

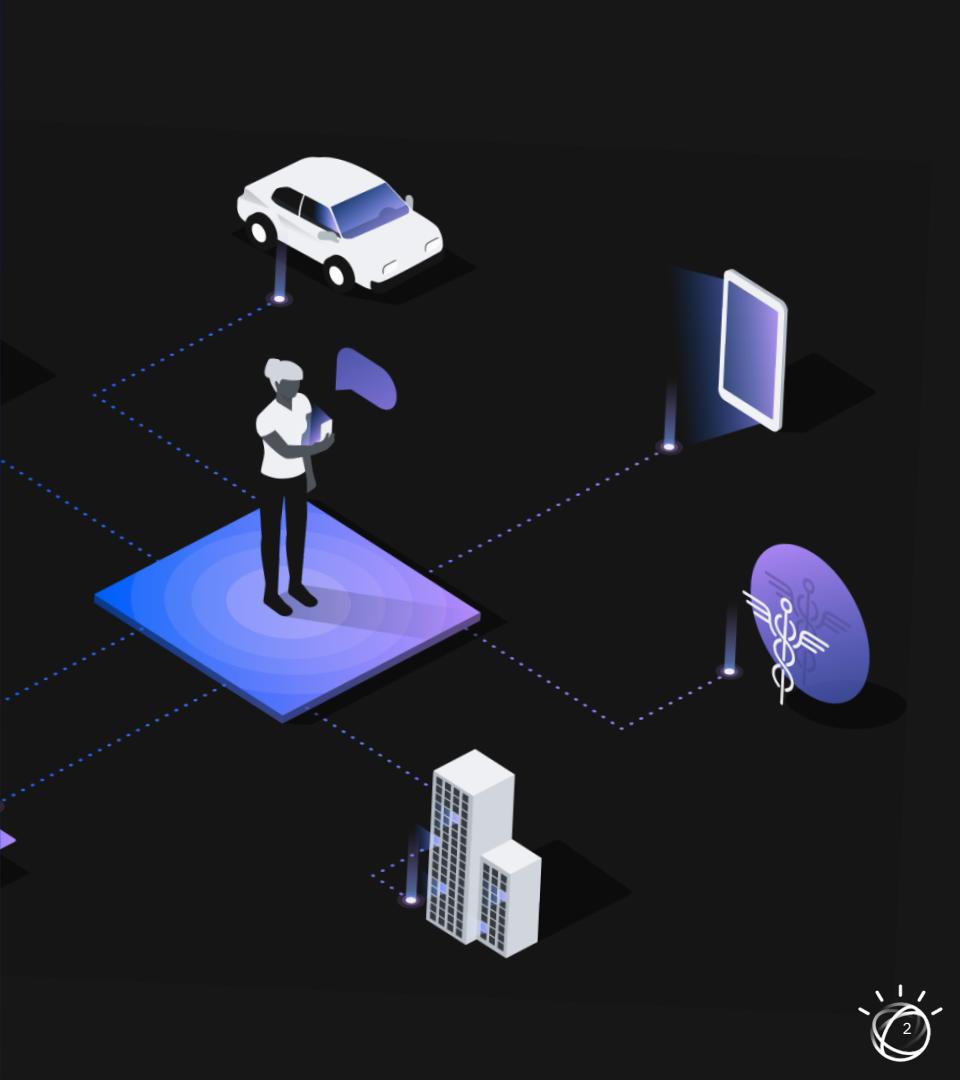
데이터 사이언스

인공지능 / 기계학습 알고리즘
&
왓슨 어플리케이션

—

Ho-Kyeong Ra, Ph.D.

나호경



인공지능 (Artificial Intelligence) 이란?

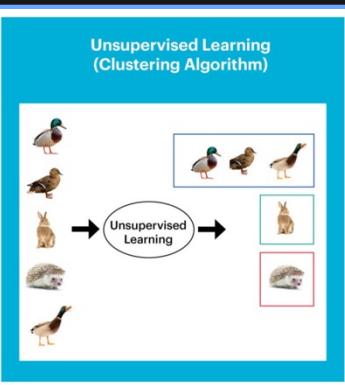
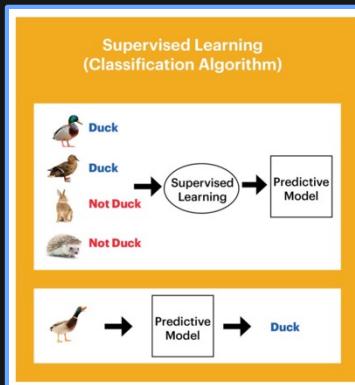


인공지능(AI)? 기계학습(Machine Learning)?

기계학습(Machine Learning)이란?

기계 학습은 컴퓨터 프로그램이 경험을 통해
자동으로 향상되도록하는 컴퓨터 알고리즘의
연구입니다.

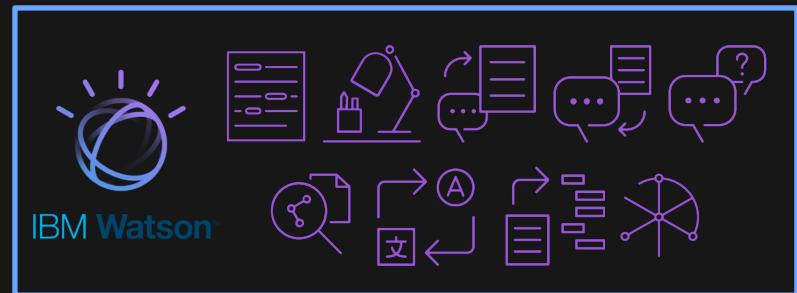
Tom Mitchell, Machine Learning Department,
Carnegie Mellon University



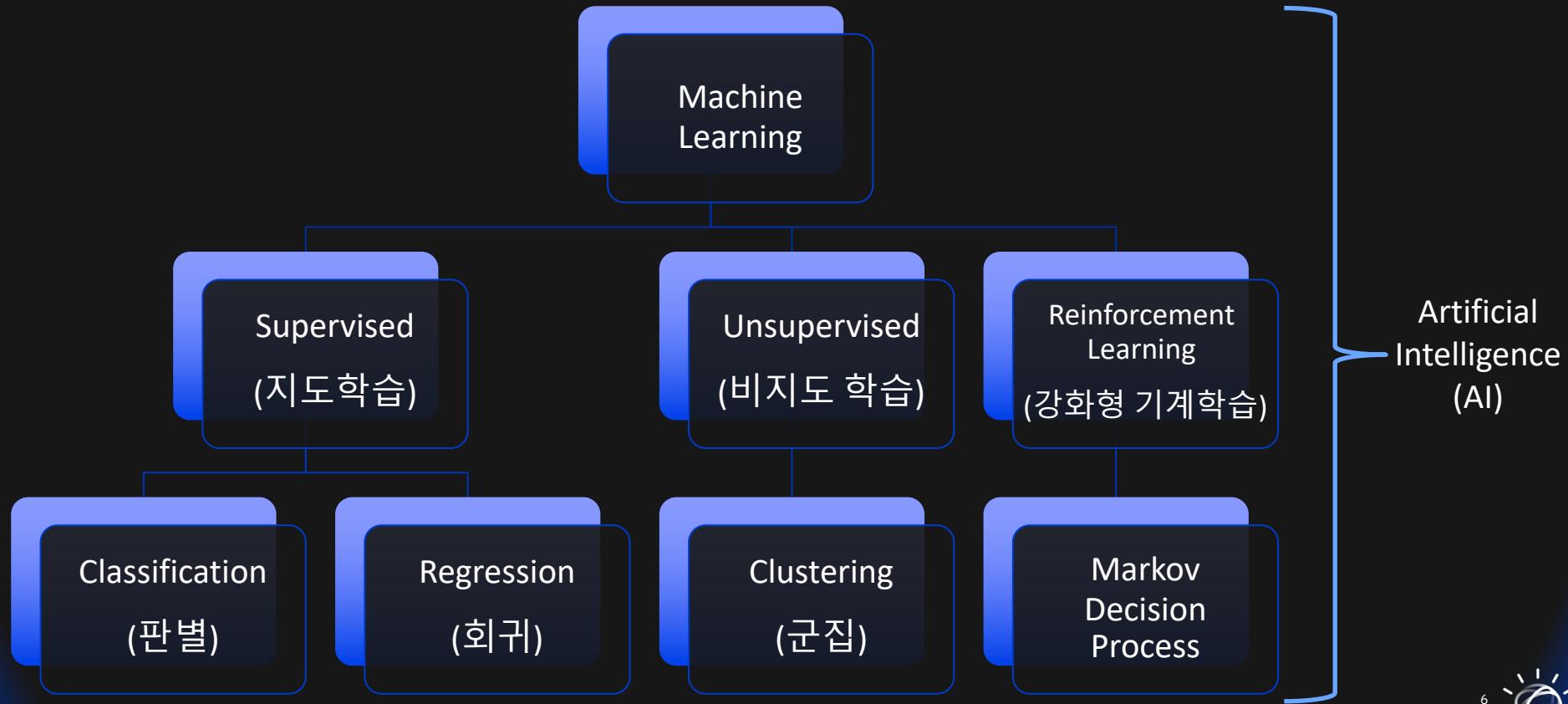
인공지능(AI) 이란?

AI는 인간이 보유하고 있지만 기계가 보유하지 않는
능력을 기반으로 움직이는 어플리케이션/기능 입니다.

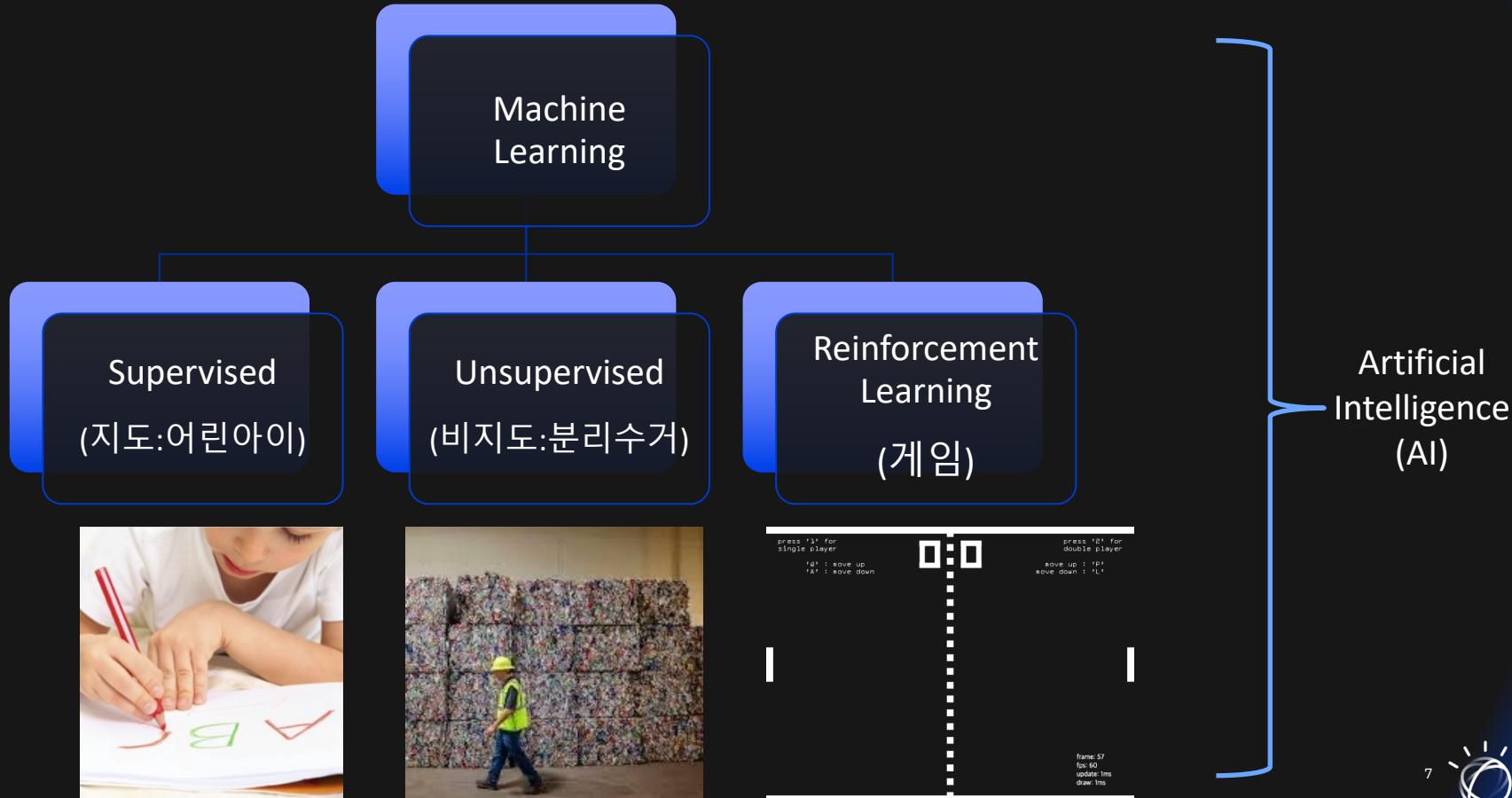
Zachary Lipton, Assistant Professor and Researcher,
Carnegie Mellon University



AI 및 기계 학습의 계층적 관점



AI 및 기계 학습의 계층적 관점 예제



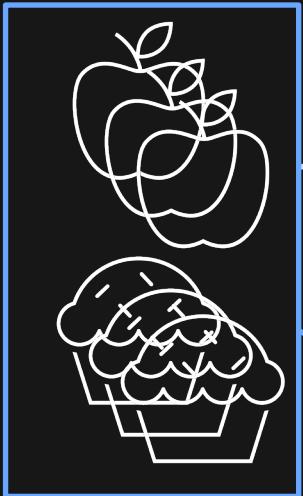
기계학습에서 “특징(feature)”이 무엇인가요?

기계 학습 및 패턴 인식에서 특징은 관찰되는 현상의 개별 측정 가능한 속성 또는 특성입니다.



기계학습 알고리즘 (지도 vs. 비지도 학습)

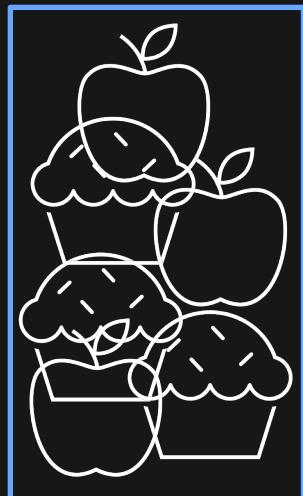
지도 학습



레이블 지정
데이터

테스트 도중에는 레이블링이
불필요하나, 모델 검증에서
이용되기도 합니다.

비지도 학습



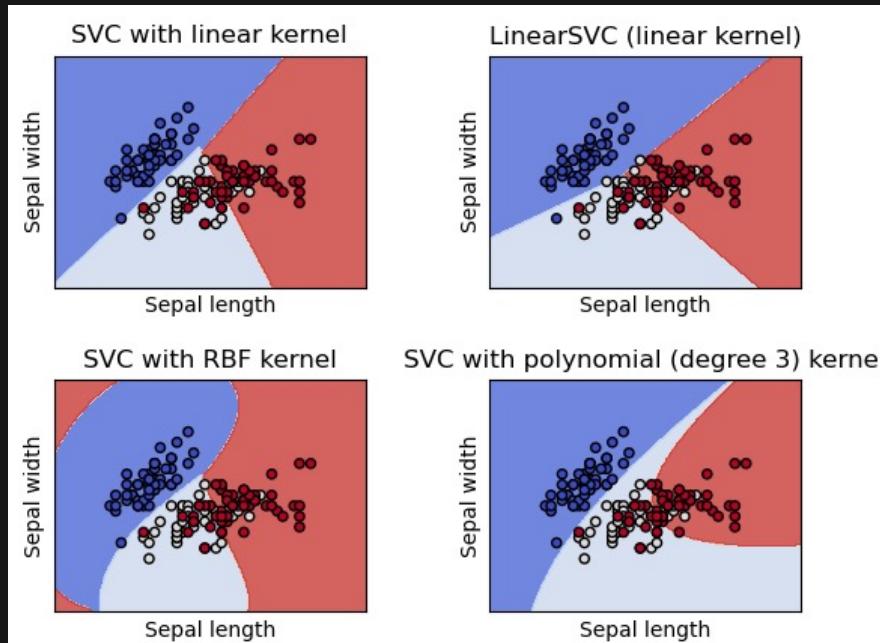
레이블 미지정
데이터



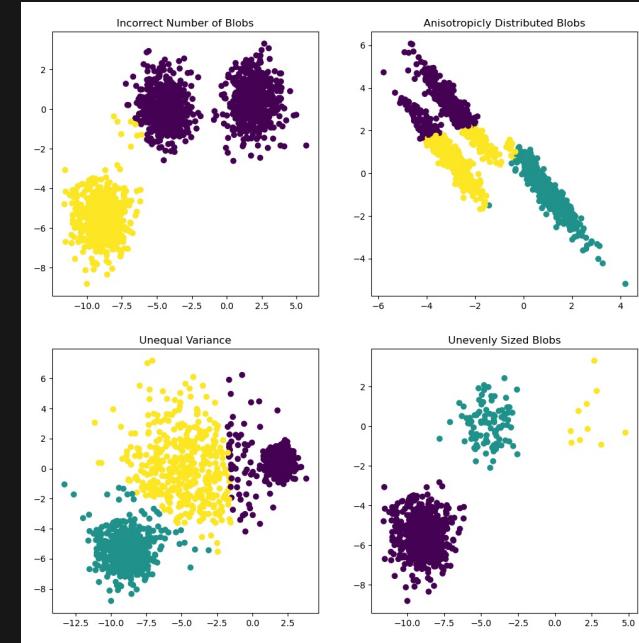
데이터 군집
부분 표시



기계학습 알고리즘 (지도 vs. 비지도 학습 실제 예제)



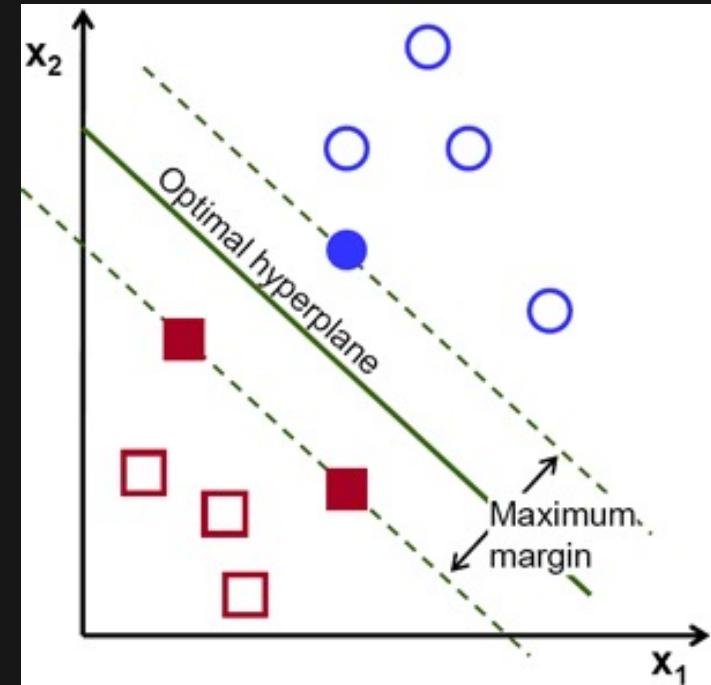
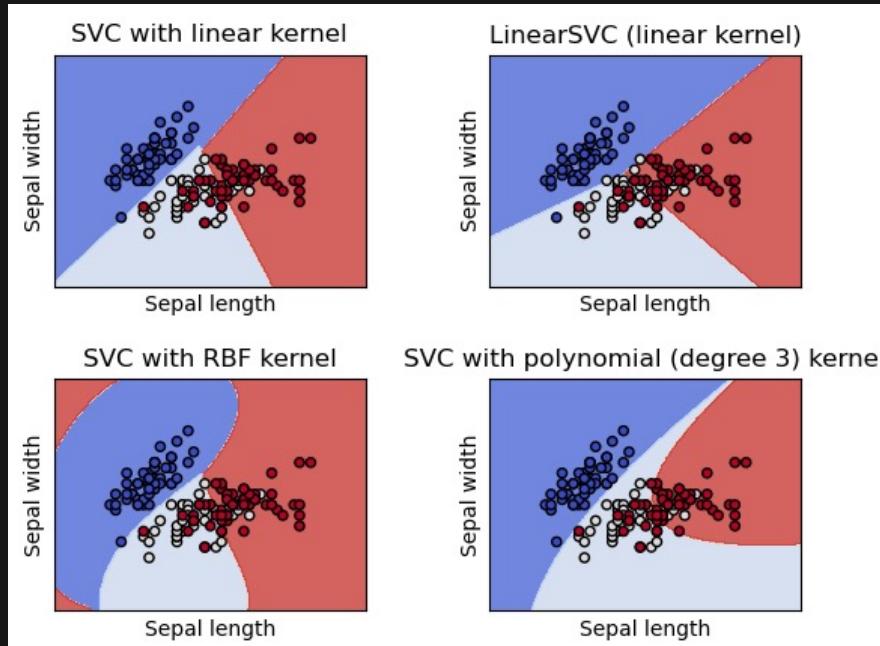
서포트 벡터 머신
(지도학습 예제: Support Vector Machine, SVM)



K - 평균 알고리즘
(비지도 학습 예제: K-Means)



기계학습 알고리즘 (지도학습 알고리즘 예제 – 서포트 벡터 머신 SVM)



기계학습 알고리즘 (구현 예제: 지도 학습 - SVM)

데이터 준비

학습 데이터(X)	클래스(y)
0, 0	0.5
2, 2	2.5

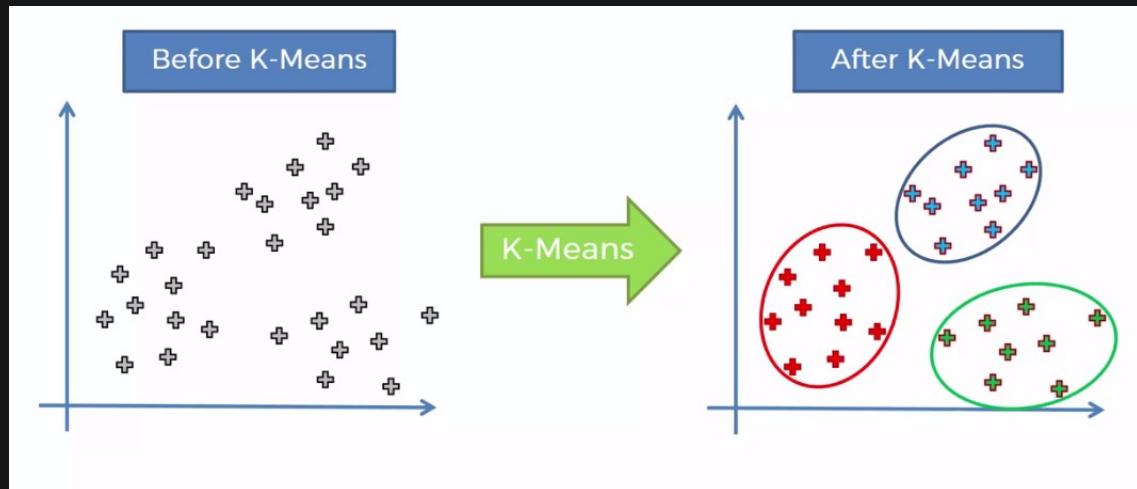
알고리즘/모델 기반 선택

학습

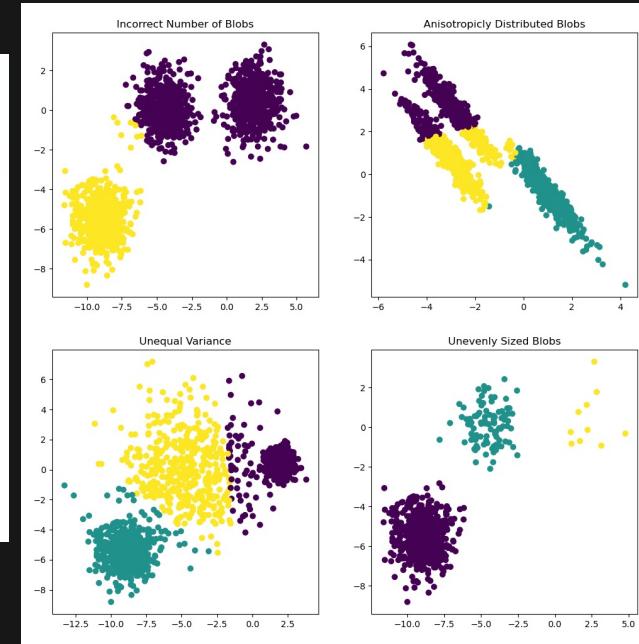
테스트

```
>>> from sklearn import svm
>>> X = [[0, 0], [2, 2]]
>>> y = [0.5, 2.5]
>>> regr = svm.SVR()
>>> regr.fit(X, y)
SVR()
>>> regr.predict([[1, 1]])
```

기계학습 알고리즘 (비지도학습 알고리즘 예제 - K - 평균)



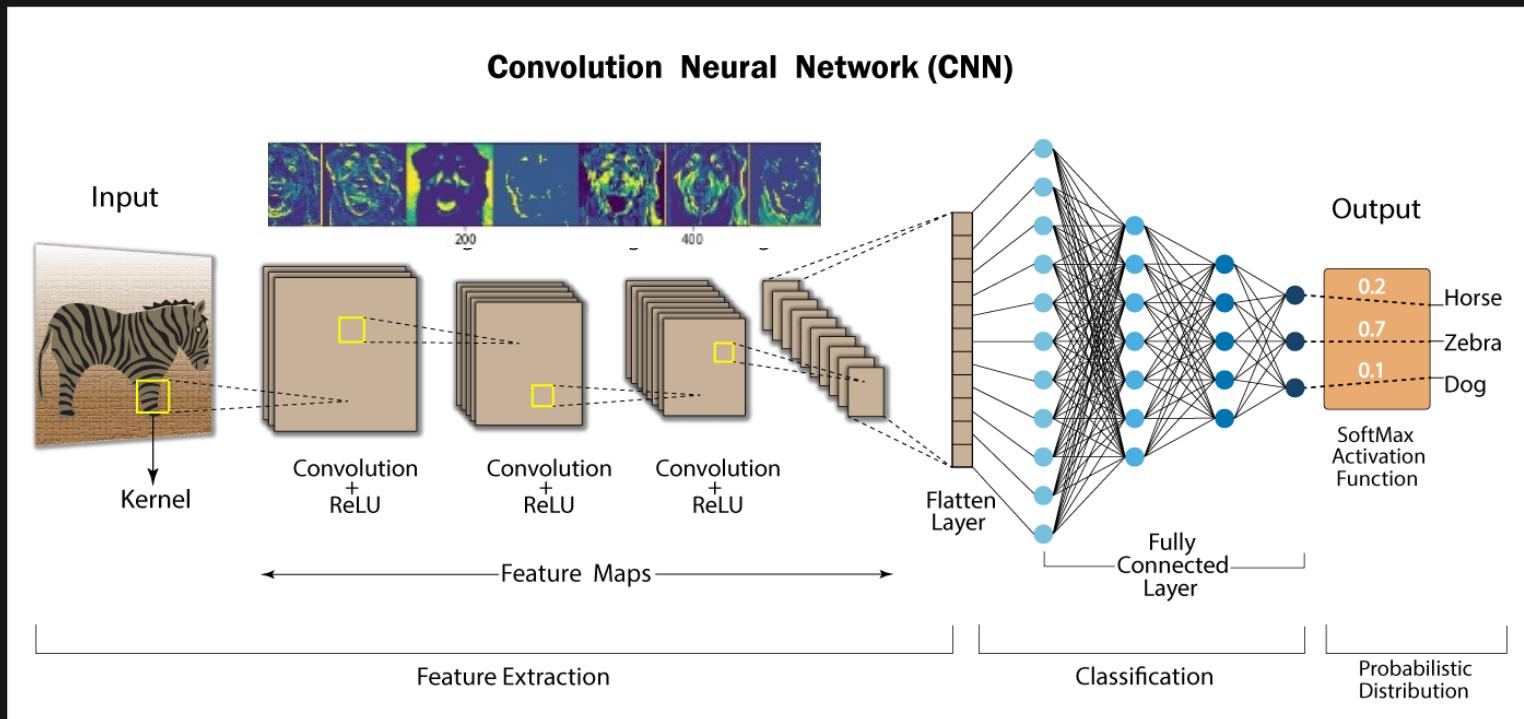
K = 3, K 값에 따라 몇 개의 군집을 구성 하는지 지정하게 됩니다.



K – Means
(Unsupervised)

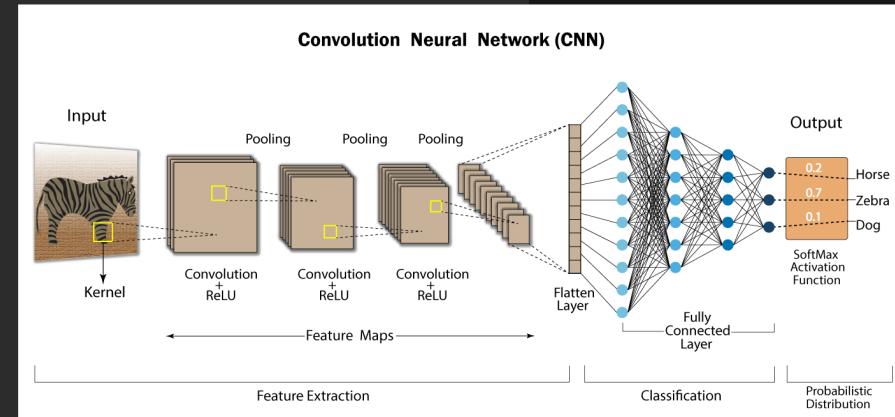


기계학습 알고리즘 (딥러닝 - 콘볼루션 신경망)



기계학습 알고리즘 (딥러닝 - 콘볼루션 신경망)

```
model=tf.keras.models.Sequential([  
  
    tf.keras.layers.Conv2D(8,(3,3),activation ='relu', input_shape=(150,150,3)),  
    tf.keras.layers.MaxPooling2D(2,2),  
  
    tf.keras.layers.Conv2D(16,(3,3),activation ='relu'),  
    tf.keras.layers.MaxPooling2D(2,2),  
  
    tf.keras.layers.Conv2D(32,(3,3),activation ='relu'),  
    tf.keras.layers.MaxPooling2D(2,2),  
  
    tf.keras.layers.Flatten(),  
  
    tf.keras.layers.Dense(1024,activation='relu'),  
    tf.keras.layers.Dense(512,activation='relu'),  
  
    tf.keras.layers.Dense(3,activation='softmax')  
])
```



기계학습 알고리즘 (이전 예제: 지도 학습 흐름)

데이터 준비

학습 데이터(x)	클래스(y)
0, 0	0.5
2, 2	2.5

알고리즘/모델 기반 선택

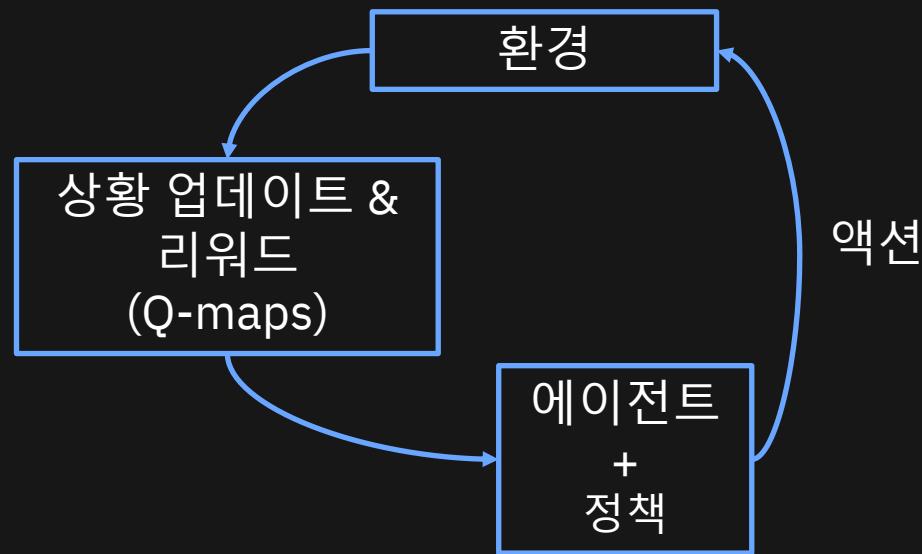
학습

테스트

```
>>> from sklearn import svm
>>> X = [[0, 0], [2, 2]]
>>> y = [0.5, 2.5]
>>> regr = 텐서 model
>>> regr.fit(X, y)
SVR()
>>> regr.predict([[1, 1]])
```



기계학습 알고리즘 (강화형 기계학습)

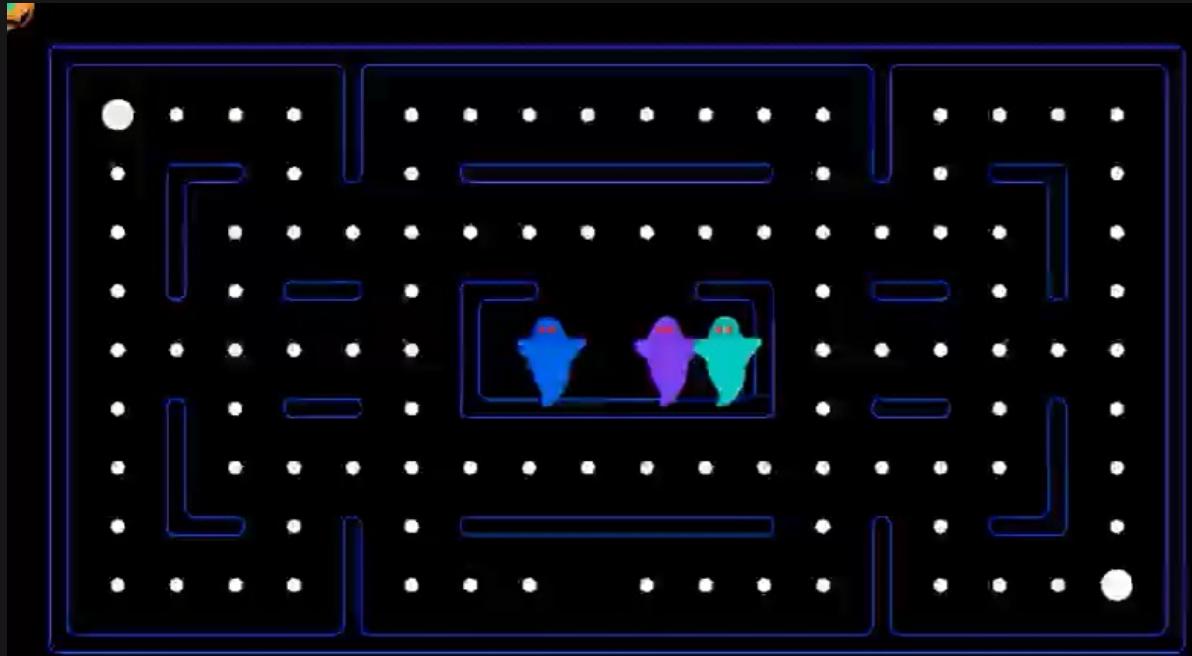


Jeopardy에서 전략적 결정을 내리기 위해
강화 학습 구현 및 적용
(IBM의 Watson 2011)



기계학습 알고리즘 (강화형 기계학습)

1. 역 강화 학습은 에이전트가 어떻게 작동할지 경로를 보여줍니다.
2. 강화 학습을 통해 에이전트는 점수를 극대화하는 방법을 배울 수 있습니다.
3. 각 상황마다 에이전트가 이 두개의 정책을 복잡한 방식으로 결합하면서 각 단계에서 어떤 행동을 해야하는지 결정되게 됩니다.



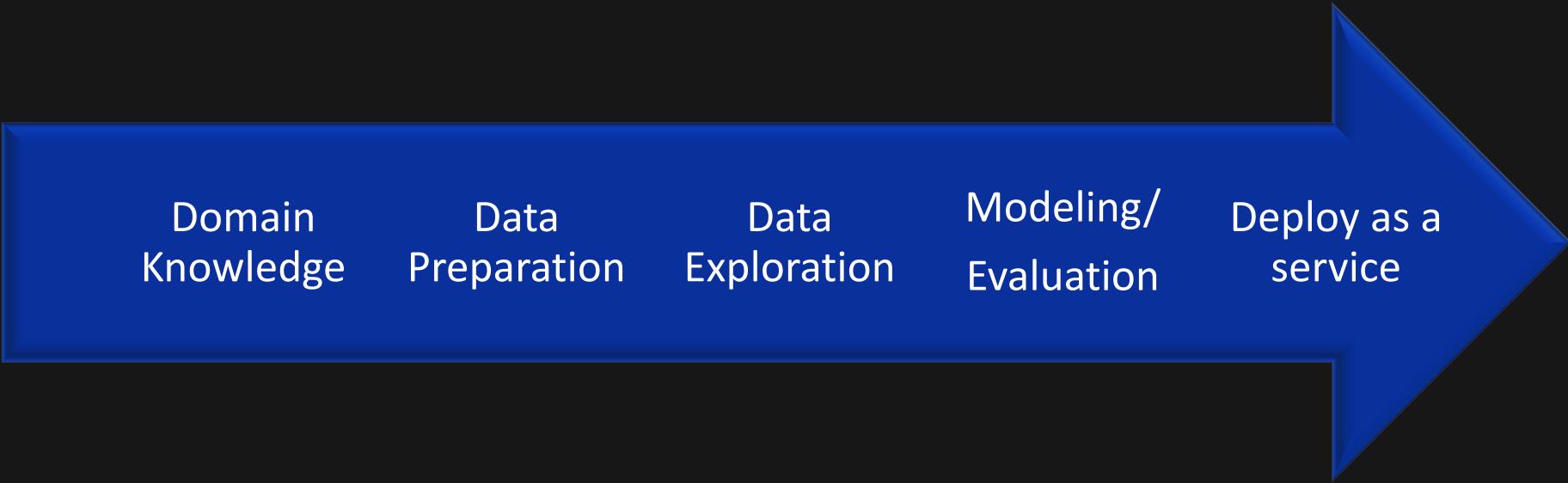
기계학습 알고리즘 (딥러닝 vs. 강화형 기계학습)



- 딥 러닝: 훈련 세트에서 학습한 다음 해당 학습을 새로운 데이터 세트에 적용
- 강화 학습: 보상을 최대화하기 위해 지속적인 피드백을 기반으로 조치를 조정하여 동적으로 학습합니다.



기계학습 알고리즘 (기본적인 AI 어플리케이션 절차)



Domain Knowledge Data Preparation Data Exploration Modeling/
Evaluation Deploy as a service



특정 분야 지식

+

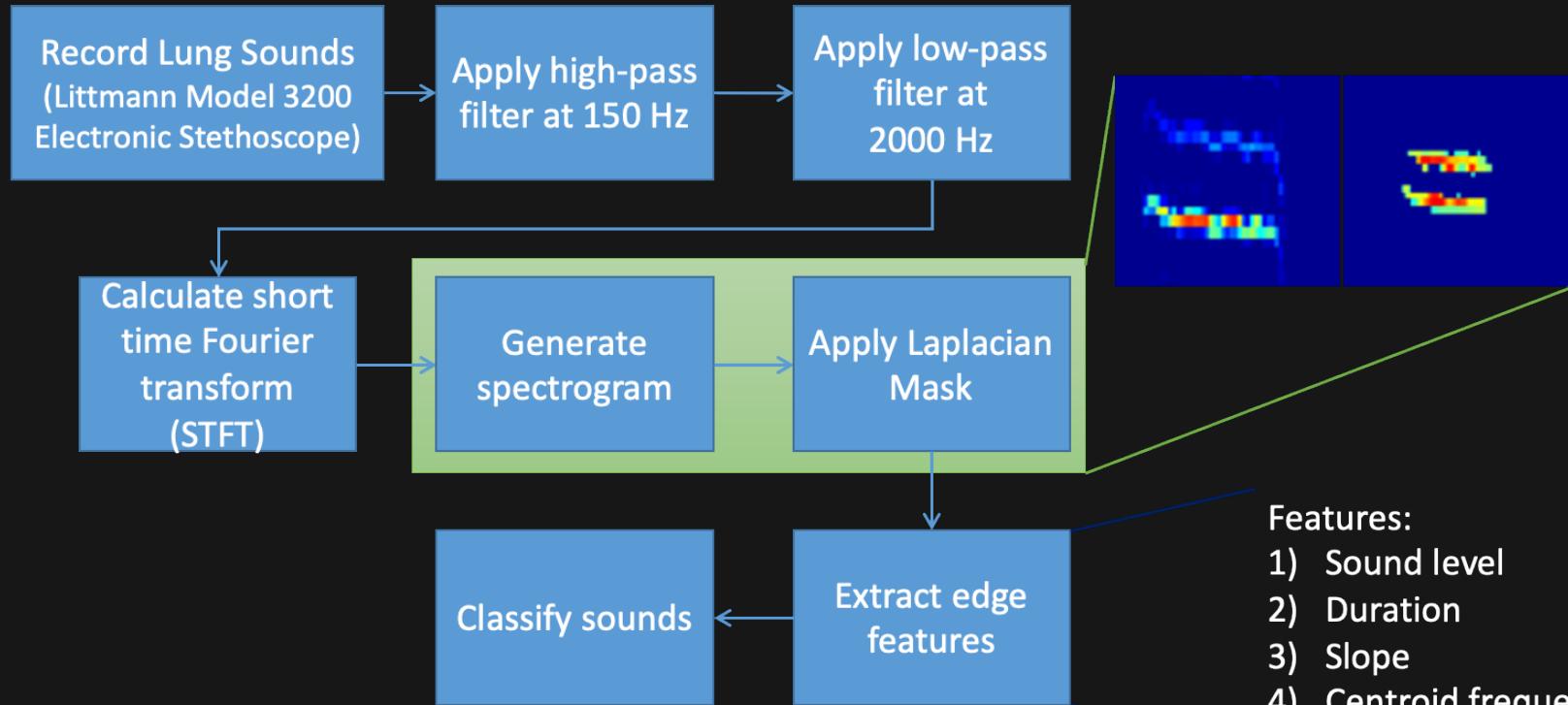
기계학습

= 인공지능



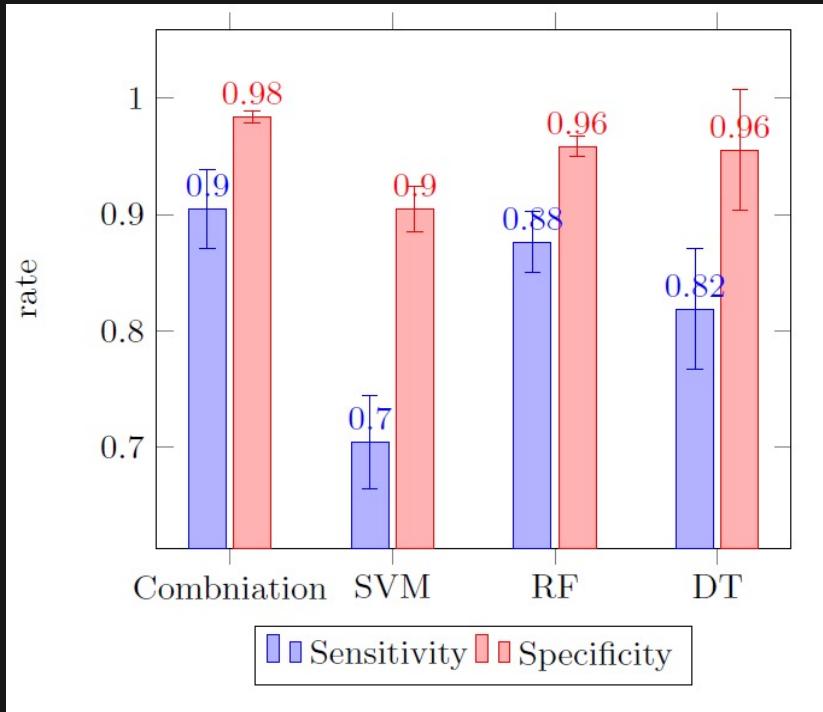
천식환자 모니터링

이상 소리 감지



천식환자 모니터링

소리 판별기 성능 평가

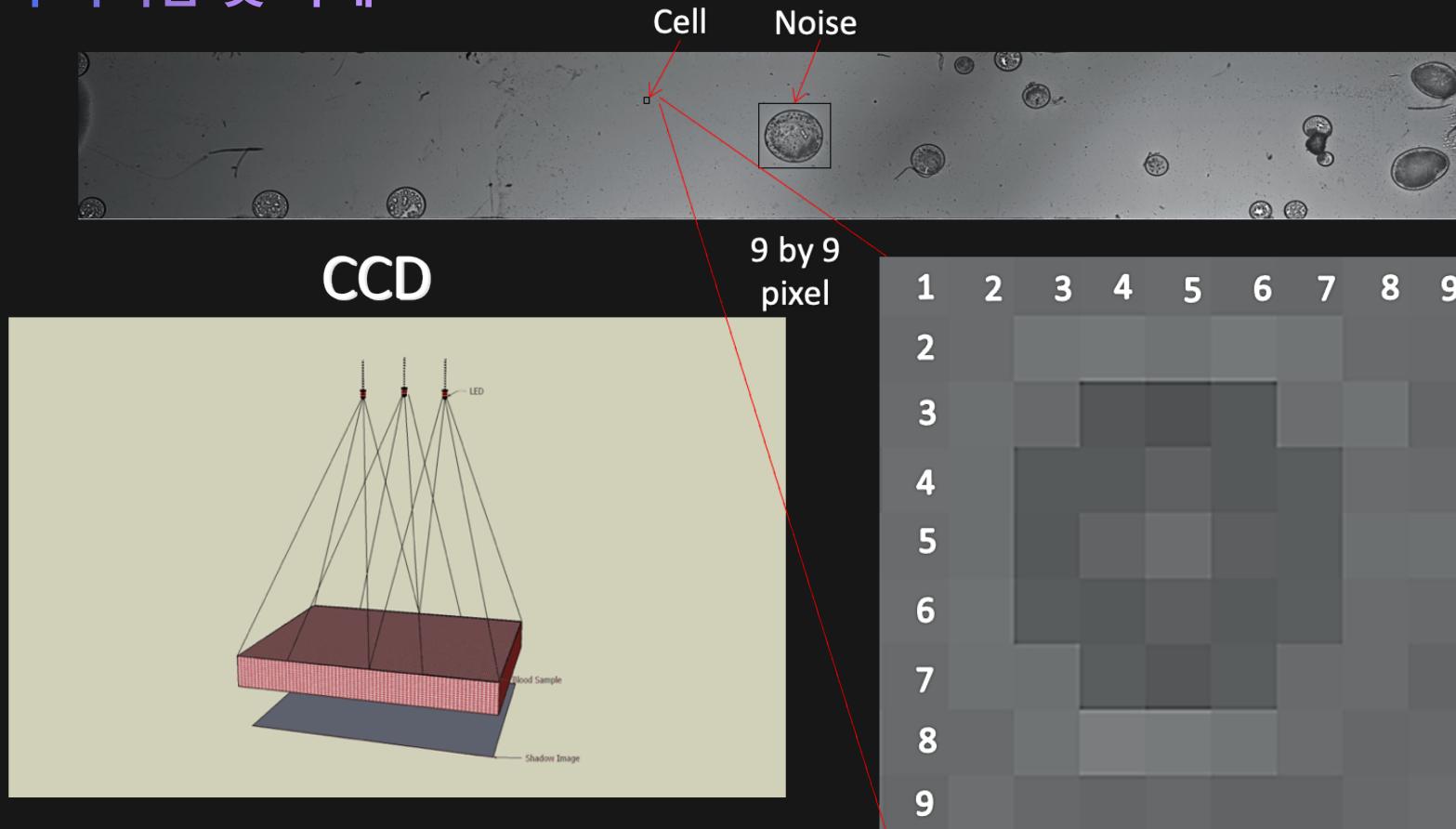


Classifier	Sensitivity (Wheezing Accuracy)	Specificity (Normal Accuracy)	Overall Accuracy
SVM	58/63 (92%)	15/21 (71%)	73/84 (87%)
Decision Tree (C4.5)	58/63 (90%)	18/21 (62%)	71/84 (85%)
Random Forest Classifier	60/63 (95%)	18/21 (86%)	78/84 (93%)
Combination	62/63 (94%)	19/21 (95%)	81/84 (96%)



세포계수 알고리즘

기술적 어려움 및 과제



스마트워치 심박수 정확도 측정 알고리즘

필터 디자인

1. Set of differences among consecutive PPG light intensity readings:

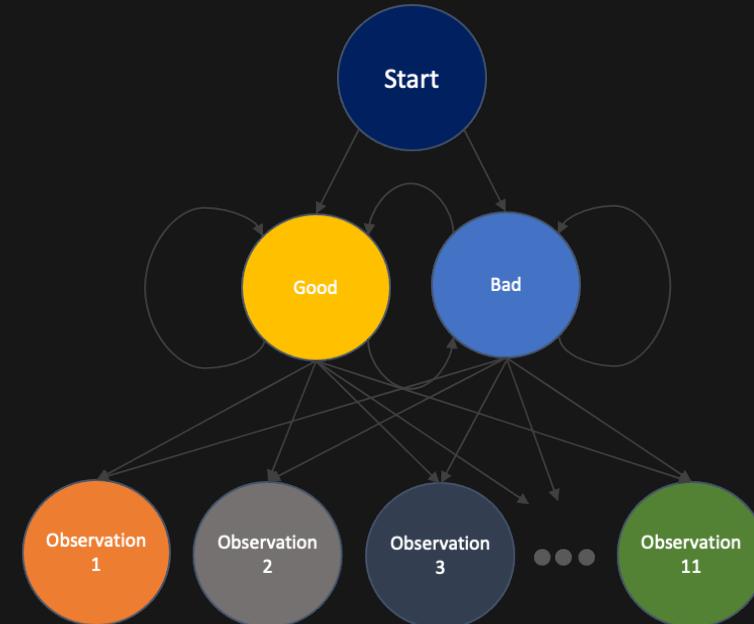
$$\Delta L_{(t:t+w)}$$

2. Identify light intensity step size between observations:

$$S_{step} = \frac{\max(|\Delta L_{(t:t+w)}|)}{N_O}$$

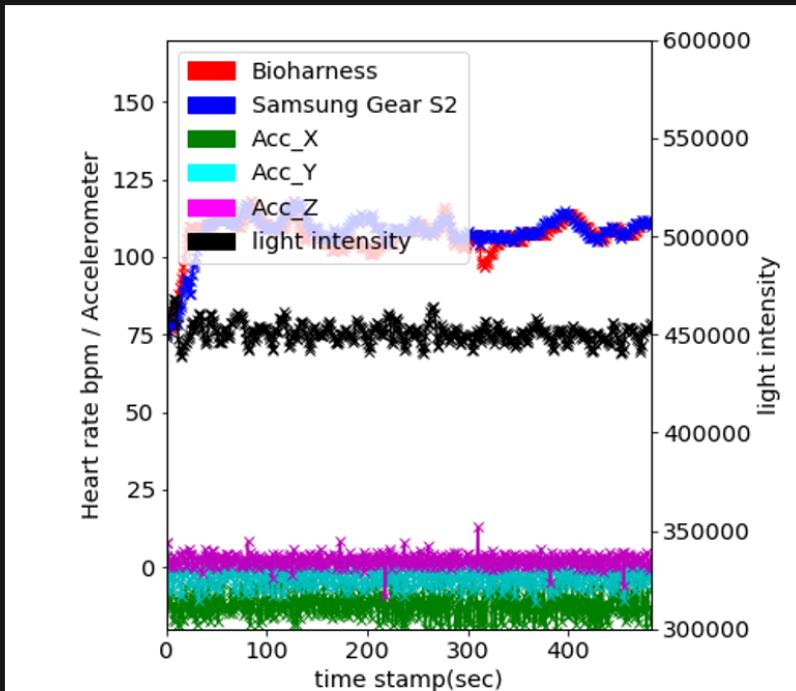
3. Compute observation for each time window:

$$O_n = \lfloor \frac{|\Delta L_{(t:t+w)}|}{S_{step}} \rfloor$$

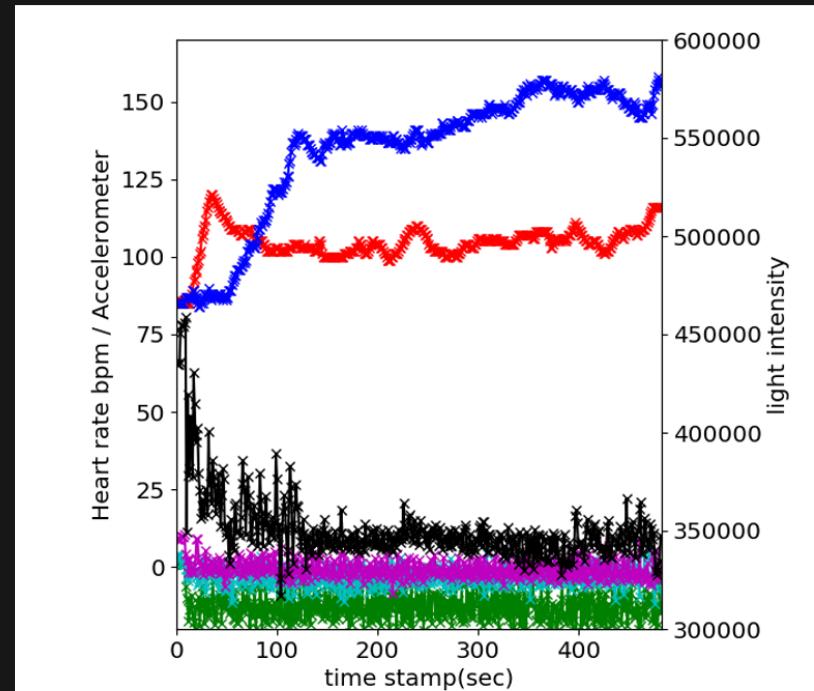


스마트워치 심박수 정확도 측정 알고리즘

Smartwatch Worn Tightly vs. Loosely



Samsung Gear S2 worn tightly



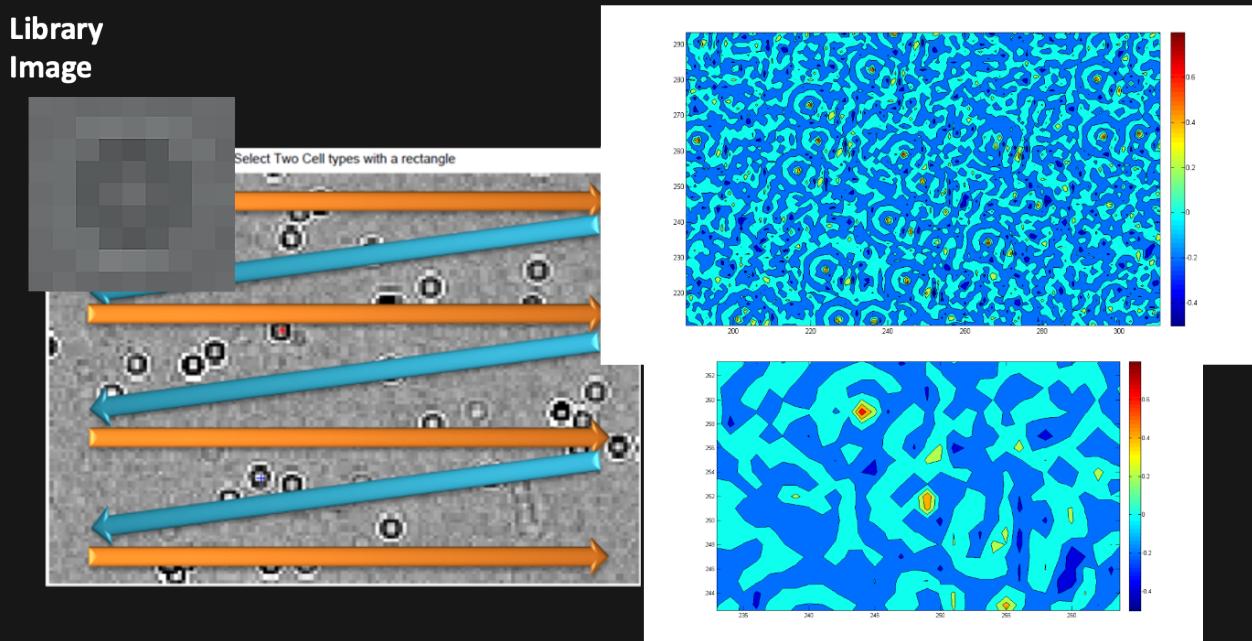
Samsung Gear S2 worn loosely



세포계수 알고리즘

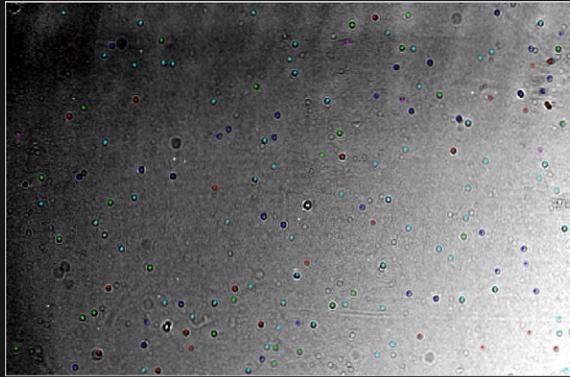
2D transform domain convolution

Similar with Green function in frequency domain and PDE solution: substitute unit delta function as matching image with Euclidian distance

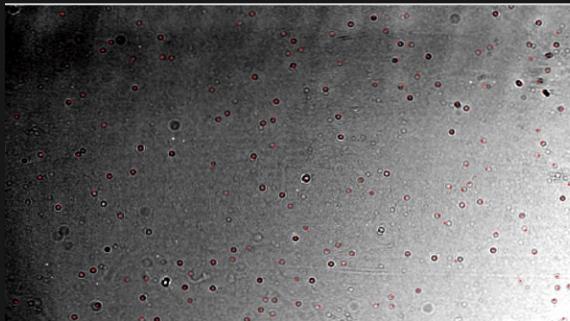


세포계수 알고리즘 (Automatic Cell-Counting)

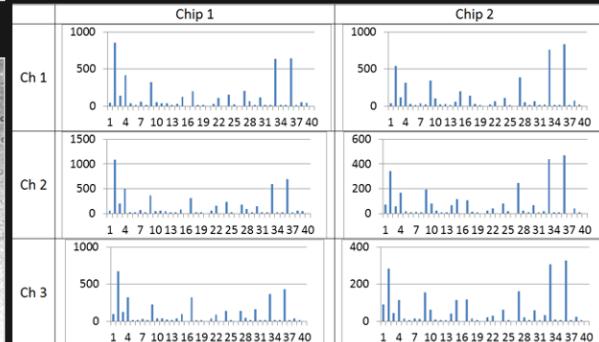
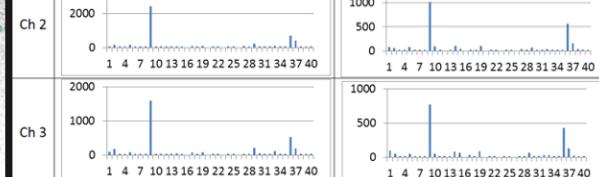
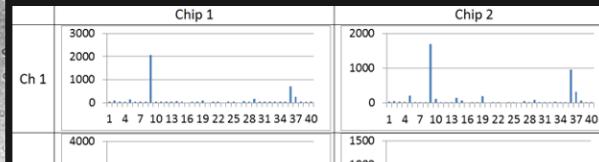
알고리즘 평가



Automatic 155



Manual 157



Biased cell type counts
10 of the most counted
cell types are removed
to avoid biased cell type
counts



IBM 왓슨

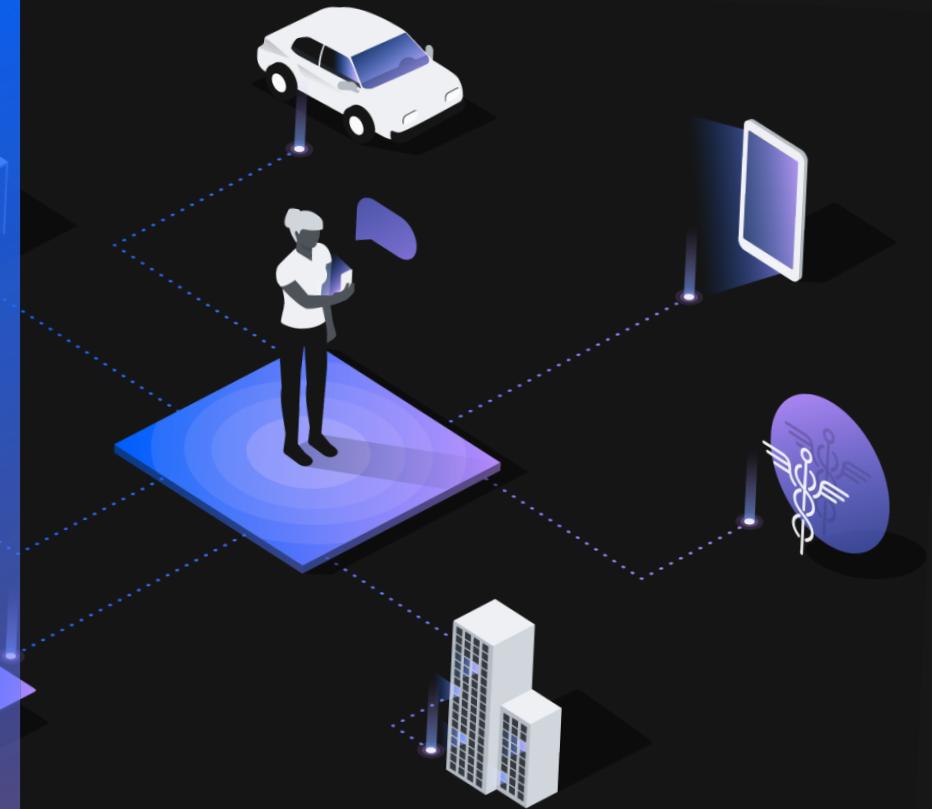
&

인공지능



IBM Watson 기능

- Assistant
- Discovery
- Natural Language Understanding
- Knowledge Studio
- Speech to Text
- Text to Speech
- Language Translator
- Watson Studio
- Watson Machine Learning



IBM Watson 어시스턴트

IBM Watson Assistant Lite Upgrade

Learning center



I want to pay for my credit bill

Customer starts with:
I want to withdraw money

Conversation steps

from what account?

1 checking accou savings accoun

↓ Continue to next step

1 = checking account

how much do you want to withdraw?

2 \$100 \$500 + 1

↓ Continue to next step

2 = \$100 , \$500

confirm? Step 2

3 yes no

↓ Continue to next step

New step +

Customer starts with

A phrase that expresses the action they want to complete.

With each example, your assistant learns when this is the right action for what a customer wants.

Enter phrases your customer might use to start this action

I want to withdraw money

give me money

Enter a phrase

Show less

Preview

Greet customer [default]

Welcome, how can I assist you?

Type something...



IBM Watson

자연어 이해(NLU) + 디스커버리

Project Debater

IBM has started the commercialization of IBM Research and Project Debater technologies to help identify, understand and analyze the most challenging aspects of human language and deliver deeper business insights. Features we plan to productize in Watson Discovery and Watson NLU include:

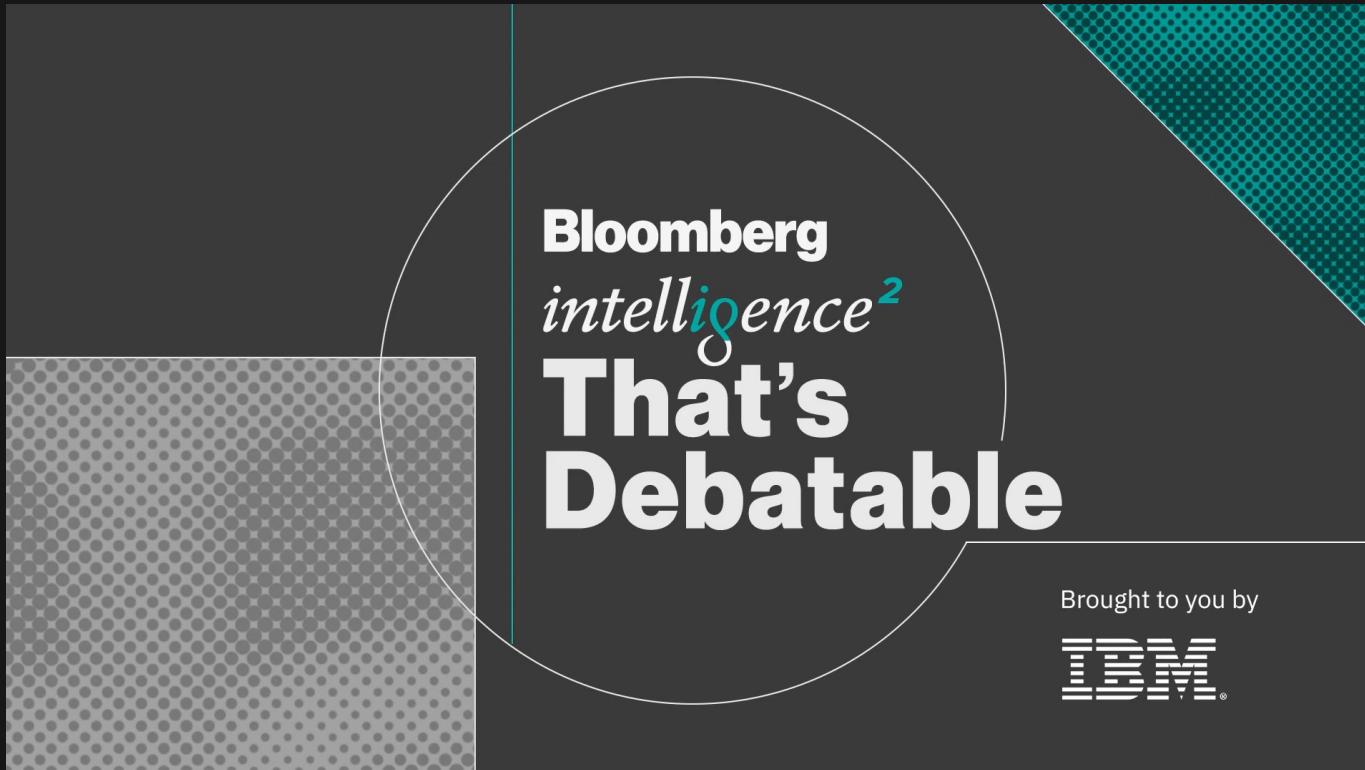
Summarization
Advanced Topic Clustering
Sentence Classification



The image shows two screenshots of IBM Research interfaces. The top screenshot displays a debate between 'We should subsidize preschools' and 'Opening Rebuttal Summary'. The bottom screenshot shows the 'IBM Science Summarizer' interface, which includes a search bar, a summary of '3D Shape Reconstruction from Free-Hand Sketches', and a topic modeling visualization. The visualization consists of five circular nodes connected by arrows, labeled 'Input Selection', 'Vocabulary Induction', 'Parameters Fine-Tuning', 'Topic Modeling', and 'Classify New Data'. Below this are several colored boxes representing different topics and their percentages: UNKNOWN (KING CLUSTER) (47%), sale_rep_change_client (29%), customer_subscription_select_product (8%), cloud_instance_admin_download (6%), job_small_big (3%), client_sustainable_business (3%), product_documentation_support_local (1%), dashboard_key_success_metric (1%), software_tool_refresh (0%), implementation_model_framework_cloudpaks (0%), modernize_internal_system_support (0%), attitude_support_ibmer_ecosystem (0%), client_perform_mutual (0%), flatten_organization_reduce_layer (0%), rely_process_limit (0%), and aim_high_productivity_metric (0%).

IBM Watson

자연어 이해(NLU) + 디스커버리



Brought to you by

IBM®



IBM Watson 음성인식(Speech to text) 화자 인식 기능 시연

The screenshot displays the 'IBM Watson Speech to Text Demo' interface. At the top left is a blue button labeled 'Stop audio'. To its right is a toggle switch labeled 'Recognize Speakers' with a greyed-out icon. Below these controls are three tabs: 'Base Model' (selected), 'With User Training' (disabled), and 'With Timings' (disabled). The main area shows two separate audio tracks. The left track, under 'Base Model', has the text 'Wohlfahrtstrasse...' and includes 'Alternatives' and 'Word Timings' buttons. The right track, also under 'Base Model', has the text 'Wohlfahrtstrasse...' and includes 'Alternatives' and 'Word Timings' buttons. Both tracks have a note below them stating 'Word alternatives will appear as the audio is playing.'



IBM Watson 음성합성(Text to speech) Data and AI 포럼 행사

IBM Watson 음성합성(Text to speech)
Data and AI 포럼 행사

실시간으로 한국어를 포함한 뉴스의 데이터를 수집하는 과정은 나호경 과장님이. 함께 트렌드를 알아보는 과정에는 인공지능 인재직업 교육의 일원인 뉴칼라스쿨 피테크 백승민 학생이, 맨토 맹윤호 쌤과 함께 해주셨습니다.

AI 사이버 : 쉽게 설명하는 AI 도전과제
CxO를 위한 최신 AI 활용가이드 : 규모에 맞게 성장 가속화
[알고리즘] 빅데이터 기반 비즈니스 가치 창출
[알고리즘] 뉴스 및 시장에 대한 디지털 혁신 가속화 및 AI 실현을 위한 새로운 여행

Watson | 함께 트렌드를 알아보는 과정에는 인공지능 인재직업교육의 일원인

나호경 과장
한국IBM Watson Developer PhD

백승민 학생
서울 뉴칼라스쿨
인공지능소프트웨어(2학년)

맹윤호 과장
한국IBM Technical Service Professional (Data & AI)

질문 입력 *
인증서 발급 조건
남은 재생 시간(분): 60

3 min(s)

04:00/08:37

PWRD BY ON24

인공지능(Artificial Intelligence) (IBM 웃슨 날씨 _weather)

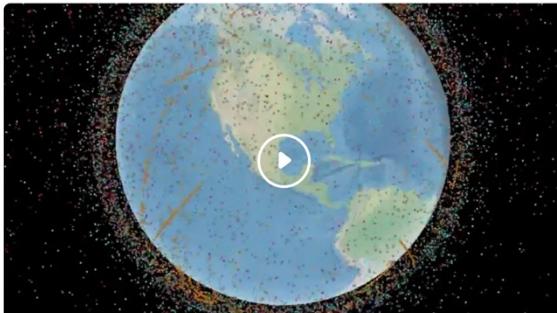
weather.com/?Goto=Redirected

The Weather Channel An IBM Business

66° Seoul, South Korea

Search City or Zip Code

Today Hourly 10 Day Weekend Monthly Radar

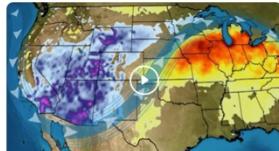


Nearly 200 'Super-Spreaders' Orbiting Earth

Threat of Severe Storms: Where to Watch



Temperature Roller Coaster Ahead This Week



Tourist Boom Likely at Site of History's Worst Nuclear Disaster

Deadly Double-Whammy for New Mexico's Trees



3:54 LTE

The Weather Channel An IBM Business

IBM

3:54 LTE

Yeongdeungpo-gu, Seoul

19° Mostly Cloudy Feels Like 19° Day 19° • Night 12°

19° 15° 19° 15° 13°

0% 70% 50%

Today's Details

Hourly Daily The Weather Channel Air Pollution Maps



인공지능(Artificial Intelligence) (IBM 웃슨 날씨 _weather)

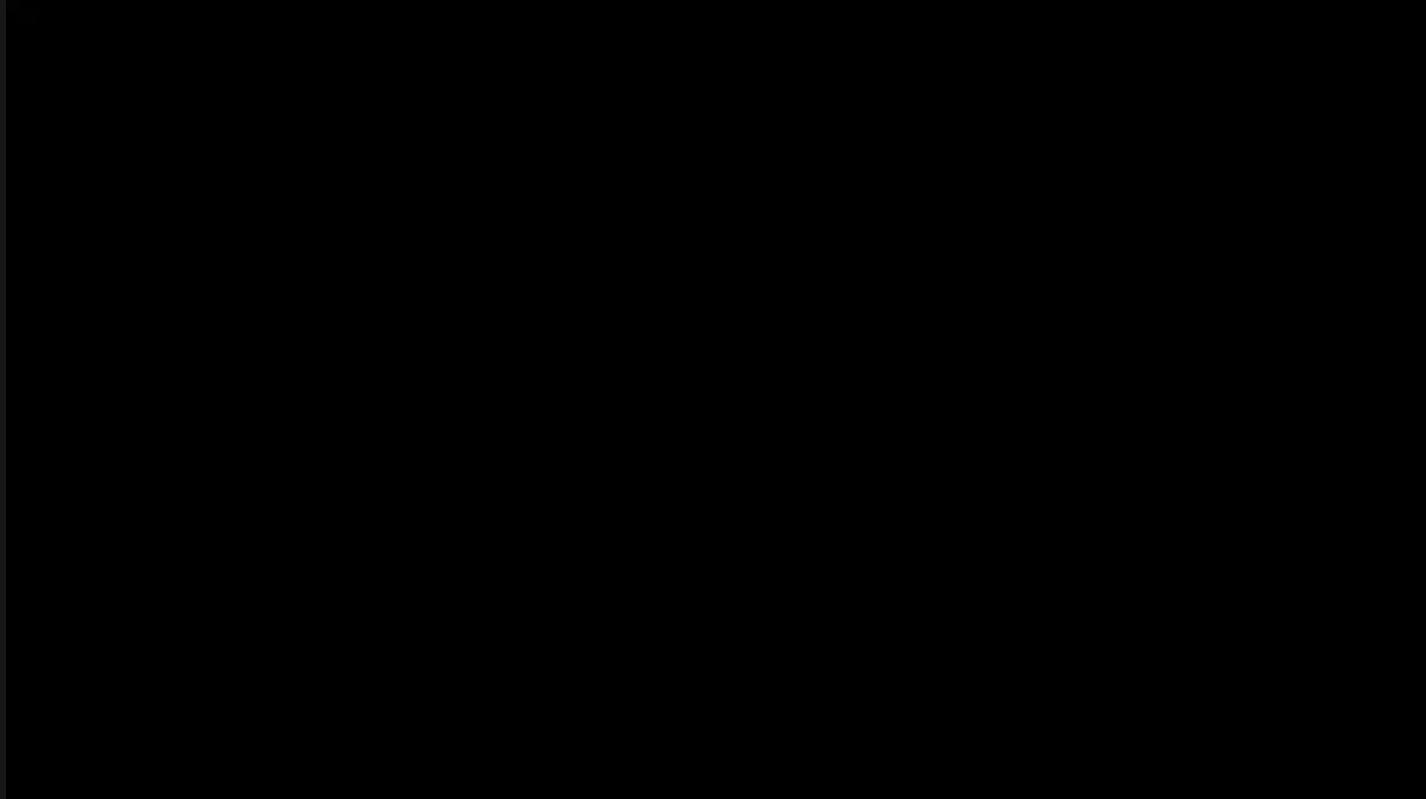


인공지능(Artificial Intelligence) (롤스로이스의 엔진 토탈 케어 서비스)

- 항공 산업은 지상 유지 보수 작업으로 인한 비행 지연으로 재정적 손실에 직면
- 클라우드와 분석을 활용하여 회사는 지능형 엔진을 만들고 비행 지연을 최소화
- 모든 장거리 비행은 수십 자리의 TB 데이터를 생성
- 빅 데이터 분석 후 해당 작업자에게 엔진 정보 제공



인공지능(Artificial Intelligence) (IBM Watson 항공기 관리 예제)



IBM Watson Studio

The screenshot shows the IBM Cloud Pak for Data interface. At the top, there is a navigation bar with the text "IBM Cloud Pak for Data", an "Upgrade" button, a bell icon, and a user account dropdown labeled "IBM - hk's account". Below the navigation bar, the page title is "Projects / hk_notebook_test_v1". The main content area has tabs for "Overview", "Assets" (which is currently selected), "Environments", "Jobs", "Access Control", and "Settings". A search bar at the top of the content area asks "What assets are you looking for?". Below the search bar, a section titled "Data assets" is expanded, showing a table with two entries:

<input type="checkbox"/>	Name ↑	Type	Created by	Last modified ↓
<input type="checkbox"/>	CSV bigml_59c28831336c6604c800002a.csv	Data Asset	HoKyeong Ra	Oct 04, 2020, 02:56 PM
<input type="checkbox"/>	hkttestbucketConnection	Connection	HoKyeong Ra	Oct 04, 2020, 02:55 PM

At the bottom right of the interface, there is a blue circular icon with a white speech bubble symbol.



IBM Watson Studio (Auto AI)

The screenshot shows the IBM Watson Studio (Auto AI) interface. At the top, there's a dark header bar with the "IBM Cloud Pak for Data" logo, an "Upgrade" button, a search icon, a notification bell, the user account "IBM - hk's account", and a profile icon labeled "HR".

The main area features a "Welcome, Hokyong!" message and a breadcrumb trail: Watson Studio • Watson Knowledge Catalog • Watson Machine Learning.

Three main sections are highlighted:

- Learn by example**: Step through solving a specific business problem in a sample project.
- Work with data**: Create a project for your team to prepare data, find insights, or build models.
- Extend your capabilities**: Add tools, databases, or other features by creating services instances.

Below these sections are buttons for "Create a project" and "Create a service".

The left sidebar includes "Quick navigation" with links to "Projects", "Catalogs", and "Deployment spaces".

The "Overview" section displays "Recent projects" (HR Doc_Test, Oct 12, 2020, 01:13 PM), "Recent catalogs" (No catalogs), and "Notifications" (Project import complete).



질문주세요!

Questions!?



인공지능!

Action!

Hands-on



