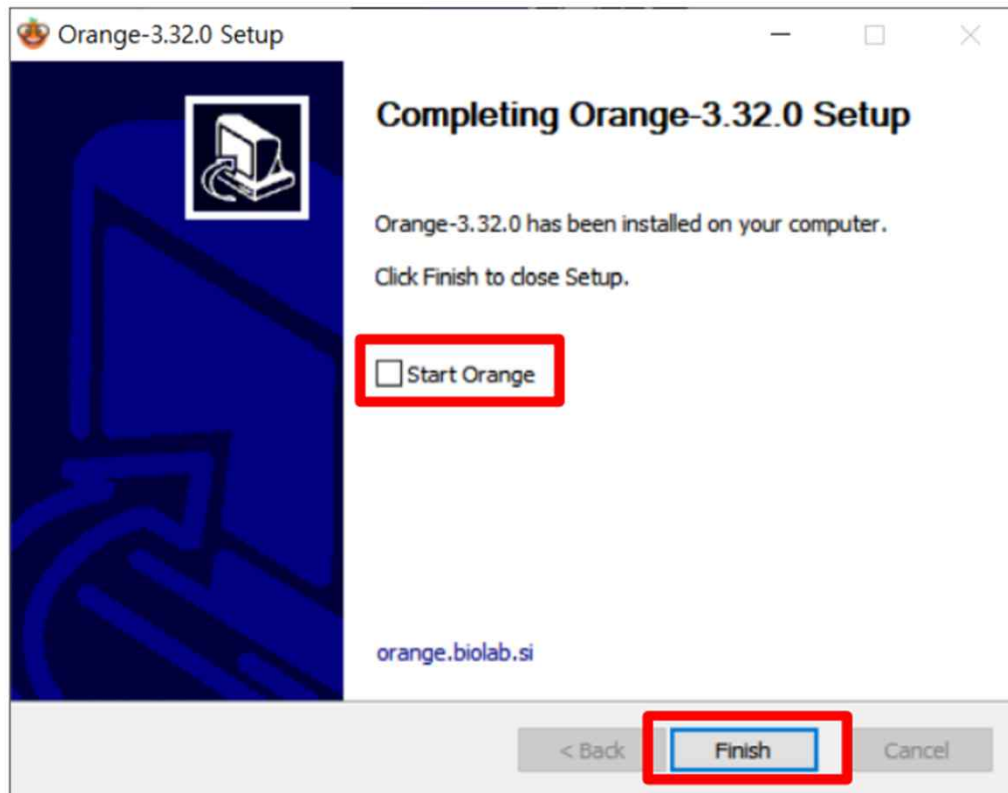
Anaconda가 설치되어 있지 않은 경우
(저는 이미 설치되어 있습니다. already installed)
Miniconda를 설치하는 과정이 더 있을 수 있습니다.
정상적인 과정이니 모두 설치하세요.



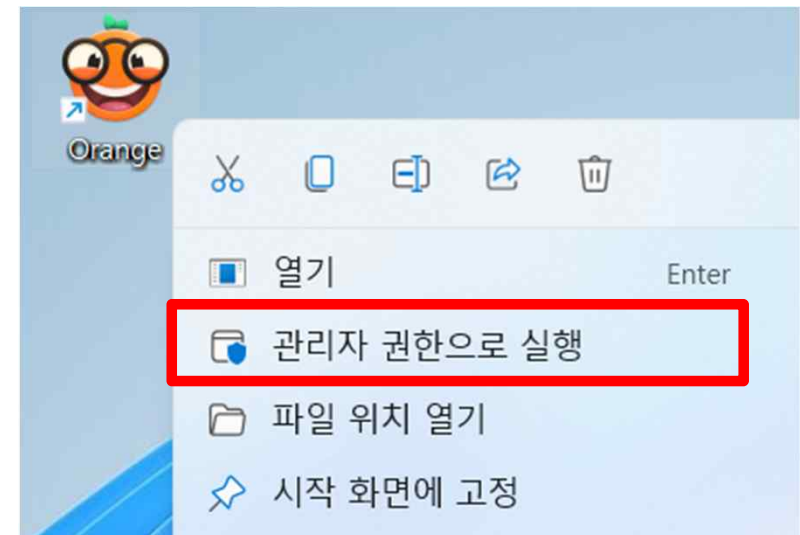
Orange 설치

<https://orangedatamining.com/>

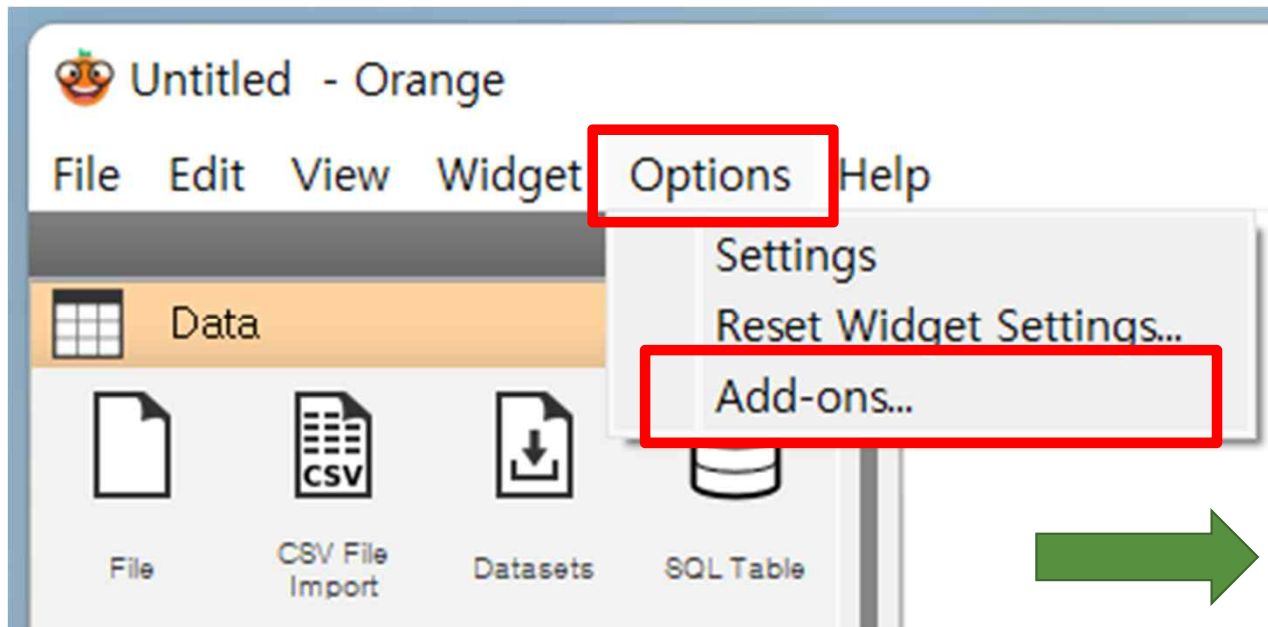
설치에 시간이 오래 걸립니다.....



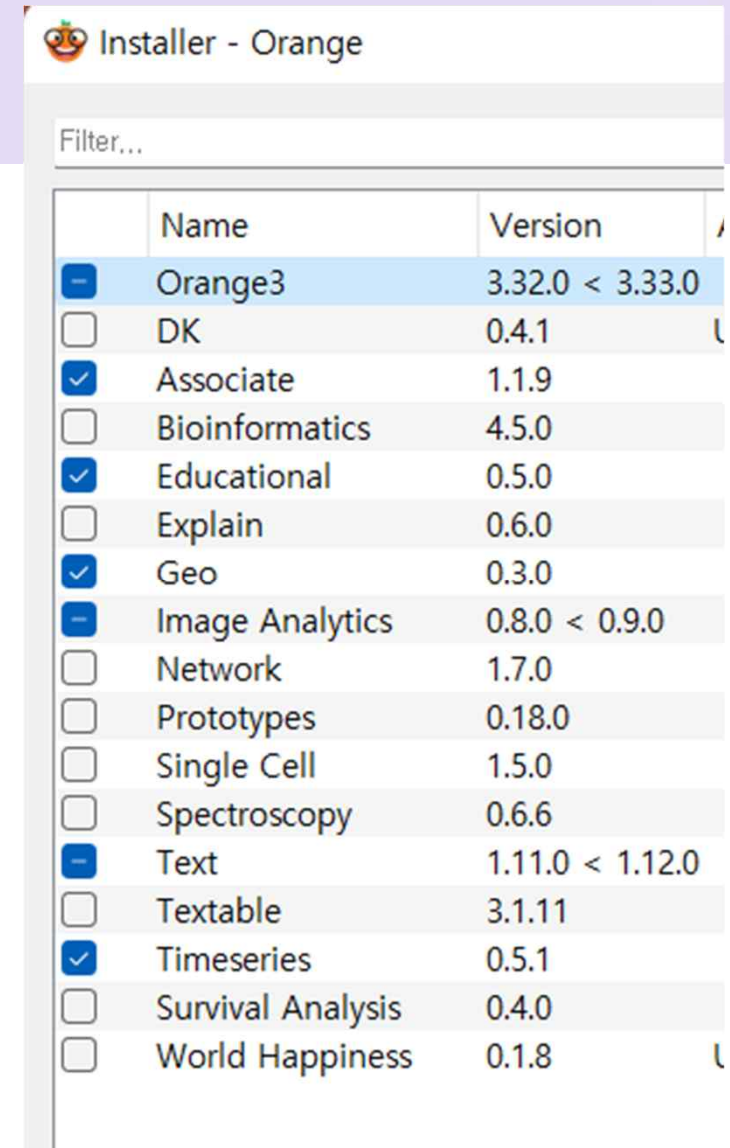
설치를 마치면 Add-on 설치를 위해 바탕화면에서 관리자 권한으로 오렌지를 실행합니다.



Orange3 Add-on 설치

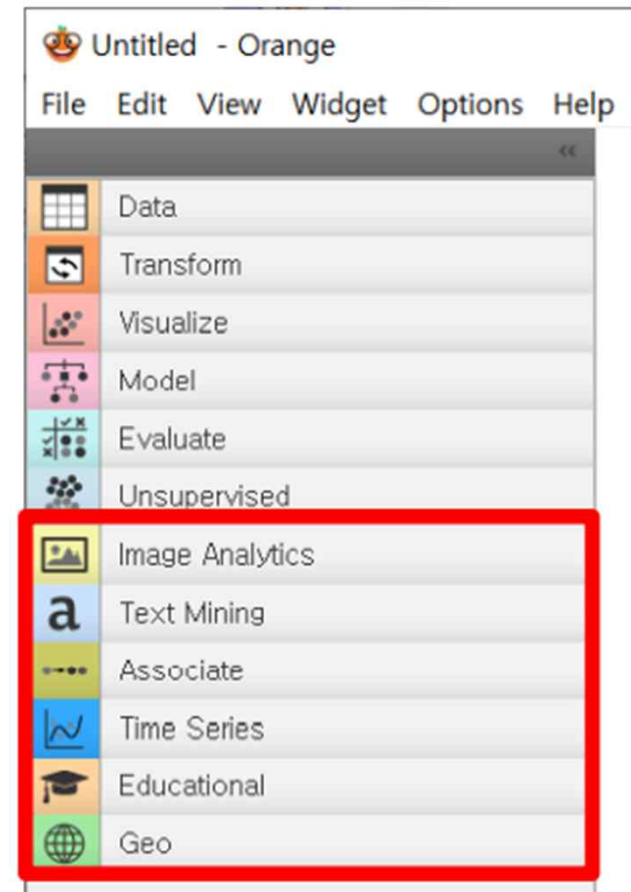
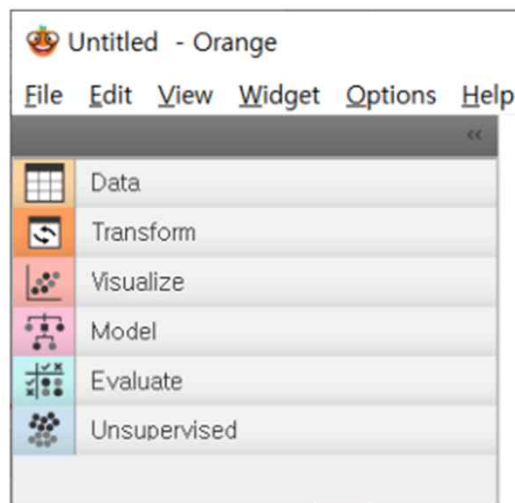
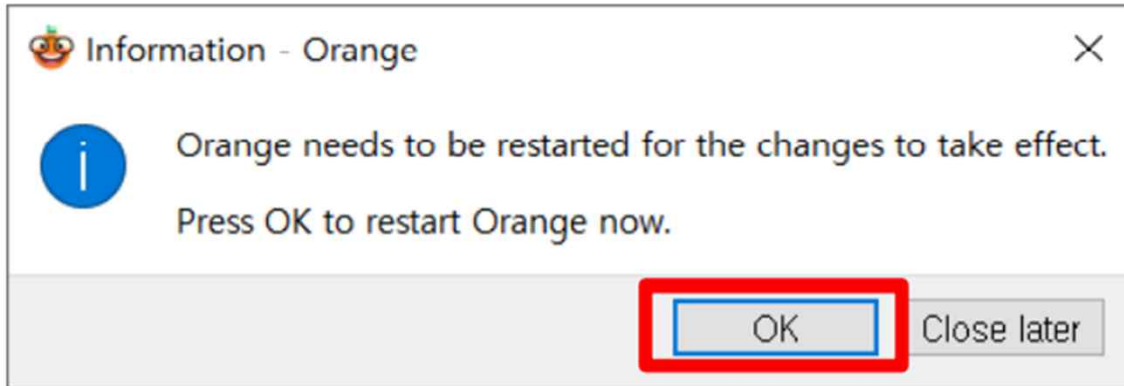


Associate, Educational, Geo, Image Analytics, Text, Timeseries 6개를 체크하고 OK를 누릅니다



Orange3 Add-on 설치

오렌지가 종료되었다가 자동으로 재실행된 후 새로운 위젯이 설치되었음을 확인할 수 있습니다.





다음 시간에는 오렌지에서의 데이터 흐름과 위
젯의 활용에 대해 알아보고 정형데이터를 활용
한 회귀 분석을 통해 회귀분석의 개념과 절차에
대해 공부해 보도록 하겠습니다. 감사합니다.

