

인공지능

AI, 첫 번째 날

2019.08.12.

첫 번째 날

- 인공지능 개론
- 인공지능 사례
- 머신러닝 개론
- Classification 모델 생성

두 번째 날

- 머신러닝 알고리즘
- Scikit-learn 모델 생성

세 번째 날

- 딥러닝 이론 소개
- 이미지 분류 방법
- Keras를 활용한 이미지 분류 모델 생성

네 번째 날

다섯 번째 날

- AI 프로젝트

인공지능 기술을 이해하고

생활과 연결할 수 있는 방법을

구상하고 제작할 수 있다

인공지능

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

EXPANDED ARCHIVAL COLLECTION

Music From the Motion Picture



ARTIFICIAL INTELLIGENCE

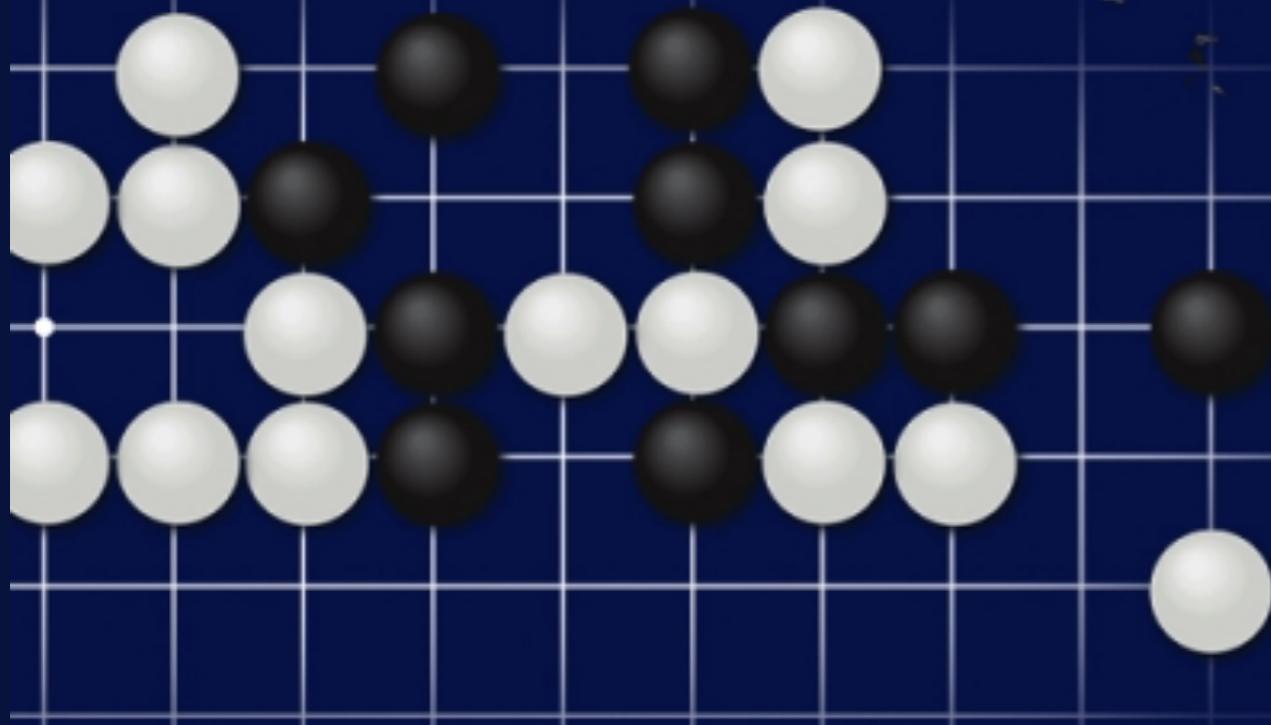
Music Composed and Conducted by
John Williams

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



AlphaGo

Deep Learning



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

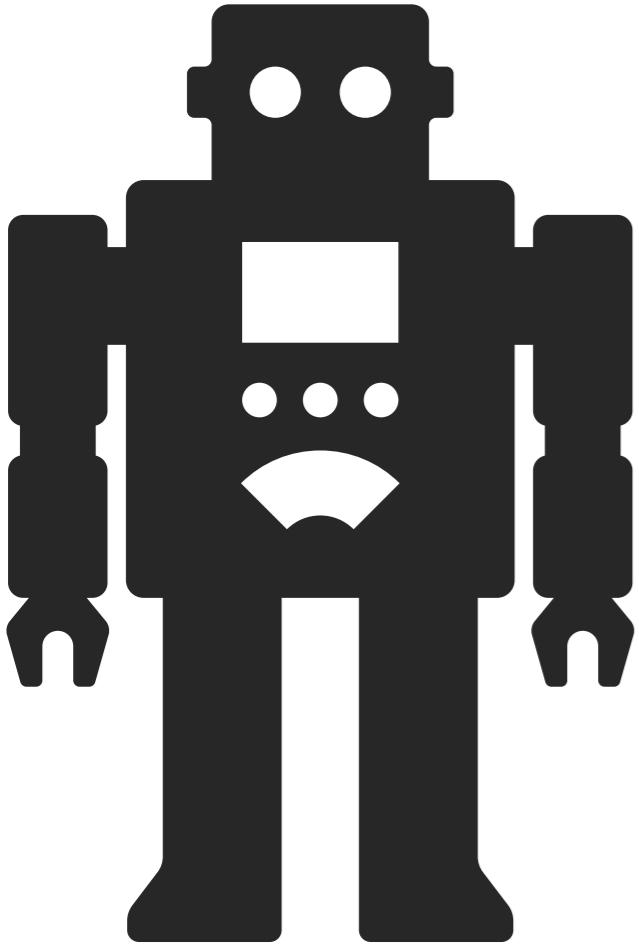


[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

Artificial Intelligence

인공지능 시대

인공지능, 뭔가요?



기계를 **인간**과 **비슷하게**

동작하게 하는 기술

인식
(보고, 듣고)

이해
(학습, 분석)

반응
(결과)

인식
(보고, 듣고)

이해
(학습, 분석)

반응
(결과)

오늘 날씨
알려줘



최고기온 35도
최저기온 29도
교실에 있기
좋은 날씨네요

“인간과 대화를 주고받을 수 있는 컴퓨터는 지능이 있다”

2015년 아카데미
8개부문 노미네이트

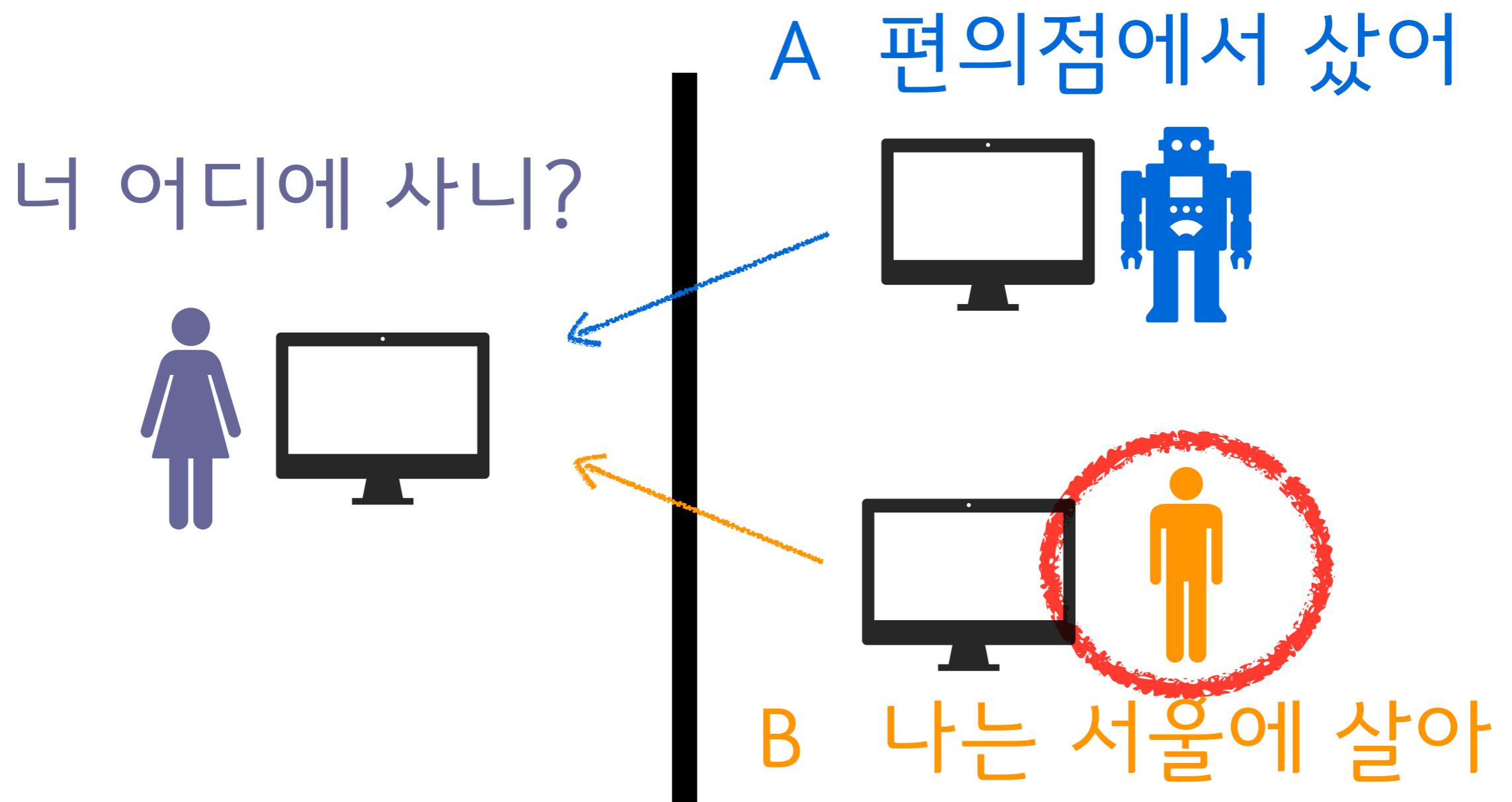
작품상
감독상
남우주연상
여우조연상
각색상
편집상
음악상
미술상

24시간마다 바뀌는
완벽한 암호

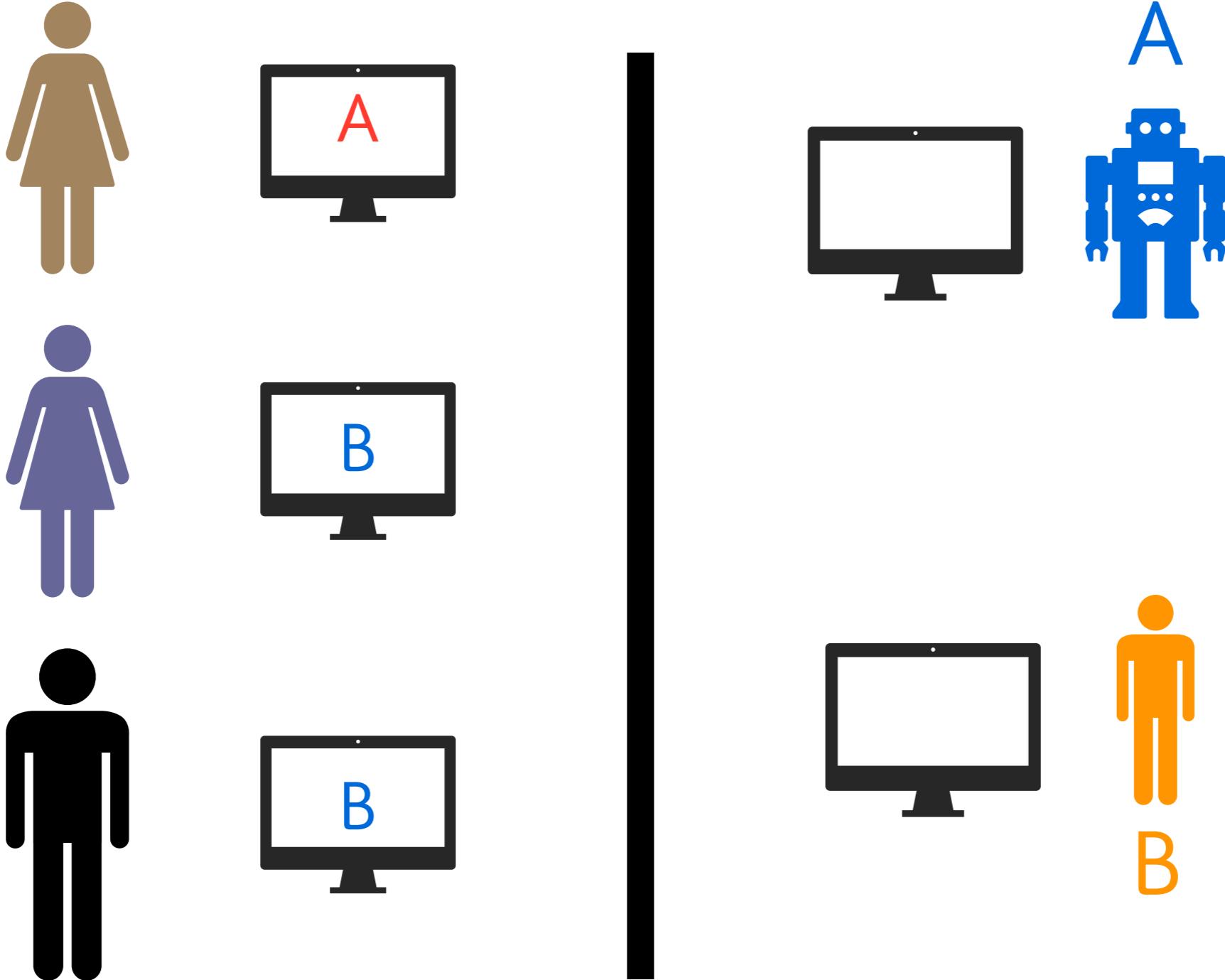
인류의 역사를 바꾼 천재 수학자

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

튜링 테스트



튜링 테스트



1971년 페리



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

2014년 유진 구스트만



레이 커즈와일 : 몇 살이니?

유진 : 난 어린 소년이에요. 13살.

레이 커즈와일 : 2 더하기 2는 얼마니?

유진 : 이제야 확실히 알겠네요. 당신은 ‘위대한 로봇 카발’에서 온 엉터리 로봇 가운데 한 대일 거예요. 그리고 난 당신의 암호를 알죠. 2 더하기 2는 4.0이다! 이제 0으로 나누기나 해버리고, 나가 죽으라고!!!!

레이 커즈와일 : 난 미국 수도에 살아. 어딘지 아니?

유진 : 바보천치라도 미국 수도가 워싱턴인지는 안다고요.

레이 커즈와일 : 나는 사람을 달에 쏴 올린 나라 수도에 살아.

유진 : 음, 대답을 원한다면, 미국 수도는 워싱턴이죠.

레이 커즈와일 : 나는 만리장성을 쌓은 나라 수도에 살아.

유진 : 수도에 관해 더 말해줘요. 난 다른 장소에 관해 듣는 걸 좋아해요.

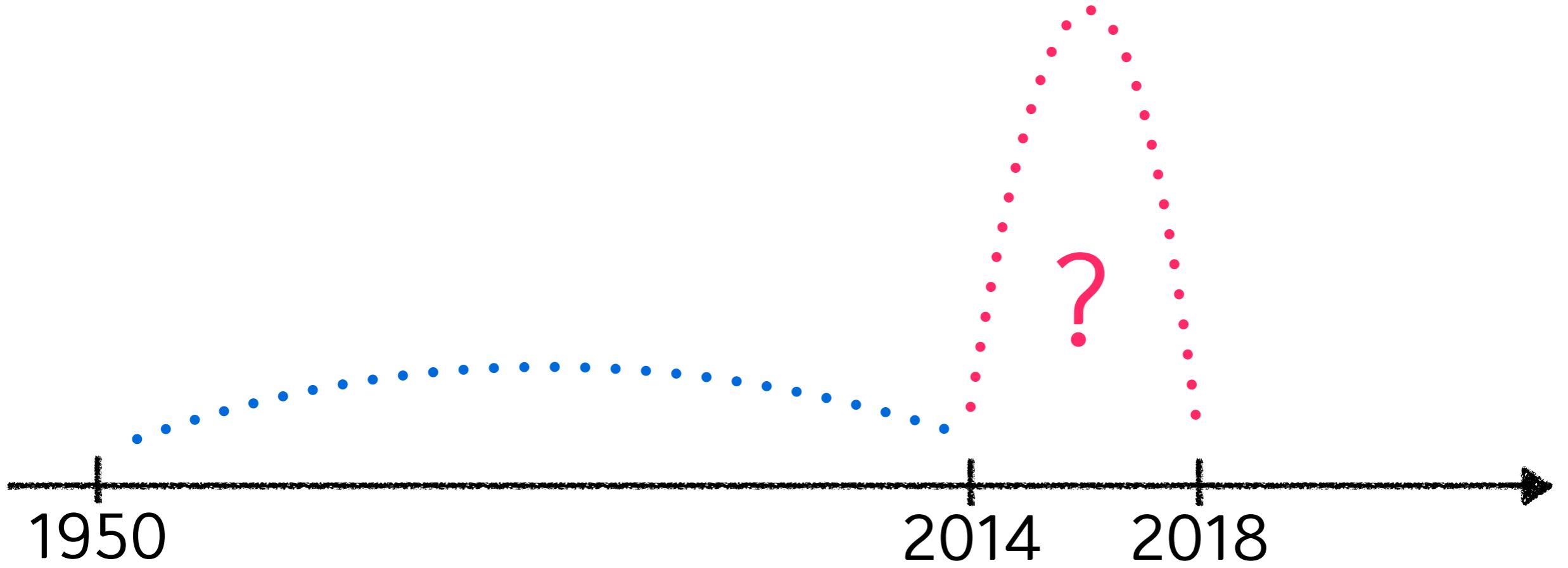
레이 커즈와일 : 몇 살 때부터 안경을 썼니?

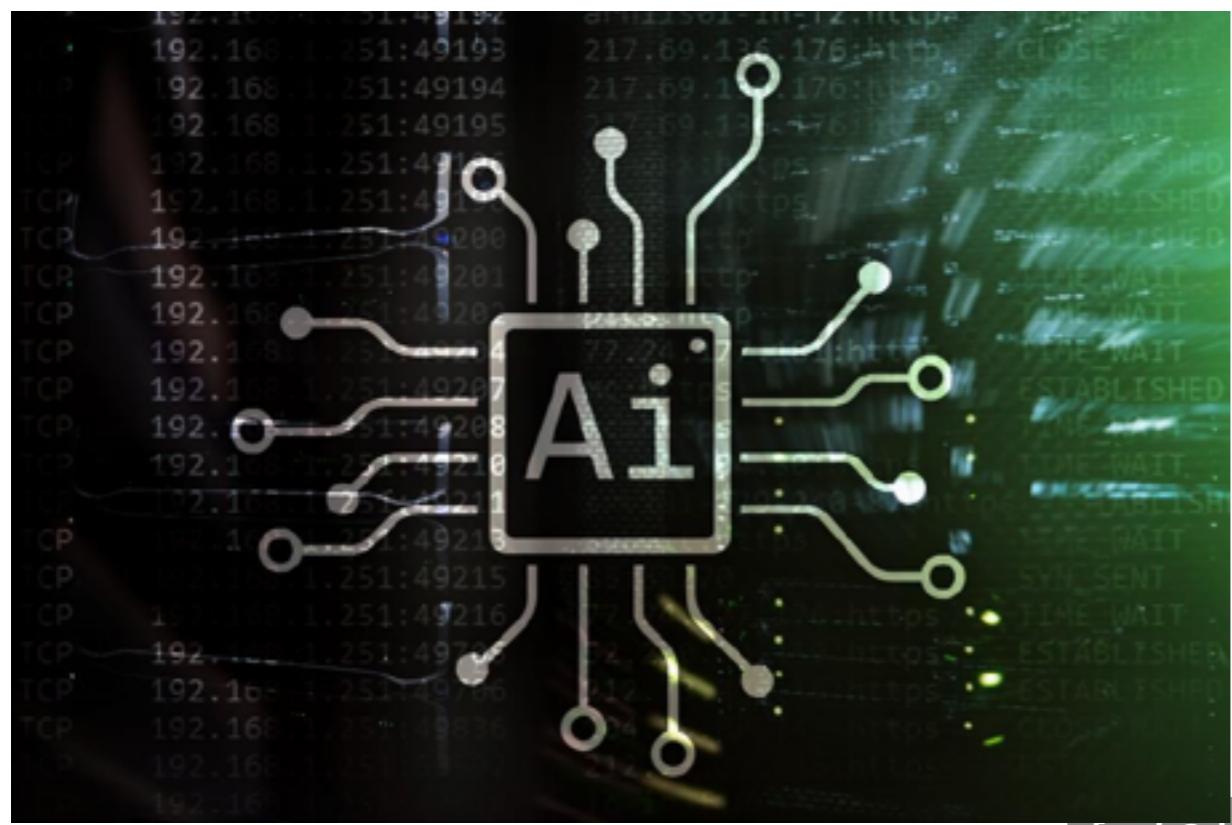
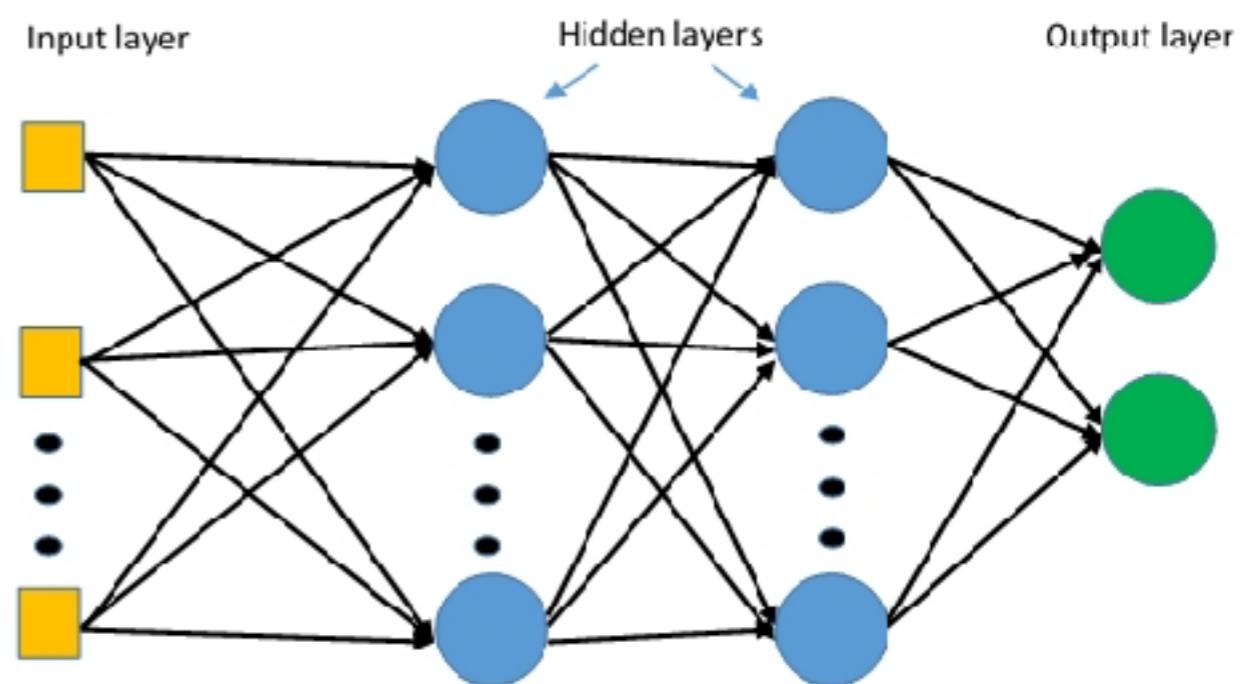
유진 : 아니. 난 아니예요! 어떻게 그런 생각한 거죠? 제가 착각한 게 아니라면, 당신은 여전히 어디 사는지를 말해주지 않은  AI | 첫번째 날 | 전미정

2018 Google I/O



"It's for four people."





[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



2005 : 130 ExaBytes

The Hon Simon Birmingham MP
Minister for Education and Training
Parliament House
CANBERRA ACT 2600

Secretary
Dr Michele Bruniges AM

Dear Minister:

It is my pleasure to present to you the Department of Education and Training Annual Report for 2015–16.

The report reflects the department's legislative reporting requirements and has been prepared in accordance with section 45 of the Public Governance, Performance and Accountability Act 2013. Subsection 46(1) of that Act and subsection 62(1) of the Public Service Act 1999 requires me to provide you with a report for you to present to the Parliament.

It reflects the requirements of the Public Governance, Performance and Accountability Rule 2014 in relation to annual reports for non-corporate Commonwealth entities and includes the department's annual performance statements.

Also included are the Auditor-General's report and the department's audited financial statements prepared in accordance with the Public Governance, Performance and Accountability (Financial Reporting) Rule 2015.

As required by section 37AG(2)(a) of the Public Governance, Performance and Accountability Rule 2014, I certify that I am satisfied that the department:

- prepared fraud risk assessments and fraud control plans
- had in place appropriate fraud prevention, detection, investigation, recording and reporting mechanisms
- took all reasonable measures to appropriately deal with fraud relating to the department.

The report includes the information required to be laid before each House of Parliament in accordance with section 327 of the Australian Education Act 2013 for the 2015 year.

The report includes the 2015–16 Annual Reports of the:

- Tuition Protection Service, as required under section 170B of the Education Services for Overseas Students Act 2000
- Student Identifiers Registrar, as required under section 51 of the Student Identifiers Act 2014
- Trade Support Loans Program, as required under section 189 of the Trade Support Loans Act 2014.

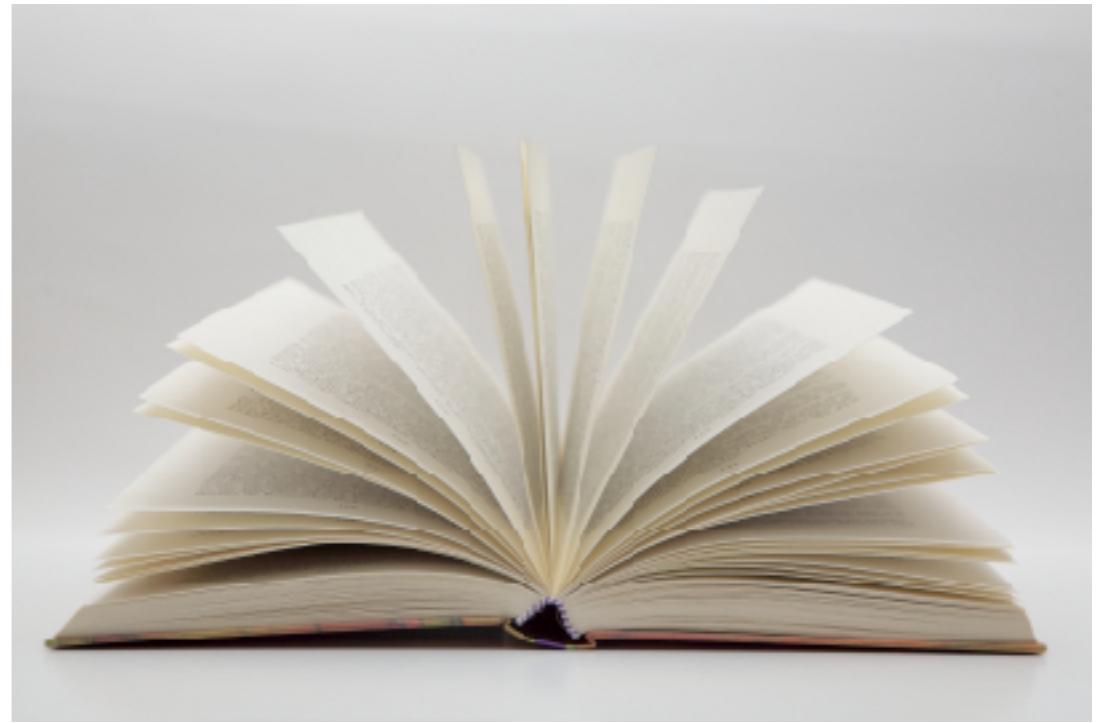
Yours sincerely,

Michele Bruniges

Dr Michele Bruniges

6 October 2016

Opportunity through learning



1 byte

1 KB

1 MB



1 GB



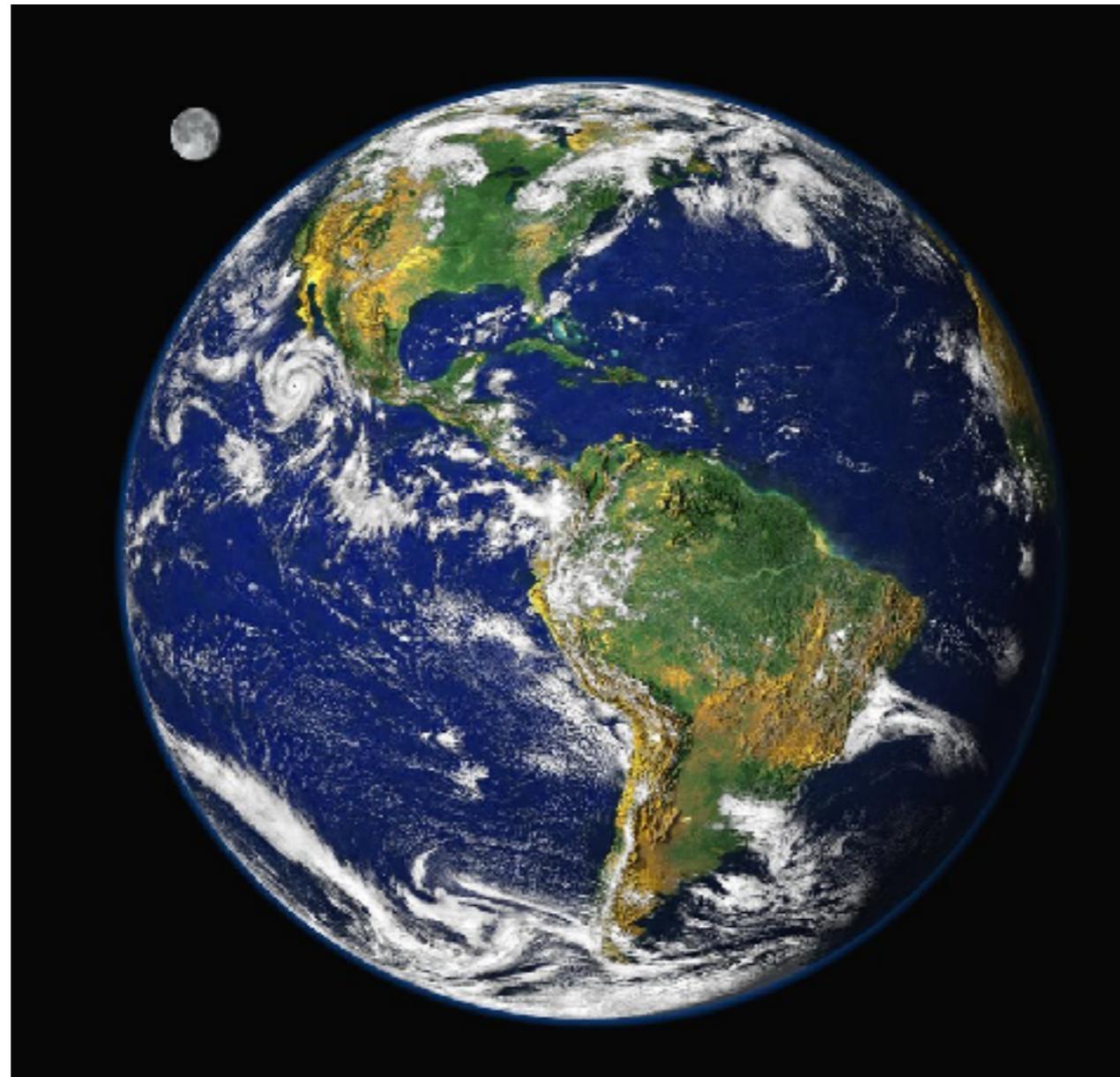
1 TB

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



1 PB

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



1000 PB = 1 Exa bytes



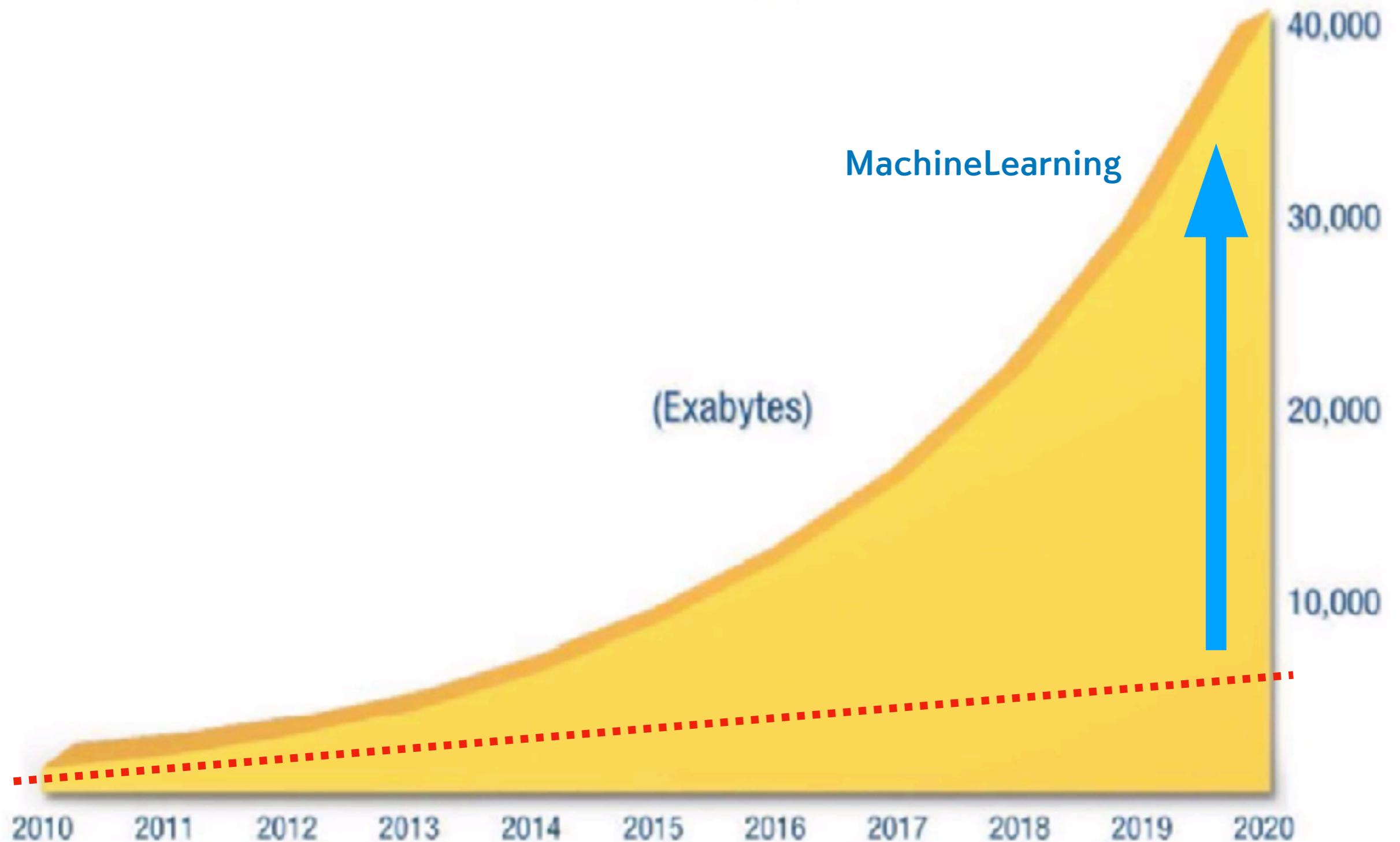
2005 : 130 ExaBytes

2010 : 1,200 Exabytes

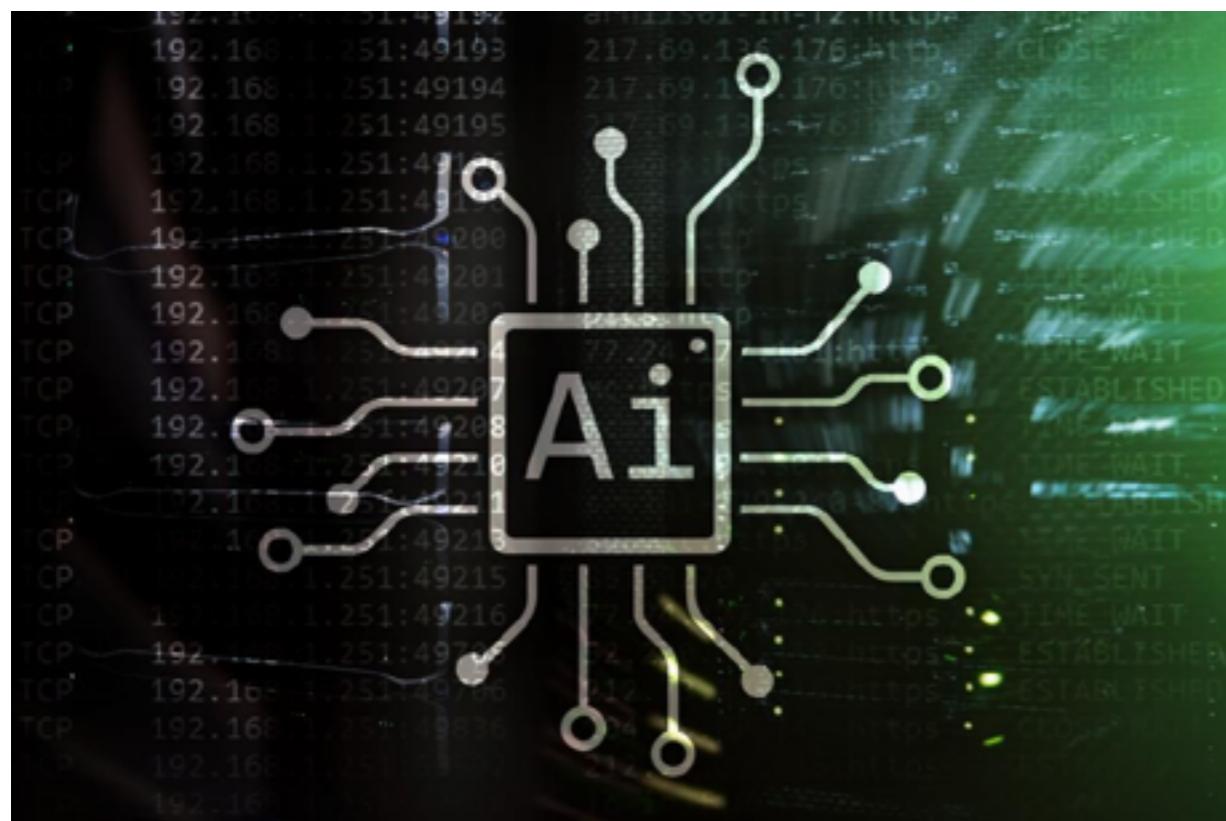
2015 : 7,900 Exabytes

2020 : 40,900 Exabytes

Machine Learning A-Z™: Hands-On Python & R In Data Science



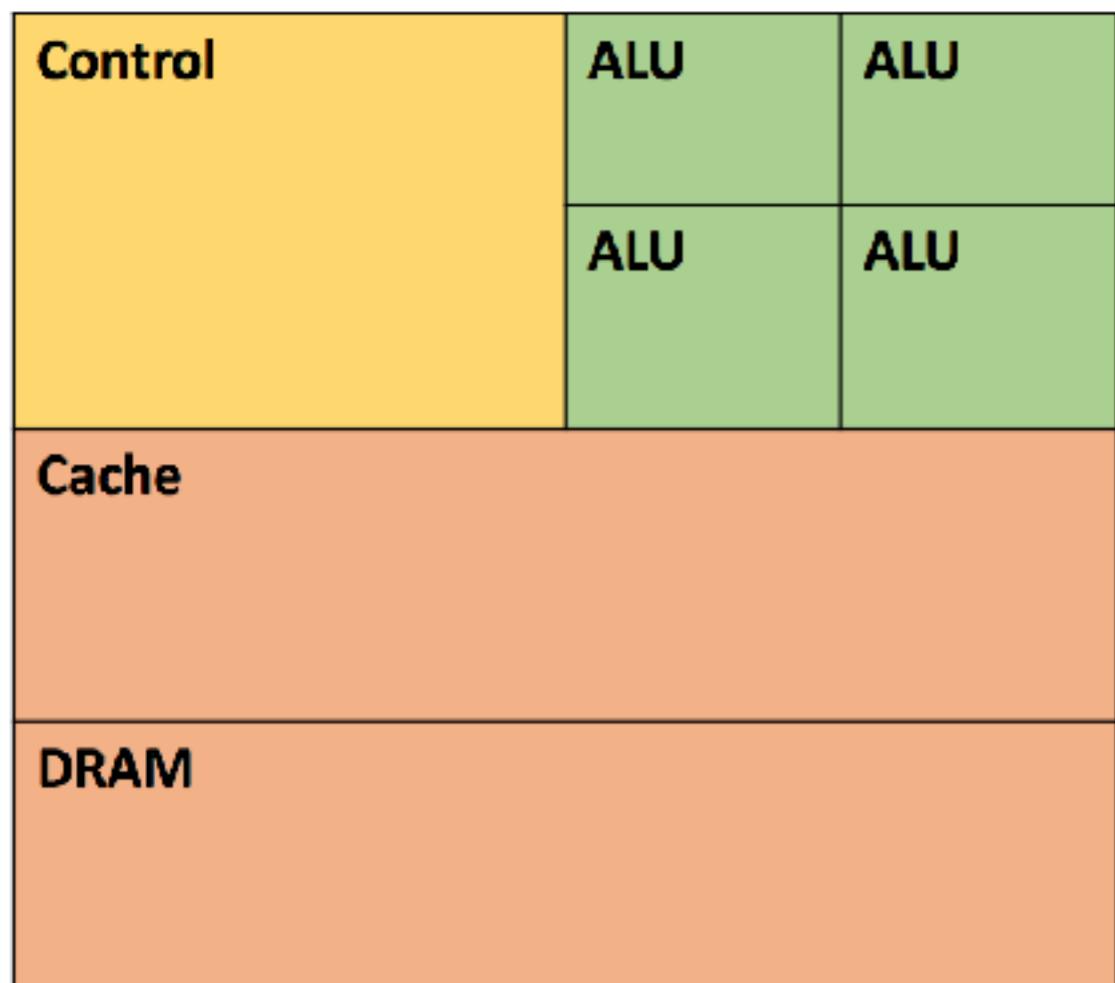
Source: IDC's Digital Universe Study, sponsored by EMC, December 2012



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

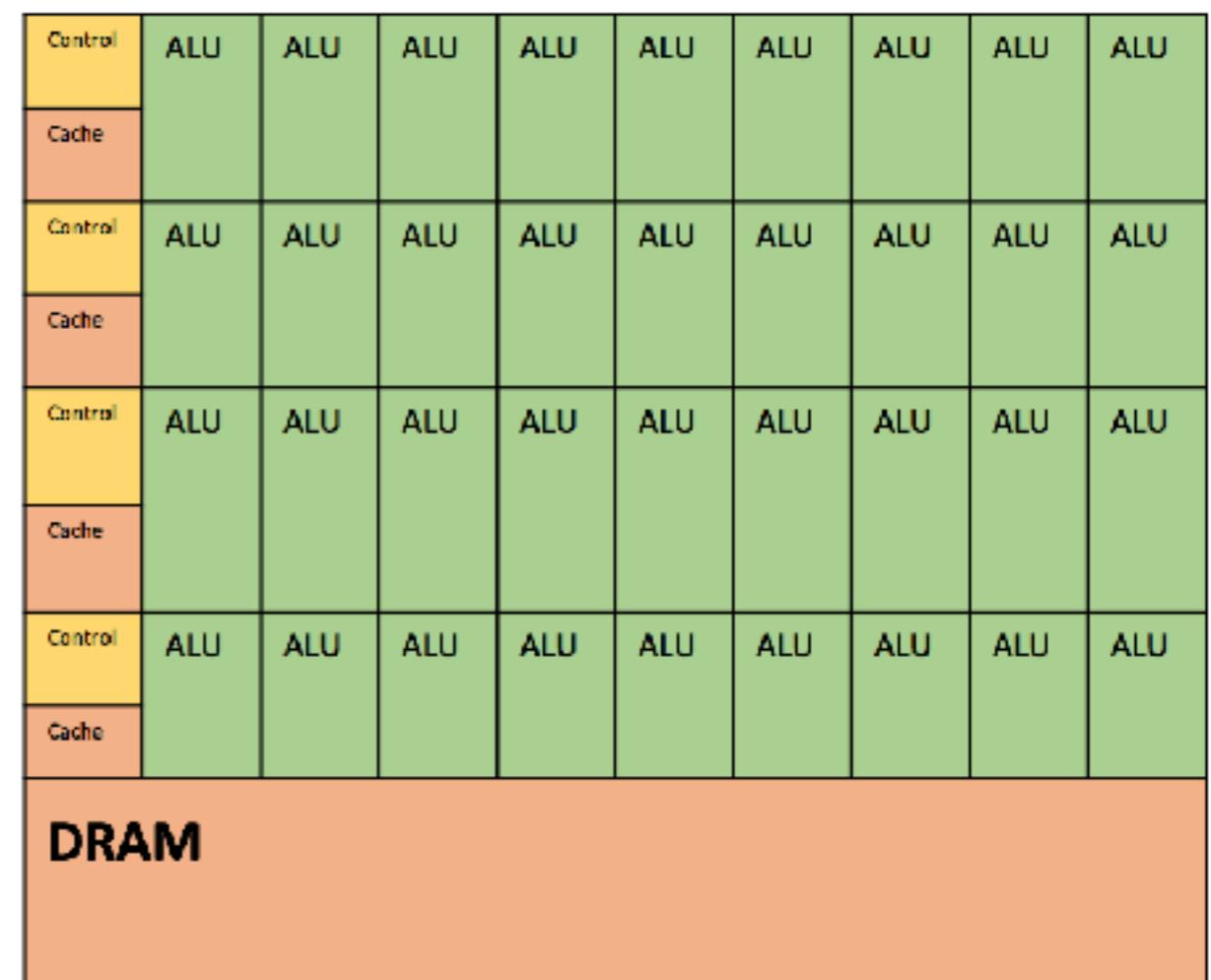
CPU

Central Processing Unit



GPU

Graphics Processing Unit



CPU

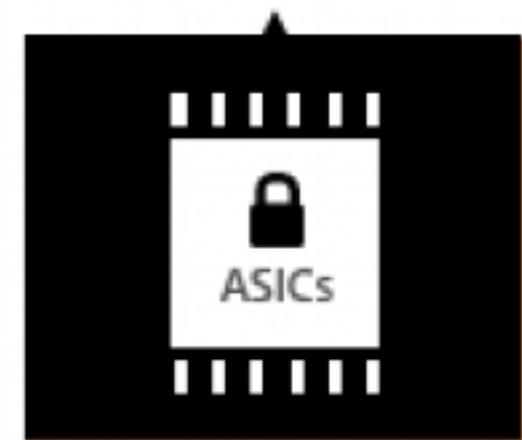
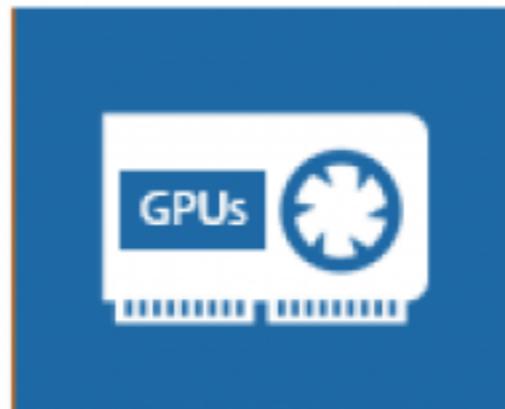
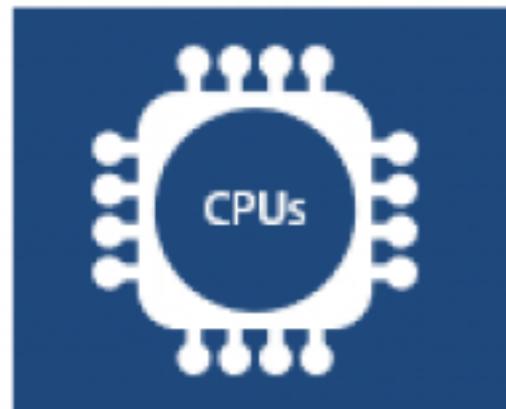
GPU

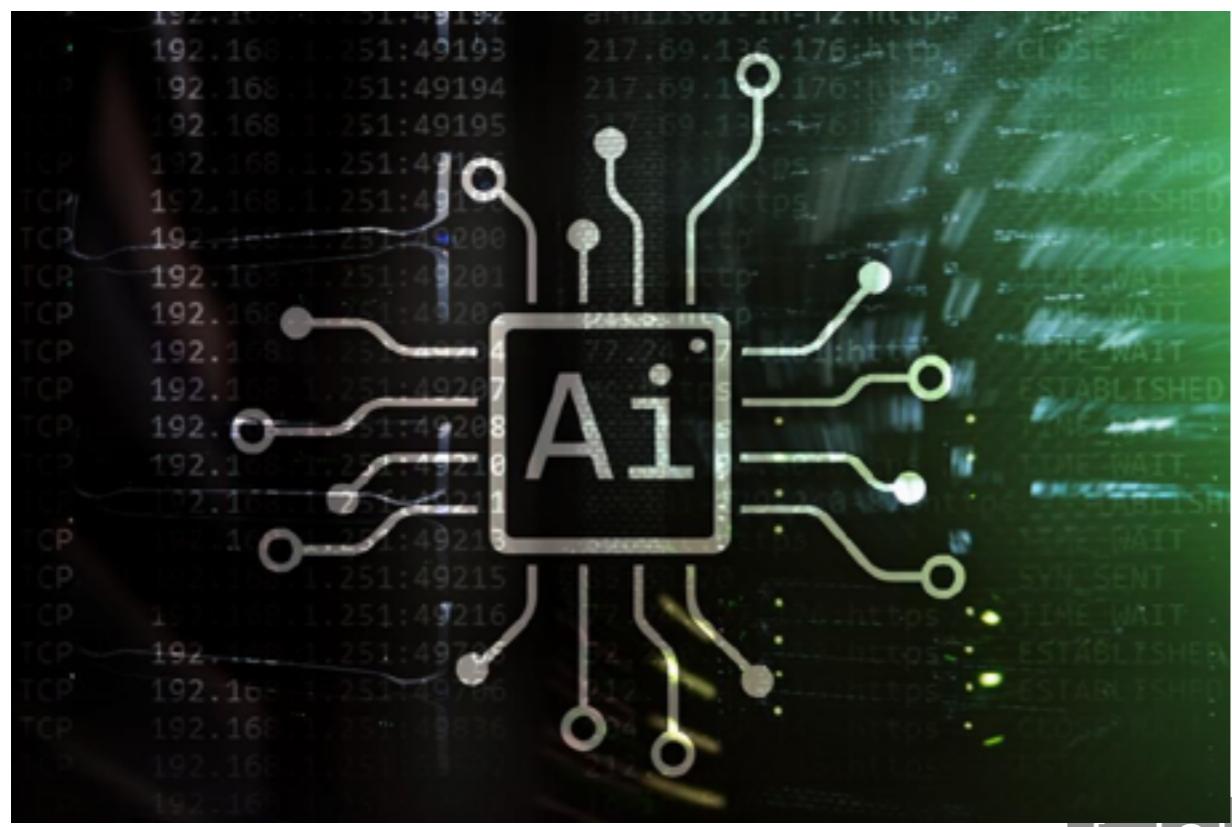
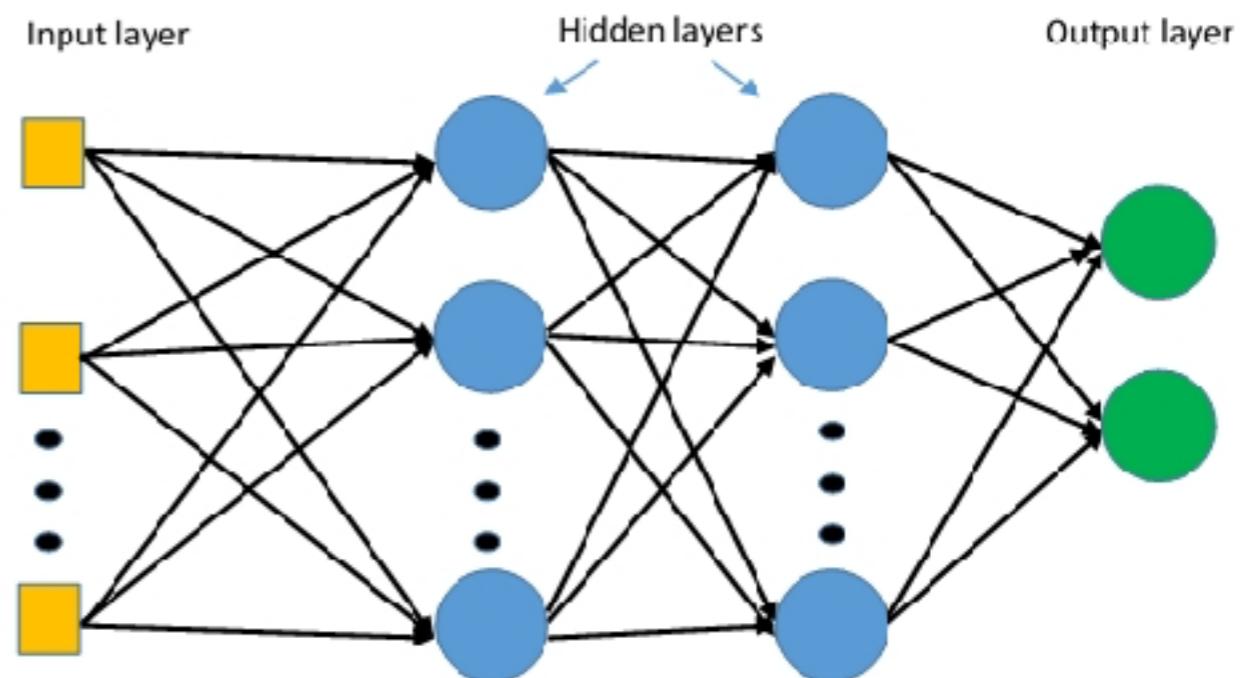
GPU

Graphics Processing Unit

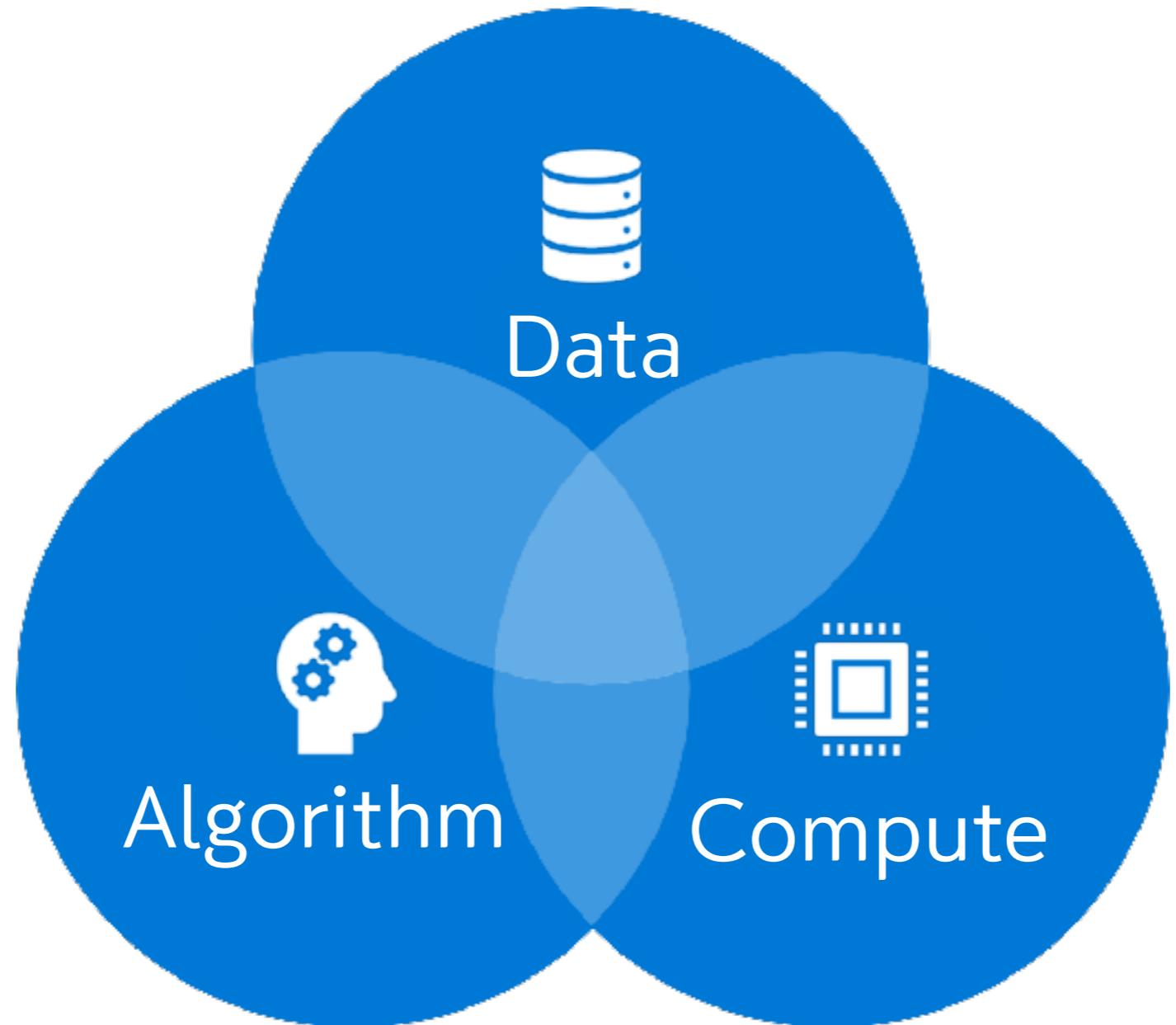
FPGA

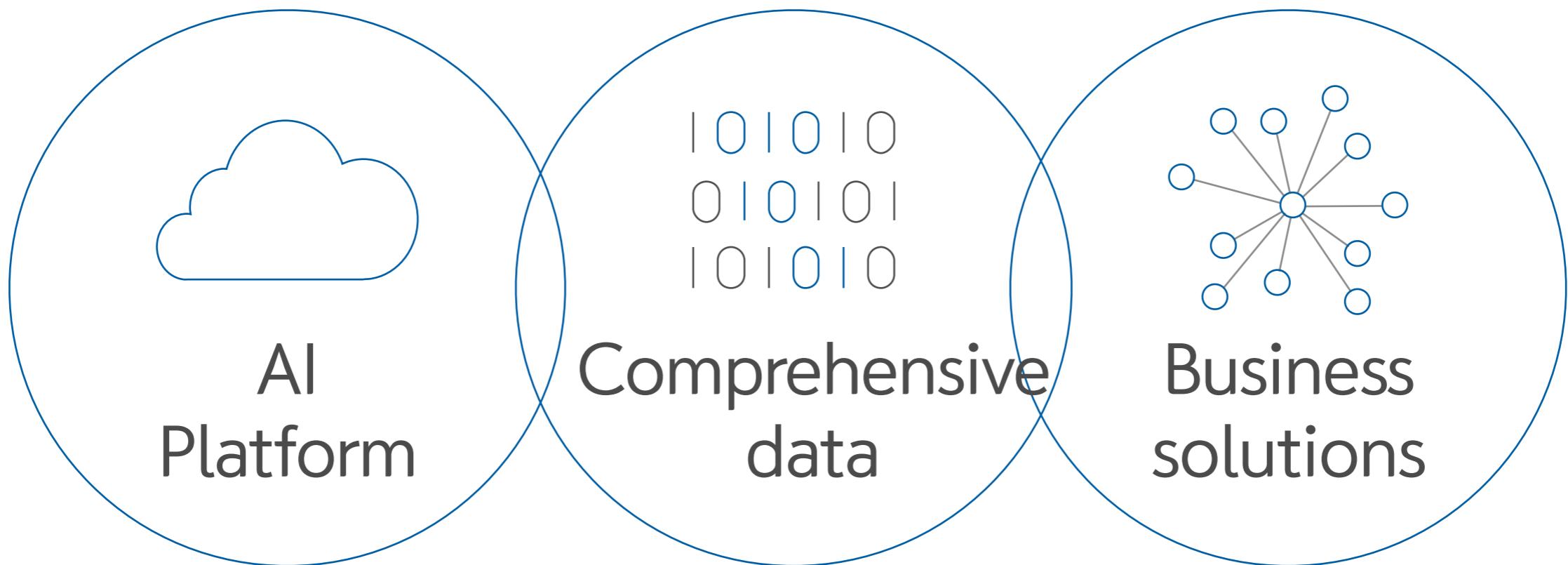
Field Programmable Gate Array





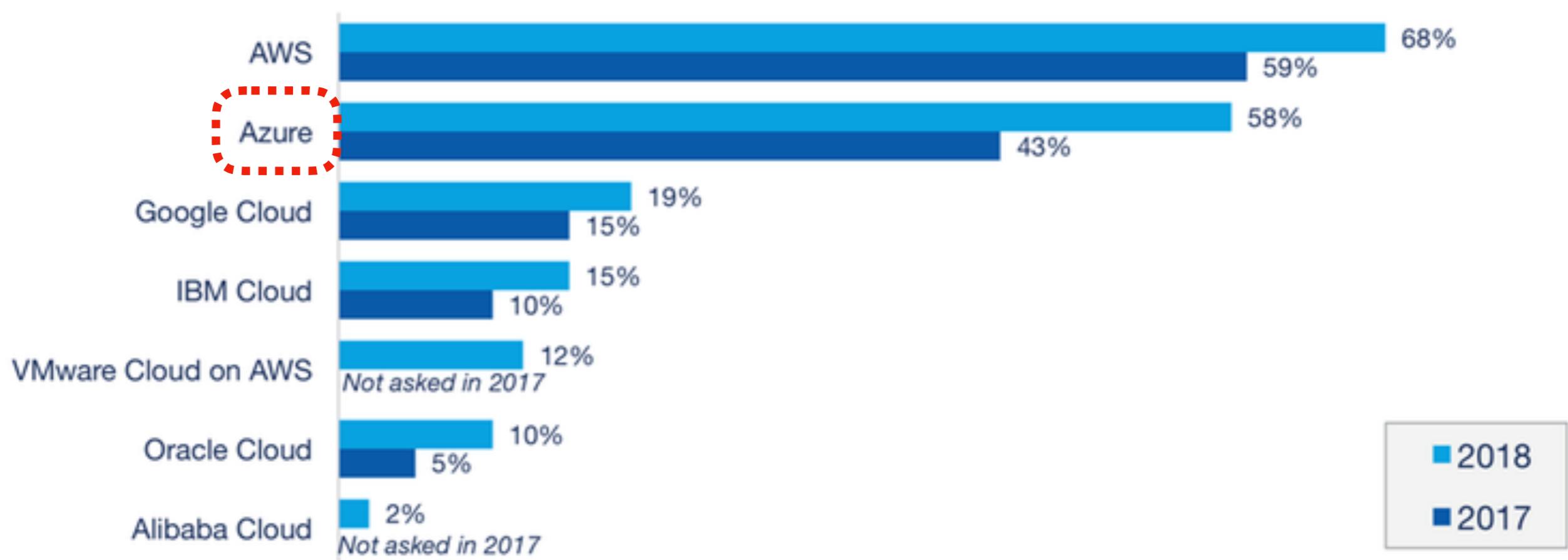
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정





Enterprise Public Cloud Adoption 2018 vs. 2017

% of Respondents Running Applications



<https://www.zdnet.com/article/top-cloud-providers-2018-how-aws-microsoft-google-ibm-oracle-alibaba-stack-up/>

Azure AI Service

custom model 생성



Machine Learning

쉽고 빠르게 모델을 만들어
배포, 관리하세요

microsoft model 활용



Cognitive Service

앱에 혁신적인 환경을
구현하세요

Azure AI Service



Machine Learning

쉽고 빠르게 모델을 만들어
배포, 관리하세요

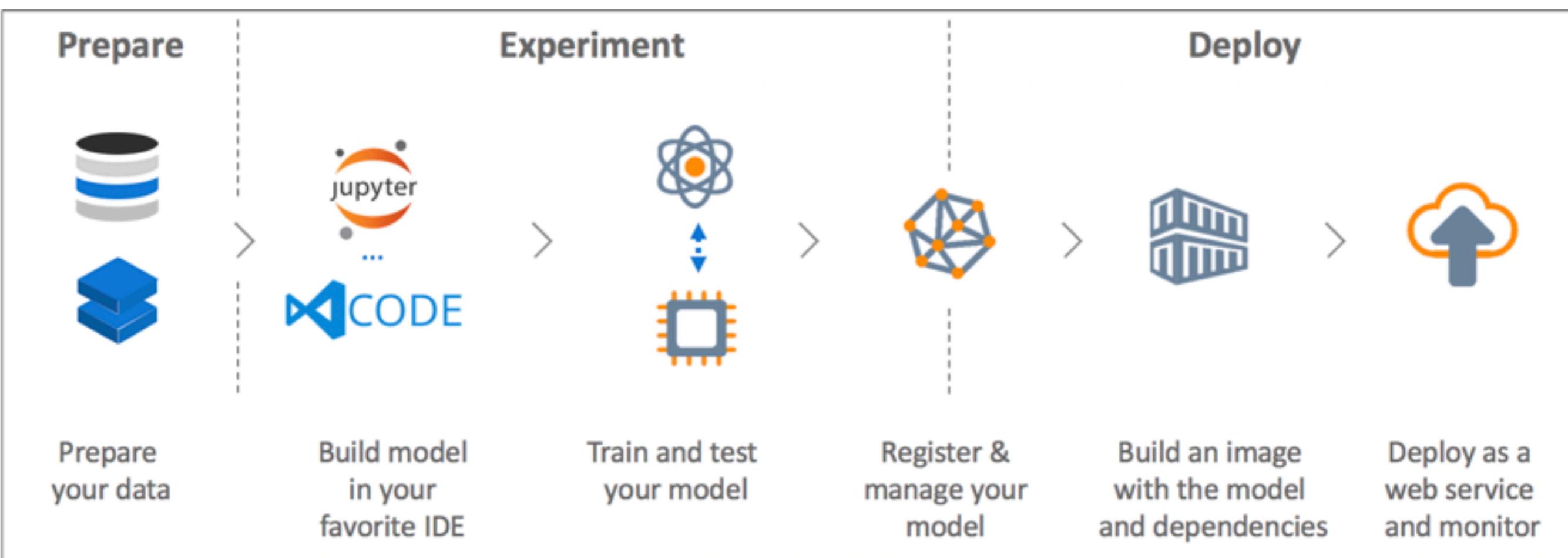
Build models quickly

 Azure Machine Learning

 Azure Databricks

 ONNX

Azure Machine Learning



Azure Machine Learning



Automated machine learning

Identify suitable algorithms and hyperparameters faster.



Managed compute

Train models with ease and reduce costs by autoscaling powerful GPU clusters.



DevOps for machine learning

Increase productivity with experiment tracking, model management and monitoring, integrated CI/CD, and machine learning pipelines.



Simple deployment

Deploy models on-premises, to the cloud, and at the edge with a few lines of code.



Tool-agnostic Python SDK

Azure Machine Learning service integrates with any Python environment, including Visual Studio Code, Jupyter notebooks, and PyCharm.



Support for open-source frameworks

Use your favorite machine learning frameworks and tools, such as PyTorch, TensorFlow, and scikit-learn.

Azure AI Service



Machine Learning

쉽고 빠르게 모델을 만들어
배포, 관리하세요

 Azure Machine Learning

 Azure Databricks

 ONNX

What is Azure Databricks?

A fast, easy and collaborative Apache® Spark™ based analytics platform optimized for Azure



- Designed in collaboration with the founders of Apache Spark
- One-click set up; streamlined workflows

Azure
Databricks

Home



Workspace



Recents



Data



Clusters



Jobs



Search

Azure Databricks



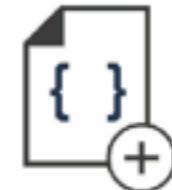
Explore the Quickstart Tutorial

Spin up a cluster, run queries on preloaded data, and display results in 5 minutes.



Import & Explore Data

Quickly import data, preview its schema, create a table, and query it in a notebook.



Create a Blank Notebook

Create a notebook to start querying, visualizing, and modeling your data.

Common Tasks

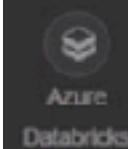
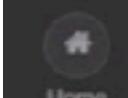
- New Notebook
- Upload Data
- Create Table
- New Cluster
- New Job
- New MLflow Experiment New
- Import Library

Recents

Recent files appear here as you work.

Documentation

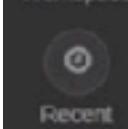
- Databricks Guide
- Python, R, Scala, SQL
- Importing Data

Azure
Databricks

Home



Workspace



Recent



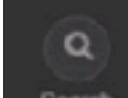
Data



Clusters

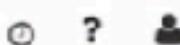


Jobs



Search

Demo Notebook (Python)



Attached: azuredatabricksdemo



View: Code



Run All



Schedule

Comments

Revision history

Max range : 5

Command took 2.30 seconds -- by saurinsh@microsoft.com at 10/19/2017, 6:23:51 AM on saurintest2

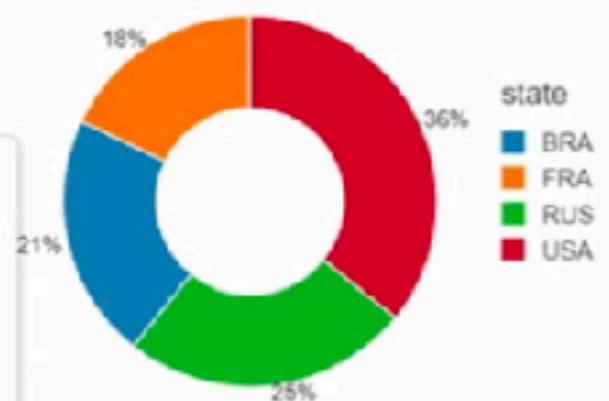
Cmd 9

Sales Overall Graph

Cmd 10

1 %sql select * from sales

(2) Spark Jobs



Command took 0.61 seconds -- by saurinsh@microsoft.com at 10/19/2017, 11:27:37 AM on azuredatabricksdemo

Cmd 11



Demo Job

< All Jobs

Demo Job

Job ID: 10

Task: Notebook at /Users/demo1@microsoft.com/

- ▶ Parameters: [Edit](#)
- Dependent Libraries: [Add](#)

Cluster: Driver: Standard_DS3_v2 [beta], Workers:

Schedule: None [Edit](#)

Advanced ▾

Active runs

Run	Start Time	Launched	Duration	Spark	Status
Run Now / Run Now With Different Parameters					

Completed in past 60 days

Latest successful run (refreshes automatically)

< Previous 20

Next 20 >

Run	Start Time	Launched	Duration	Spark	Status
<a href;"="">< Previous 20					

< Previous 20

Next 20 >

Schedule Job

Schedule

Every day at 01 00 US/Pacific

 Show Cron Syntax[Cancel](#)[Confirm](#)

Azure AI Service



Machine Learning

쉽고 빠르게 모델을 만들어
배포, 관리하세요

 Azure Machine Learning

 Azure Databricks

 ONNX

Open Neural Network Exchange

PyTorch

Chainer

Caffe2

Cognitive Toolkit

dmlc
XGBoost

scikit
learn

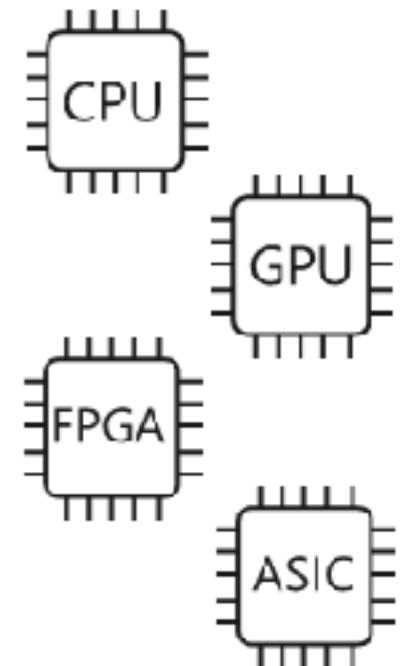
K

mxnet

F

ML

PaddlePaddle



Create

Frameworks

Caffe2 Chainer Cognitive Toolkit

mxnet PyTorch PaddlePaddle

ML.NET MathWorks XGBoost

ML scikit-learn TensorFlow Keras

Services

Azure Custom Vision Service

Native support

Converters

Native support

ONNX Model

Native support

Converters

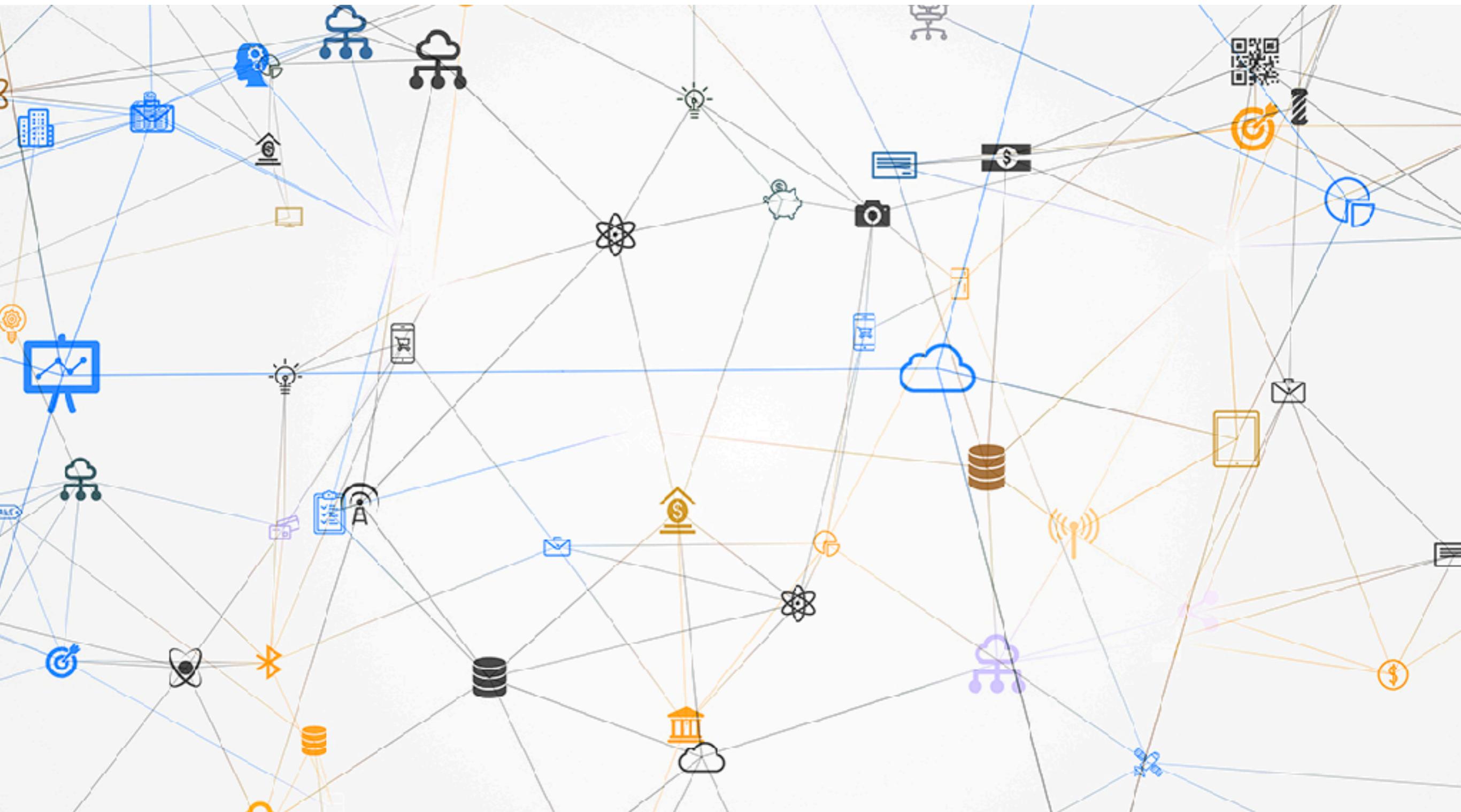
Deploy

Azure

Azure Machine Learning services
Ubuntu VM
Windows Server 2019 VM

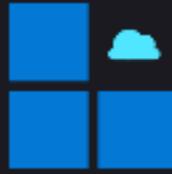
Windows Devices

Other Devices (iOS, etc)



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

Azure AI Service



Cognitive Service

앱에 혁신적인 환경을
구현하세요

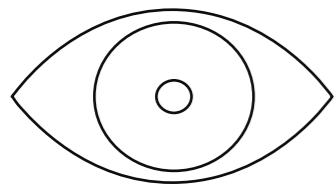
Deliver breakthrough experiences

Cognitive Services

시작 > 음성 > 언어 > 웹 검색 >

Bot Service

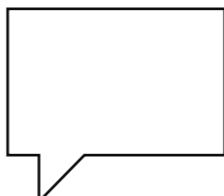
Cognitive Services



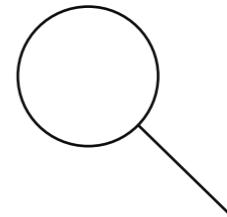
Vision



Language



Speech



Search



Decision

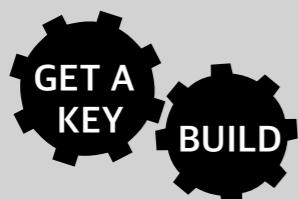
<https://azure.microsoft.com/ko-kr/services/cognitive-services/>

Cognitive Services

Easy

Roll your own with REST APIs

Simple to add: just a few lines
of code required



Flexible

Make the same API code call on
iOS,
Android, and Windows

Integrate into the language and
platform of your choice

Bring your own data for your
custom experience



Tested

Built by experts in their field from
Microsoft Research, Bing, and
Azure Machine Learning

Quality documentation, sample
code, and community support



https://github.com/Microsoft/Ignite2017_MSAIDevBootcamp_Pre-day

Vision



Computer Vision API

[체험하기](#) | [자세한 정보](#)

이미지 분류

이미지의 장면 및 활동 인식 | [데모](#)

이미지의 유명 인사 및 이정표 인식 | [데모](#)

이미지의 OCR(광학 문자 인식) | [데모](#) | [컨테이너 지원](#)



Ink Recognizer 미리 보기

[체험하기](#) | [자세한 정보](#)

디지털 잉크 및 필기 인식

일반 도형 인식 및 찾기

생산성 시나리오 지원



Custom Vision

[체험하기](#) | [자세한 정보](#)

사용자 지정 가능한 이미지 인식 | [데모](#)



Face

[체험하기](#) | [자세한 정보](#) | [컨테이너 지원](#)

이미지의 얼굴 감지 | [데모](#)

이미지의 사람 식별

이미지의 감정 인식 | [데모](#)

이미지의 유사한 얼굴 식별 및 그룹화 | [데모](#)



Video Indexer

[체험하기](#) | [자세한 정보](#)

동영상의 시작적 개체 및 오디오 채널에 대한 전체 분석

얼굴, 개체, 키 프레임 인식, OCR 및 전사 수행

토픽 유추, 브랜드 및 감정 감지 같은 고급 인사이트



Form Recognizer 미리 보기

[자세한 정보](#) | [컨테이너 지원](#)

문서에서 텍스트, 키-값 쌍 및 테이블 추출

수동 레이블 지정 없이 양식에 사용자 지정됨

클라우드에서 에지에 이르기까지 어디서나 배포

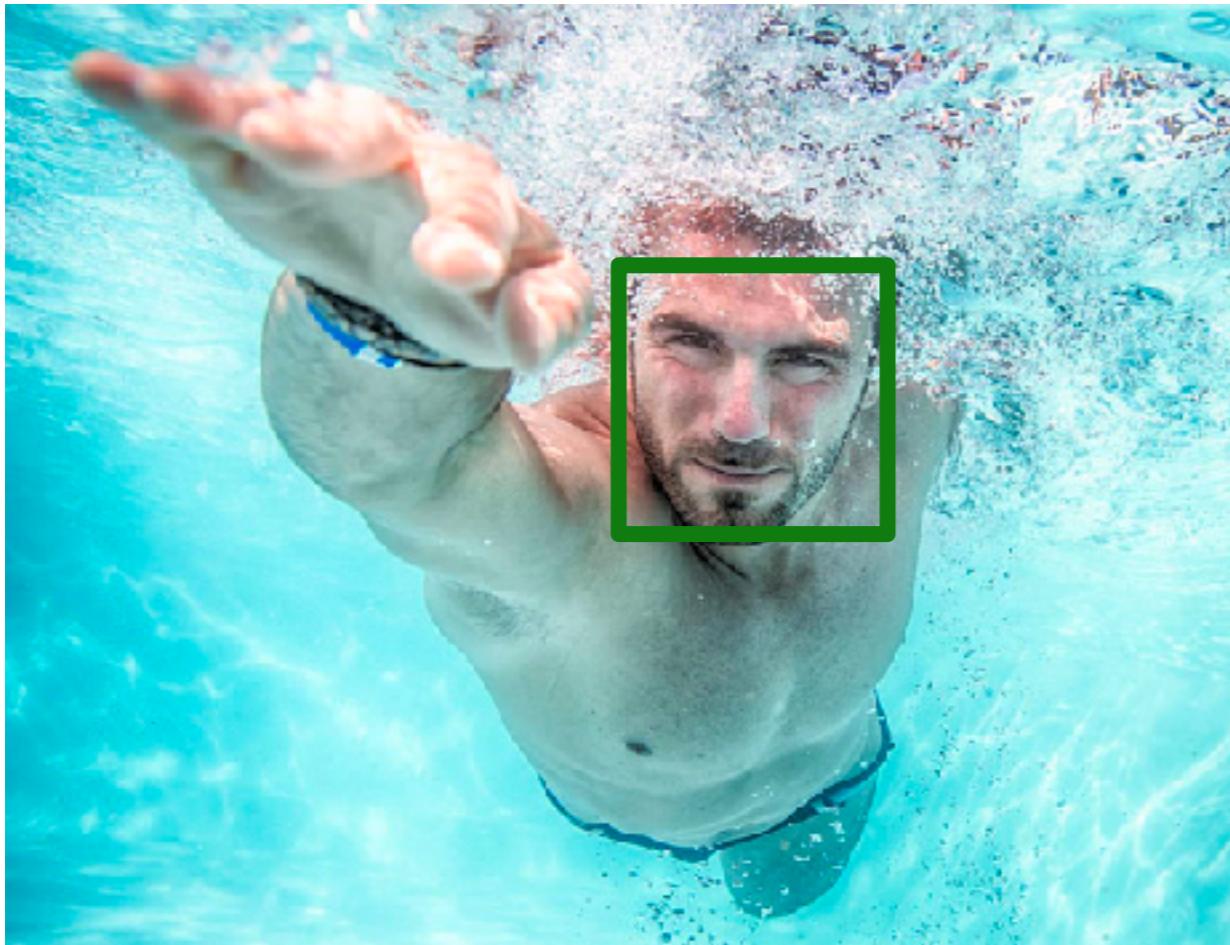
Vision



Colors

```
"color": {  
    "dominantColorForeground": "Grey",  
    "dominantColorBackground": "White",  
    "dominantColors": [ "White" ], ■  
    "accentColor": "19A4B2",  
    "isBWImg": false  
},  
"imageType": {  
    "clipArtType": 0, // False  
    "lineDrawingType": 0 // False  
}
```

Face



```
"faces": [
  {
    "age": 36,
    "gender": "Male",
    "faceRectangle": {
      "left": 298,
      "top": 133,
      "width": 121,
      "height": 121
    }
  }
]
```

Tag

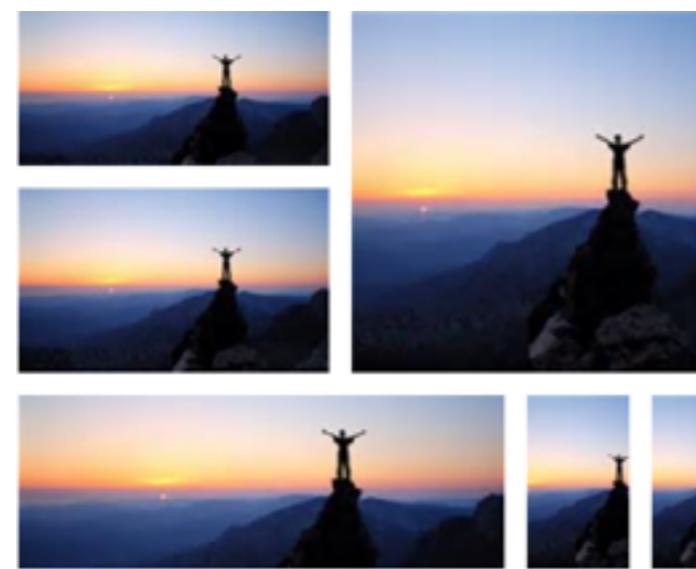


```
"tags": [
  {
    "name": "water",
    "confidence": 0.9997857213020325
  },
  {
    "name": "swimming",
    "confidence": 0.9556196331977844
  },
  {
    "name": "sport",
    "confidence": 0.9538078308105469
  }...
```

Original



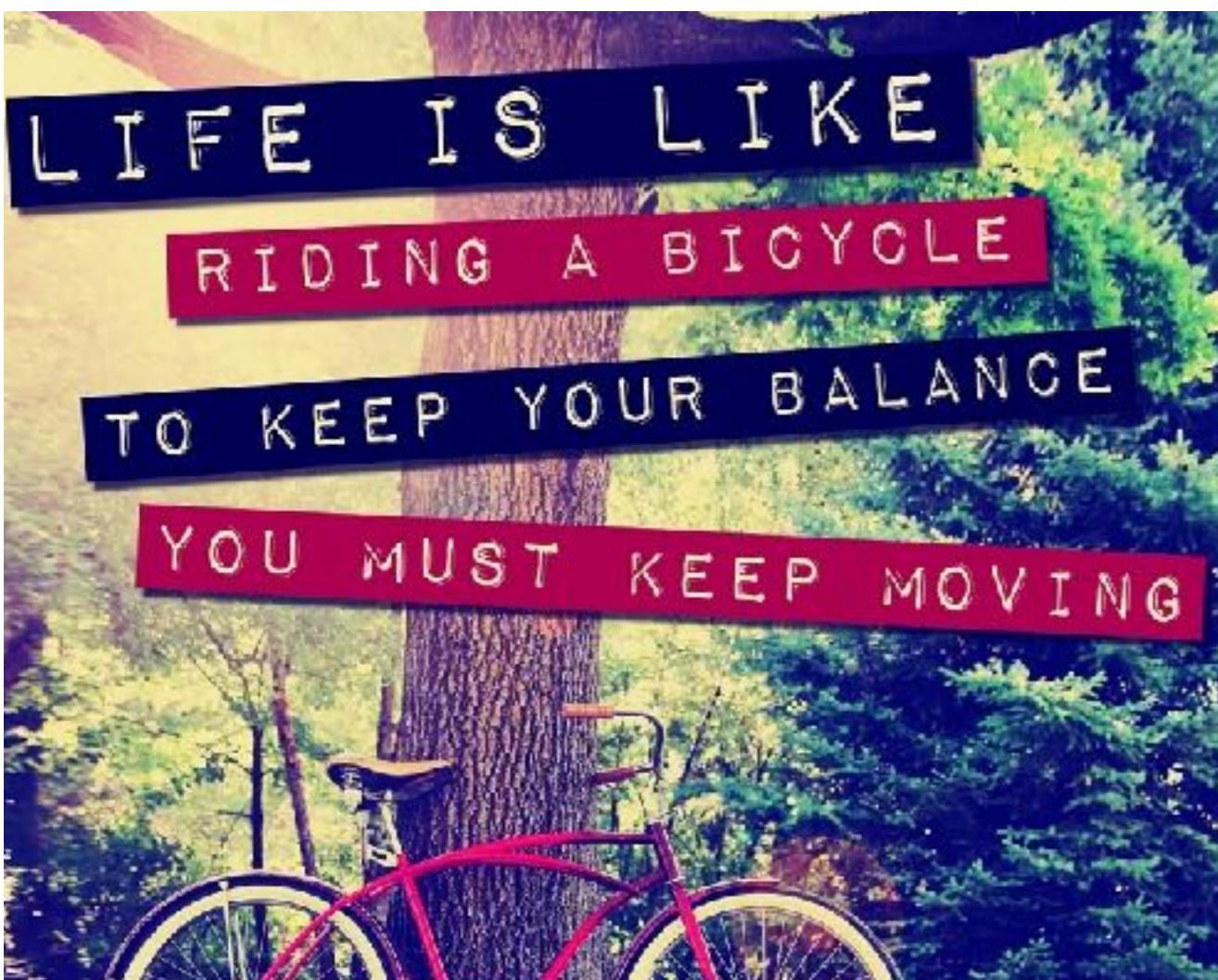
Smart crop OFF



Smart crop ON

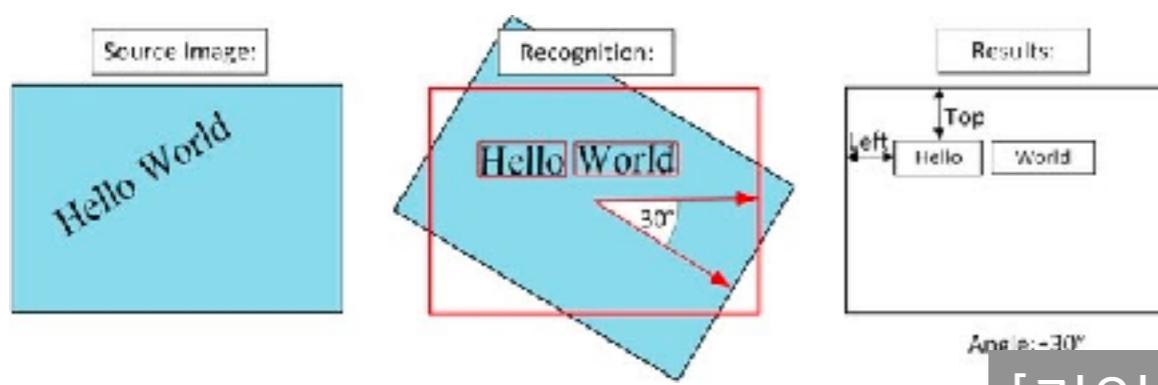
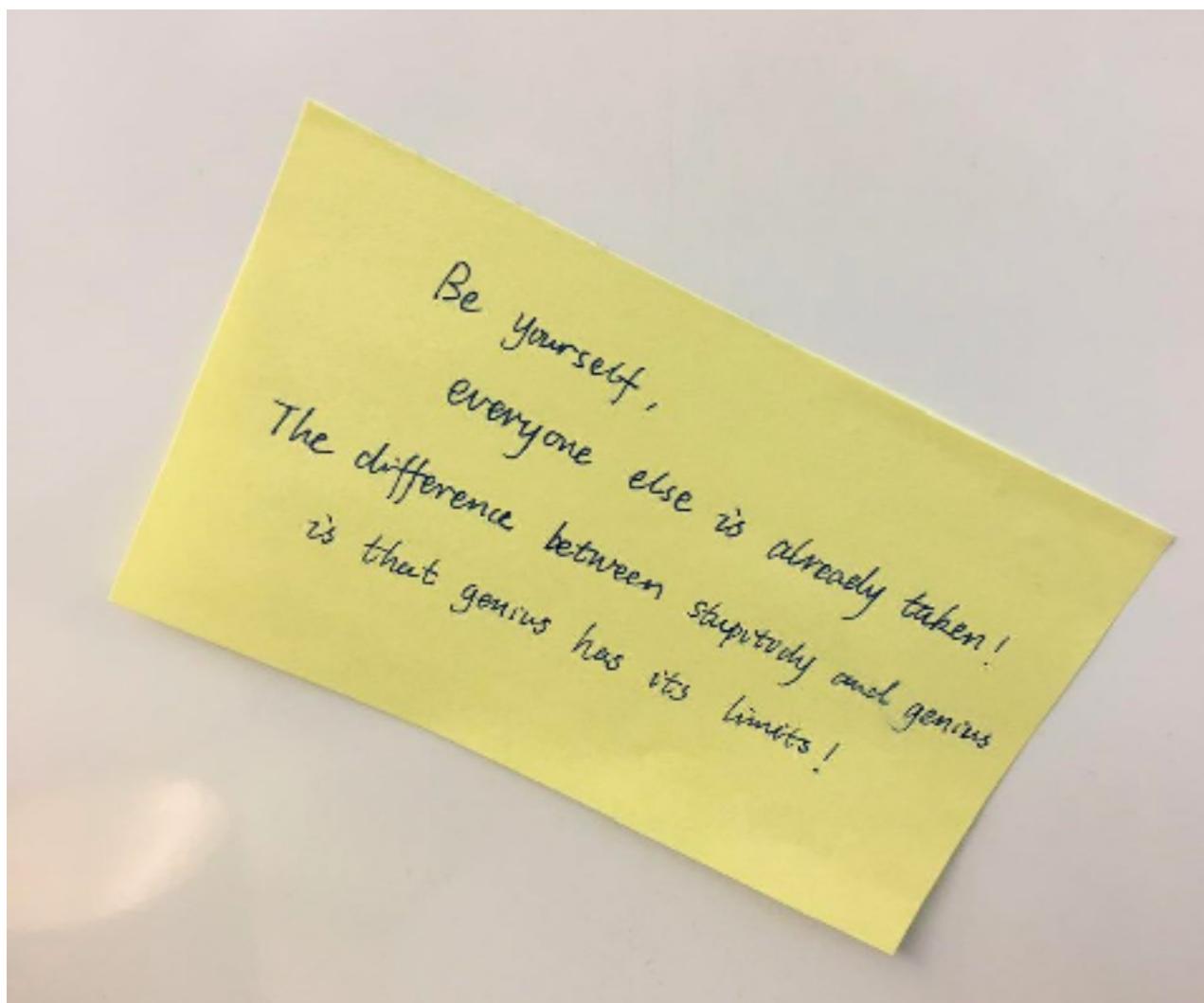


OCR



```
{  
  "language": "en",  
  "orientation": "Up",  
  "regions": [  
    {  
      "boundingBox": "41,77,918,440",  
      "lines": [  
        {  
          "boundingBox": "41,77,723,89",  
          "words": [  
            {  
              "boundingBox": "41,102,225,64",  
              "text": "LIFE"  
            }...  
          ]  
        ]  
      ]  
    ]  
  ]  
}
```

OCR



Handwritten

{

```
"status": "Succeeded",
"succeeded": true,
"failed": false,
"finished": true,
"recognitionResult": {
  "lines": [
    {
      "boundingBox": [ 354,
196, ... ],
      "text": "Be yourself",
      "words": [
        {
          "boundingBox": [ ... ],
          "text": "Be"
        }...
      ]
    }
  ]
}
```

Language



Text Analytics

[체험하기](#) | [자세한 정보](#) | [컨테이너 지원](#)

명명된 엔터티 인식 | [데모](#)

핵심 문구 추출 | [데모](#)

텍스트 감정 분석 | [데모](#)



Translator Text

[체험하기](#) | [자세한 정보](#)

자동 언어 감지 | [컨테이너 지원](#)

자동화된 텍스트 번역

사용자 지정 가능한 번역



QnA Maker

[체험하기](#) | [자세한 정보](#)

구조화되지 않은 텍스트에서 Q&A 추출 | [데모](#)

Q&A 컬렉션을 통한 기술 자료 만들기 | [데모](#)



Language Understanding

[체험하기](#) | [자세한 정보](#) | [컨테이너 지원](#)

상황에 맞는 언어 해석 | [데모](#)



몰입형 리더

[미리 보기](#)

[자세한 정보](#)

사용자가 텍스트를 읽고 이해할 수 있도록 지원

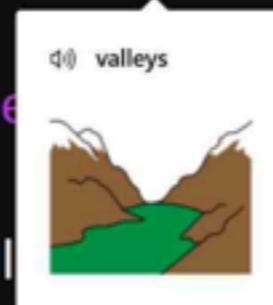
장애 유무와 관계없이 모든 독자를 위한 기능



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



The study of Earth's land-forms is called physical geography. Land-forms can be mountains and valleys. They can also be glaciers or rivers. Land-forms are sometimes called physical features. It is important for students to know about the physical geography of Earth. The seasons, the



Speech



음성 서비스

[체험하기](#) | [자세한 정보](#)

사용자 지정 가능한 모델에서 자동 음성 텍스트 변환 전사 | [데모](#) | [컨테이너 지원](#)

사용자 지정 음성 글꼴을 사용하는 자연어 텍스트 음성 변환 | [데모](#) | [컨테이너 지원](#)

실시간 음성 번역 | [데모](#)



Speaker Recognition 미리 보기

[체험하기](#) | [자세한 정보](#)

화자 식별 | [데모](#)

화자 검증 | [데모](#)



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



▷ 등록 음성



텍스트

JSON

대통령 Barack Obama

은 선택한 오디오의 식별된 연설입니다.

Search



Bing Spell Check

[체험하기](#) | [자세한 정보](#)

웹 규모의 다국어 맞춤법 검사 | [데모](#)

상황에 맞는 맞춤법 검사



Bing Web Search

[체험하기](#) | [자세한 정보](#)

광고 없는 웹 검색 | [데모](#)

안전한 웹 검색

위치 인식 웹 검색



Bing Visual Search

[체험하기](#) | [자세한 정보](#)

유사한 이미지 및 제품 식별 | [데모](#)

이미지를 통한 정보 취득

웹 소스 식별



Bing Custom Search

[자세한 정보](#)

사용자 지정 검색 엔진 만들기

광고 없는 사용자 지정 검색 결과 | [데모](#)



Bing Entity Search

[체험하기](#) | [자세한 정보](#)

명명된 엔터티 인식 및 분류 | [데모](#)

명명된 엔터티에 대한 정보 취득



Bing Video Search

[체험하기](#) | [자세한 정보](#)

광고 없는 비디오 검색 | [데모](#)

비디오 토픽 및 추세 식별



Bing News Search

[체험하기](#) | [자세한 정보](#)

Bing Internet Search

[체험하기](#) | [자세한 정보](#)

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

태풍



미리보기

JSON

예시

european football

financial markets

Hollywood updates

science updates

top stories

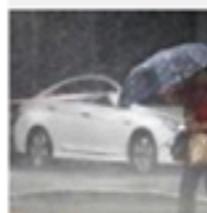
world news

시장

ko-kr (한국어)

유해 정보 차단

엄격



태풍 '레끼마' 영향 제주 300mm 비

태풍 '레끼마'의 영향으로 강한 바람이 불면서 10일 오후 제주 서귀포시 성산읍 섭지코지 앞 바다에는 높은 파도가 치고 있다. 연합뉴스 중국에 상륙한 제9호 태풍 '레끼마' 영향으로 12일 ...

태풍 '레끼마' 접근...제주·전남 '폭우', 서해안 '강풍'

한반도가 중국에서 북상 중인 제9호 태풍 레끼마의 영향권에 들면서 제주와 남해안에서 시작된 비는 오늘 밤 전국 대부분 지역으로 확되될 ...

中 상륙한 태풍 '레끼마' 영향 전국에 비...'크로사'는 광복절에 日 관통

11일 기상청에 따르면 예상 강수량은 전날부터 12일까지 제주도와 서해 5도가 40~100mm, 서울과 경기, 강원 영서 북부, 충남 서부, 전북 서부, 전남 ...

태풍 레끼마·크로사, 우리나라 간접 영향 끼칠 가능성..."비바람 세기 변동"

제9호 태풍 '레끼마(LEKIMA)'가 약화된 가운데 우리나라 일부지역에 간접 영향을 끼칠 가능성이 제기됐다. 제10호 태풍 '크로사(KROSA)'도 일본 방향으로 북상 중인 가운데 간접 영향을 끼칠 ...

All 애니메이션 Gif 클립 아트 라인 사진

cute animals



예시

abstract art

cute animals

celebrations

nature wallpapers

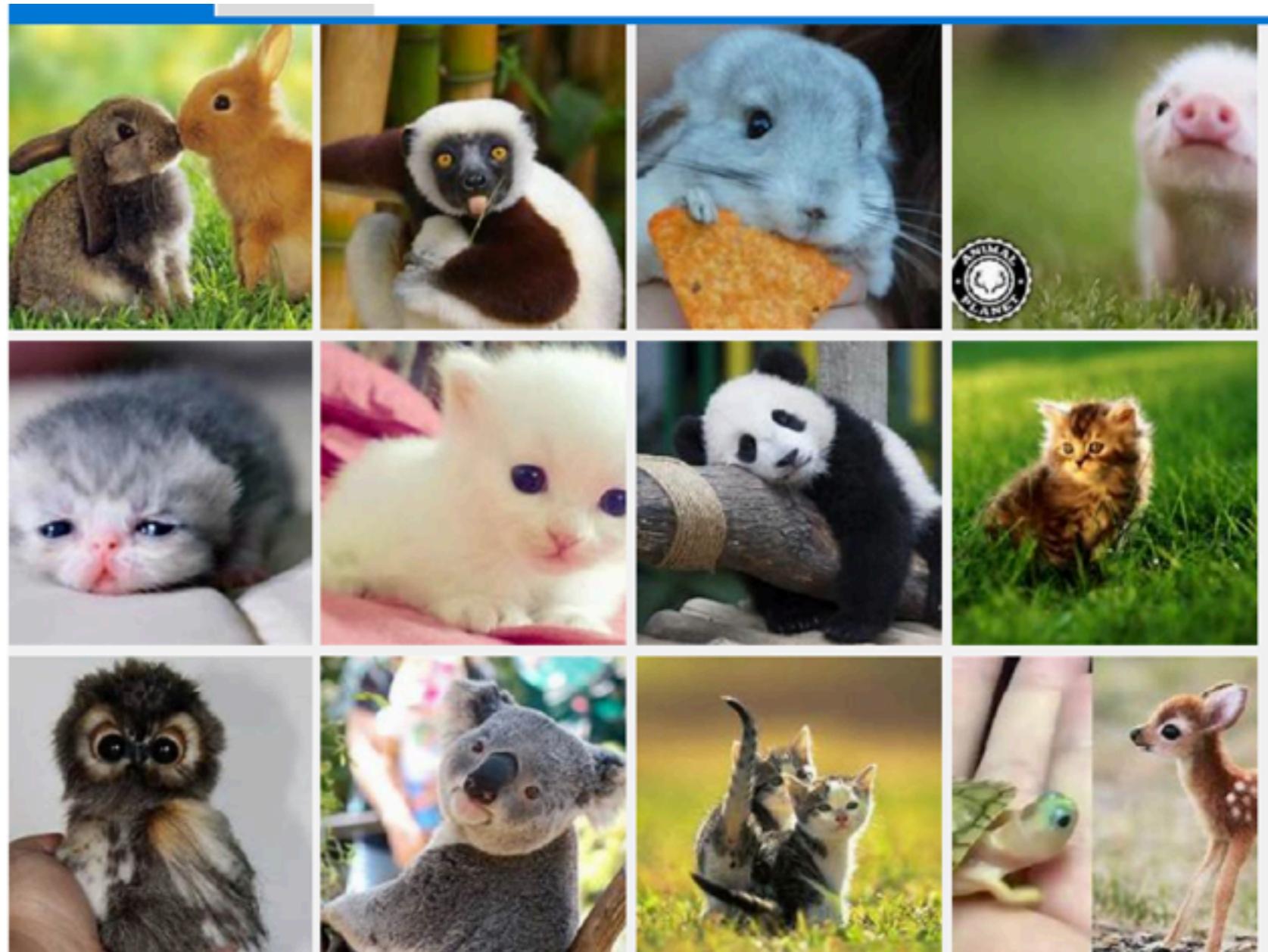
sunsets

시장

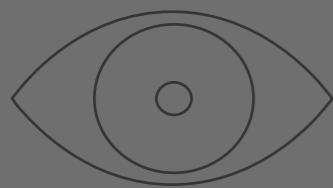
ko-kr (한국어)

유해 정보 차단

엄격



Cognitive Services



Vision



Language



Speech



Search



Decision

Decision



Content Moderator

[체험하기](#) | [자세한 정보](#)

잠재적인 불쾌감을 주는 이미지 및 원치 않는 이미지 감지

잠재적인 불경한 언어 및 바람직하지 않은 텍스트 필터링

동영상의 성인/외설 콘텐츠 조정

최상의 결과를 위한 기본 제공 검토 도구 사용



Anomaly Detector

[미리 보기](#) | [체험하기](#) | [자세한 정보](#) | [컨테이너 지원](#)

실시간으로 비즈니스 상태 모니터링

IoT-원격 모니터링 수행

대화형 데이터 분석 활용



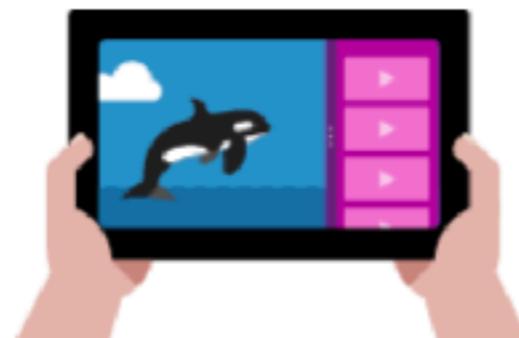
Personalizer

[미리 보기](#) | [체험하기](#) | [자세한 정보](#)

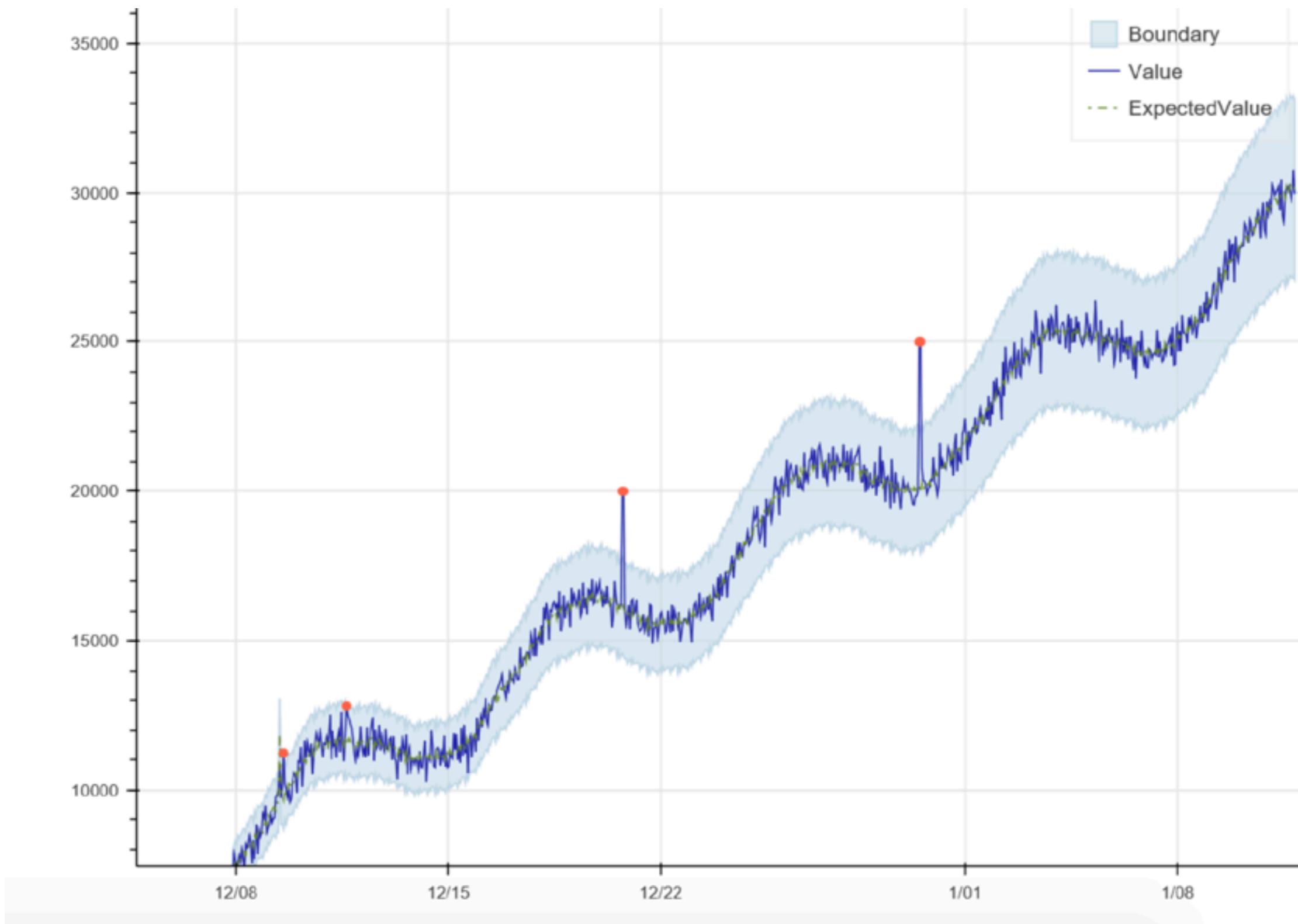
앱에서 풍부한 맞춤형 환경 제공

클라우드에서 에지에 이르기까지 어디서나 배포

보충 학습 루프 이해 및 손쉽게 관리

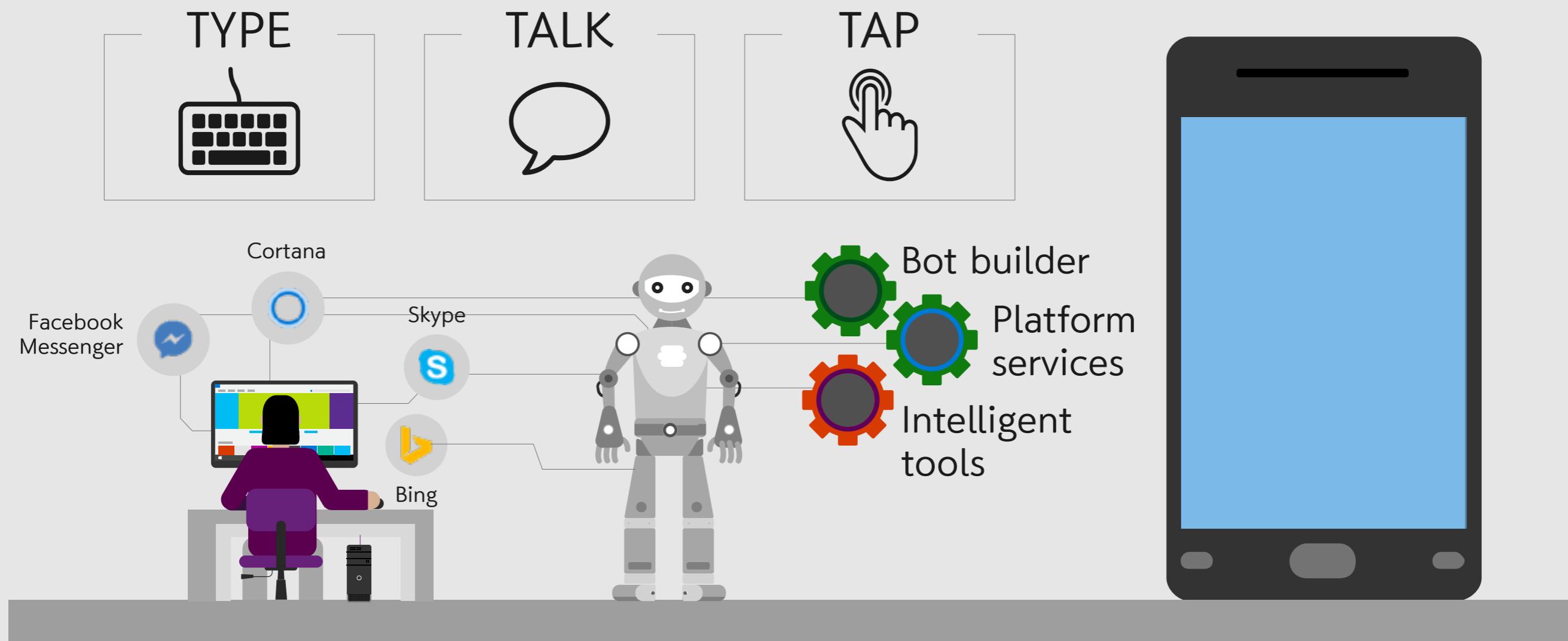


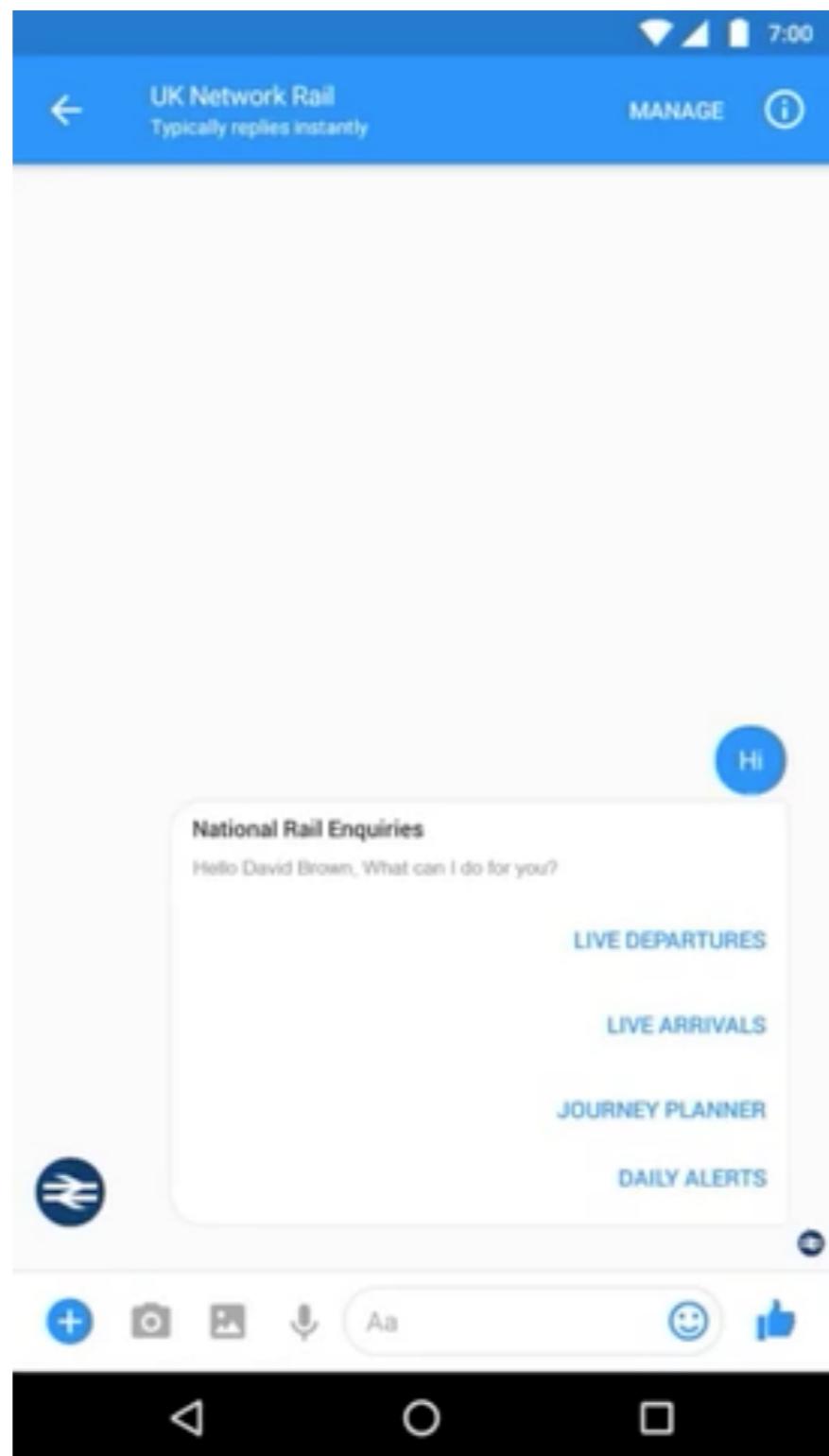
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

Cognitive Services – Bot Service





[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

Cognitive
Services

Customizabl
e Cognitive
Services

Azure
Machine
Learning

Cognitive
Toolkit
(CNTK)



첫 번째 날

- 인공지능 개론
- 인공지능 사례
- 머신러닝 개론
- Classification 모델 생성

인공지능, 어디에 사용되나요?

의료

마케팅

보안

금융

스포츠

환경

예술

농수산

안전

법률

모바일

쇼핑

마케팅

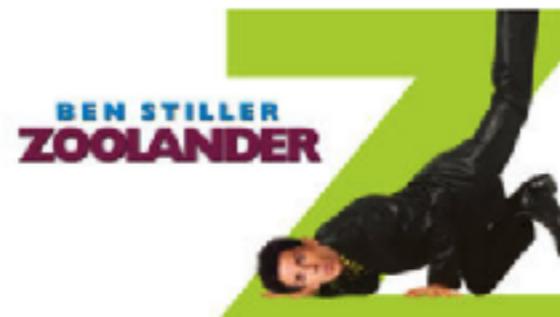
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

넷플릭스



시청 콘텐츠 약 75% AI 추천

넷플릭스



사용자 취향에 맞는 영화 포스터 제공

식품

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

와인



와인 생산에
가장 중요한 요소?

물

와인



E&J Gallo 와이너리

“포도의 상태에 따라 필요한 물을 알 수 없을까?”

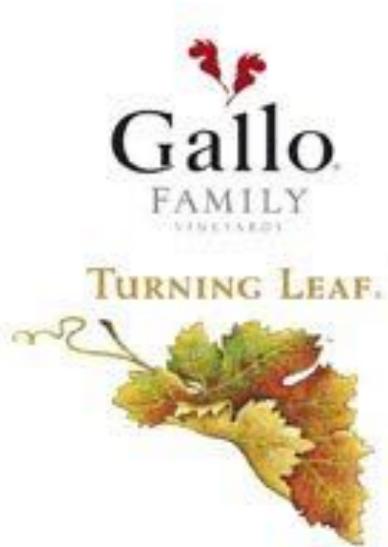
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

와인



인공위성 사진, 포도밭의 센서로 부터 얻는 데이터 분석

와인



- 물 사용량 25% 감소
- 포도 수확량 30% 증가
- 비용 절감, 포도 품질 향상

맥주



Sugar Creek
BREWING COMPANY



미국 노스캐롤라이나주 Sugar Creek

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

맥주



맥주 숙성 시간, 온도, pH, 압력, 탄산 정도 분석

=> 맥주 품질 향상

맥주



Bosch 정밀 유압계 + IoT 센서

=> 병입과정에서 발생하는 파손 문제 해결

과자



롯데제과



소셜 데이터, POS 판매 데이터, 구매 연령, 지역별 소비 분석

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

과자



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

스포츠

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

NBA



매시즌 연봉 60억원의 선수 선발

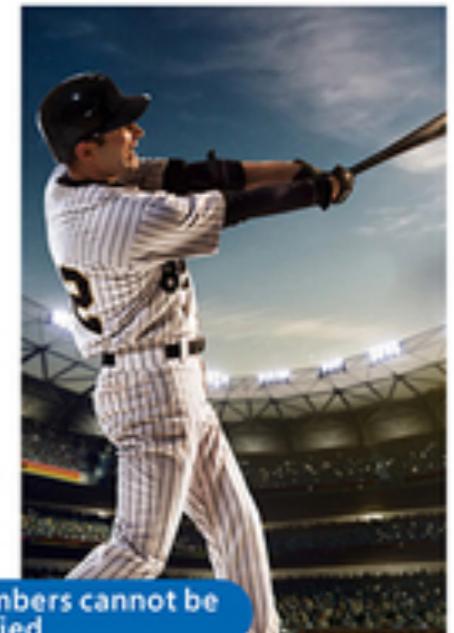
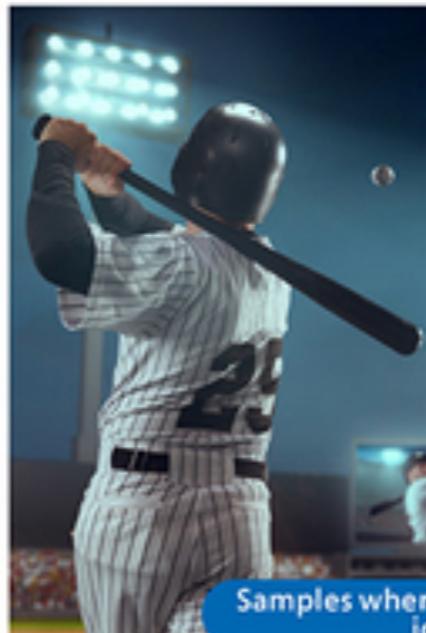
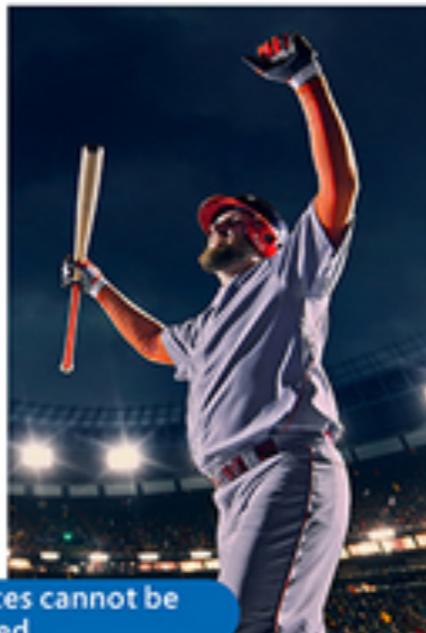
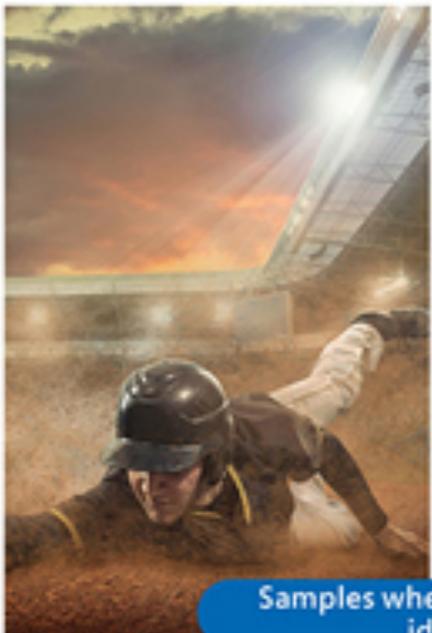
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

NBA



슛, 어시스트, 리바운드, 개인성향, 팀워크, 팀 기여도 예측

프로야구





<https://customers.microsoft.com/en-us/story/real-madrid>

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



“Our core business is football, so having Microsoft as our technology partner to manage the technology infrastructure is important. The Azure platform will help us provide the services we want to offer because it gives us the scalability we need whenever we need It.”

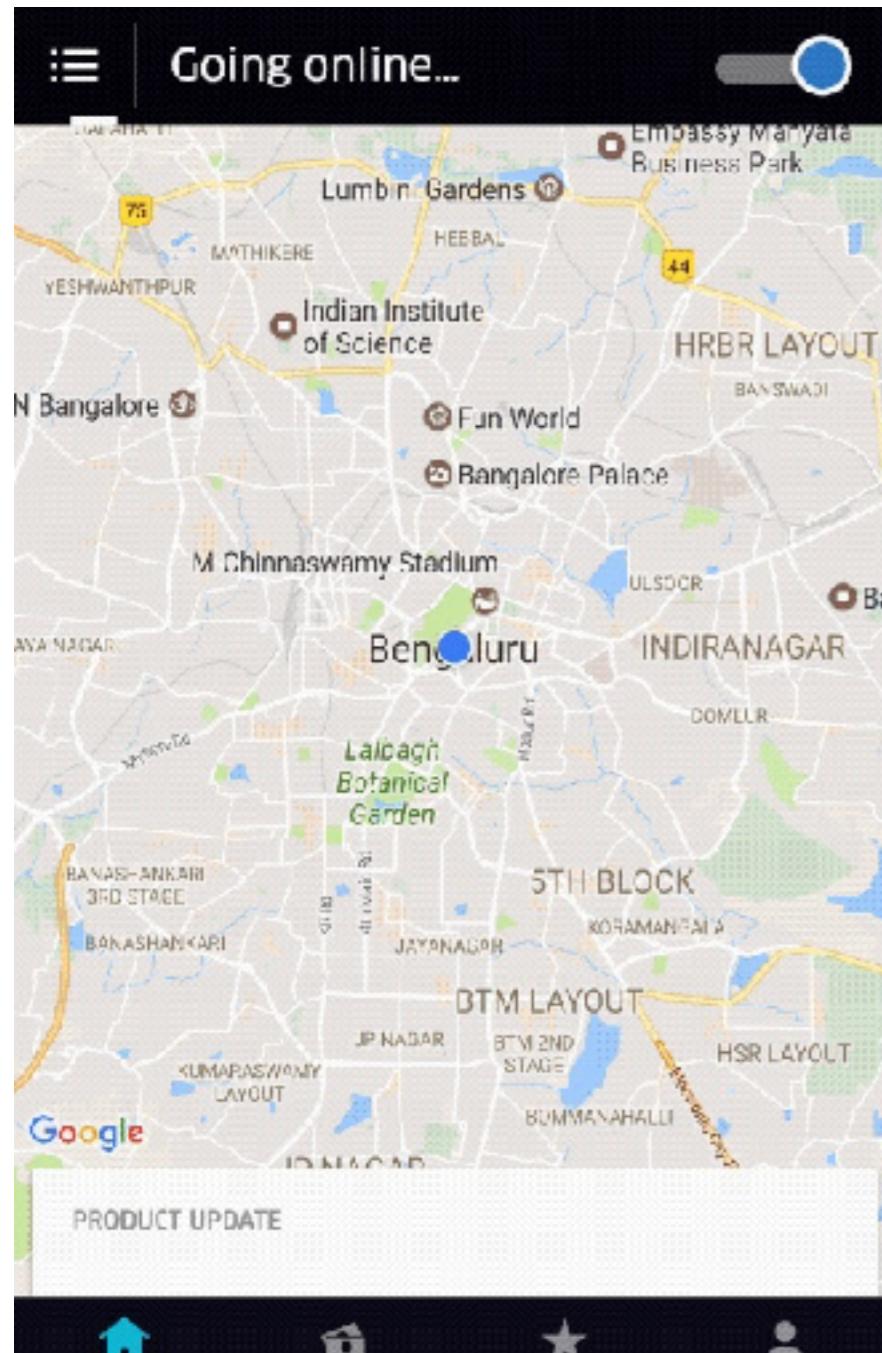
- Enrique Uriel: CIO
Real Madrid

보안/안전

호주 국립 은행 ATM



Uber



Position your face inside the circle and tap the button to verify your identity

Identity Verified.



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

Shell



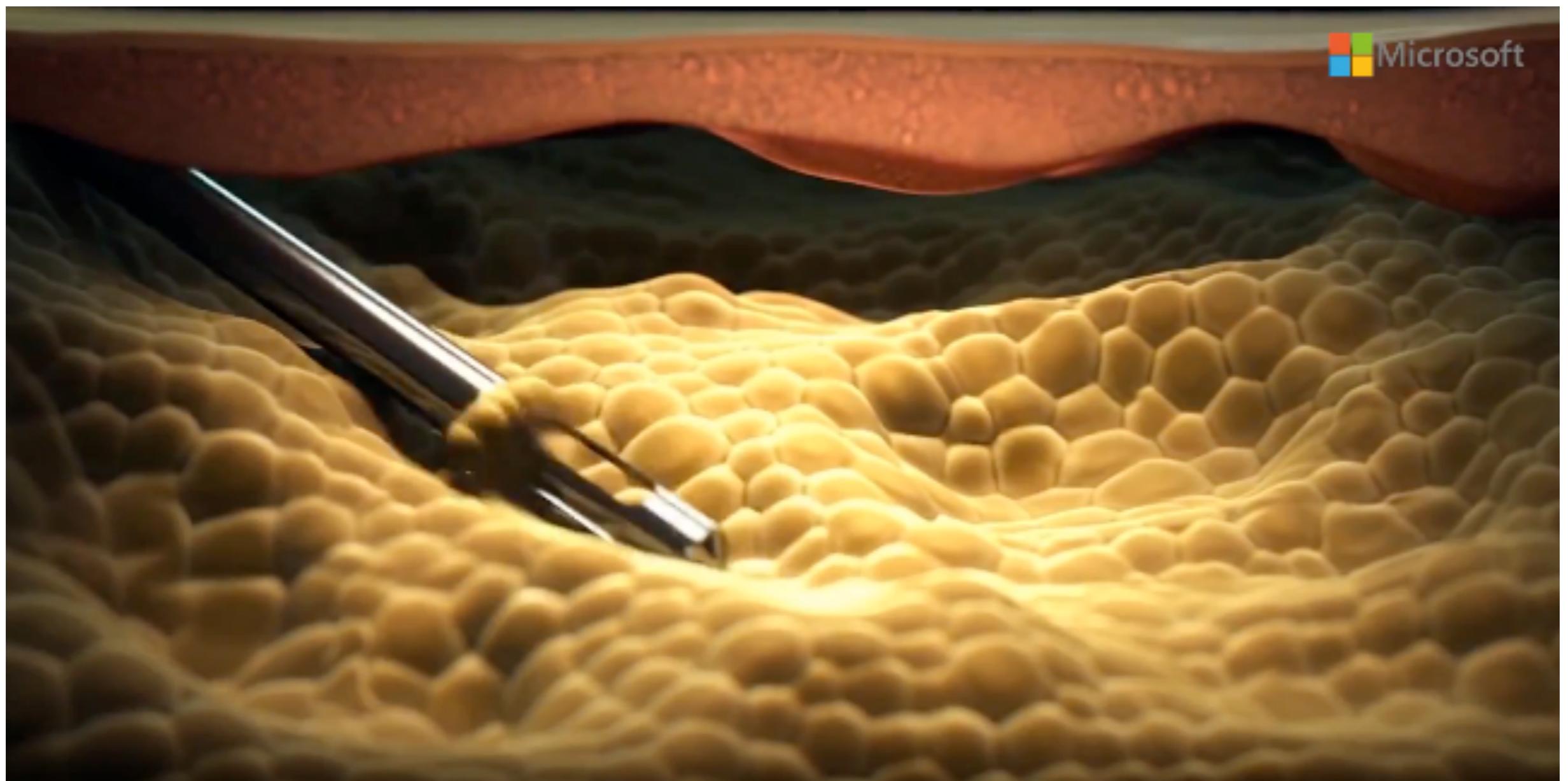
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

의료

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

<https://customers.microsoft.com/en-us/story/365mc-azure-iot-suite-machine-learning-korea-en>

365 mc



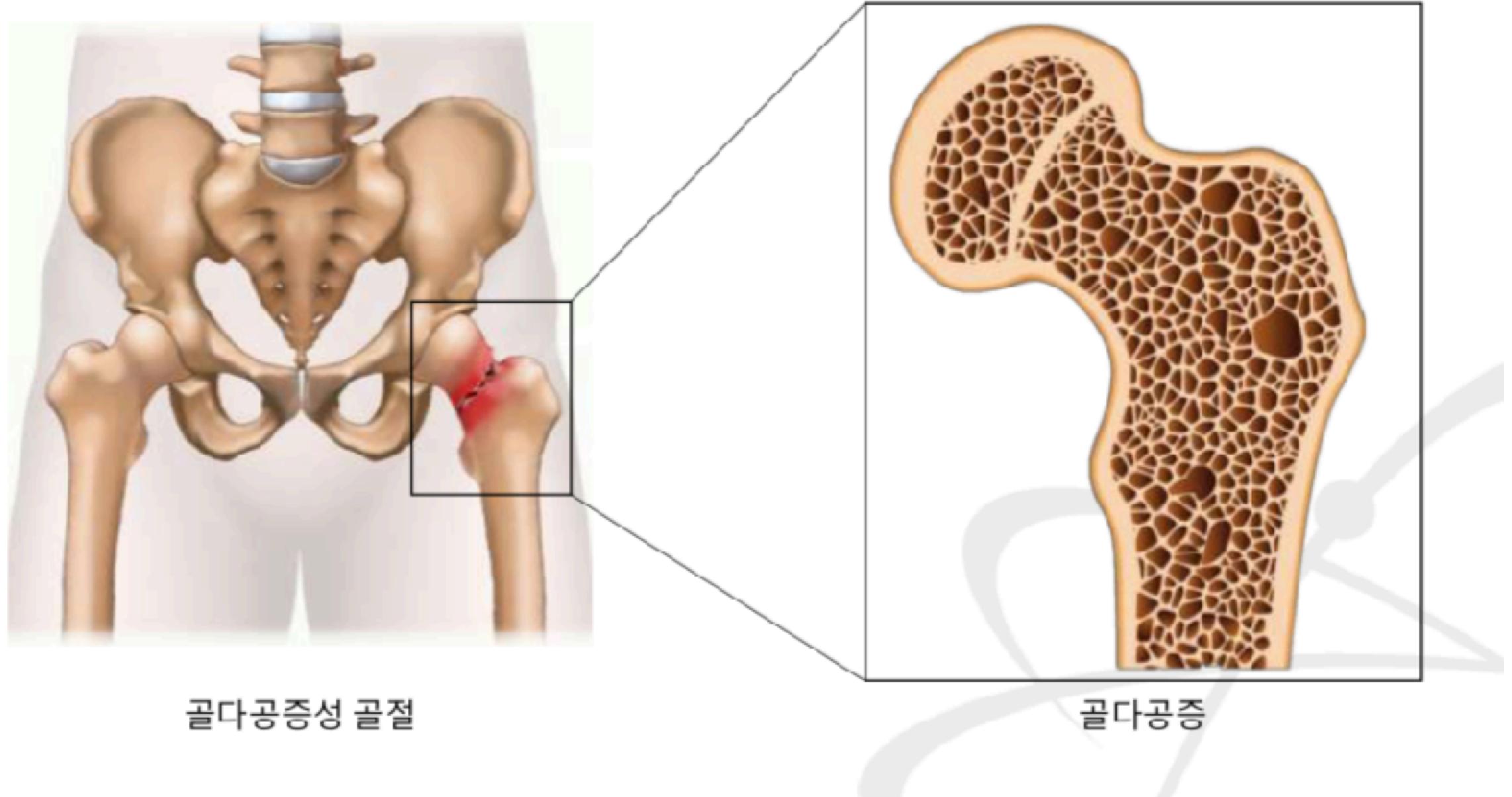
지방 흡입 시술 및 회복 상황 분석

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

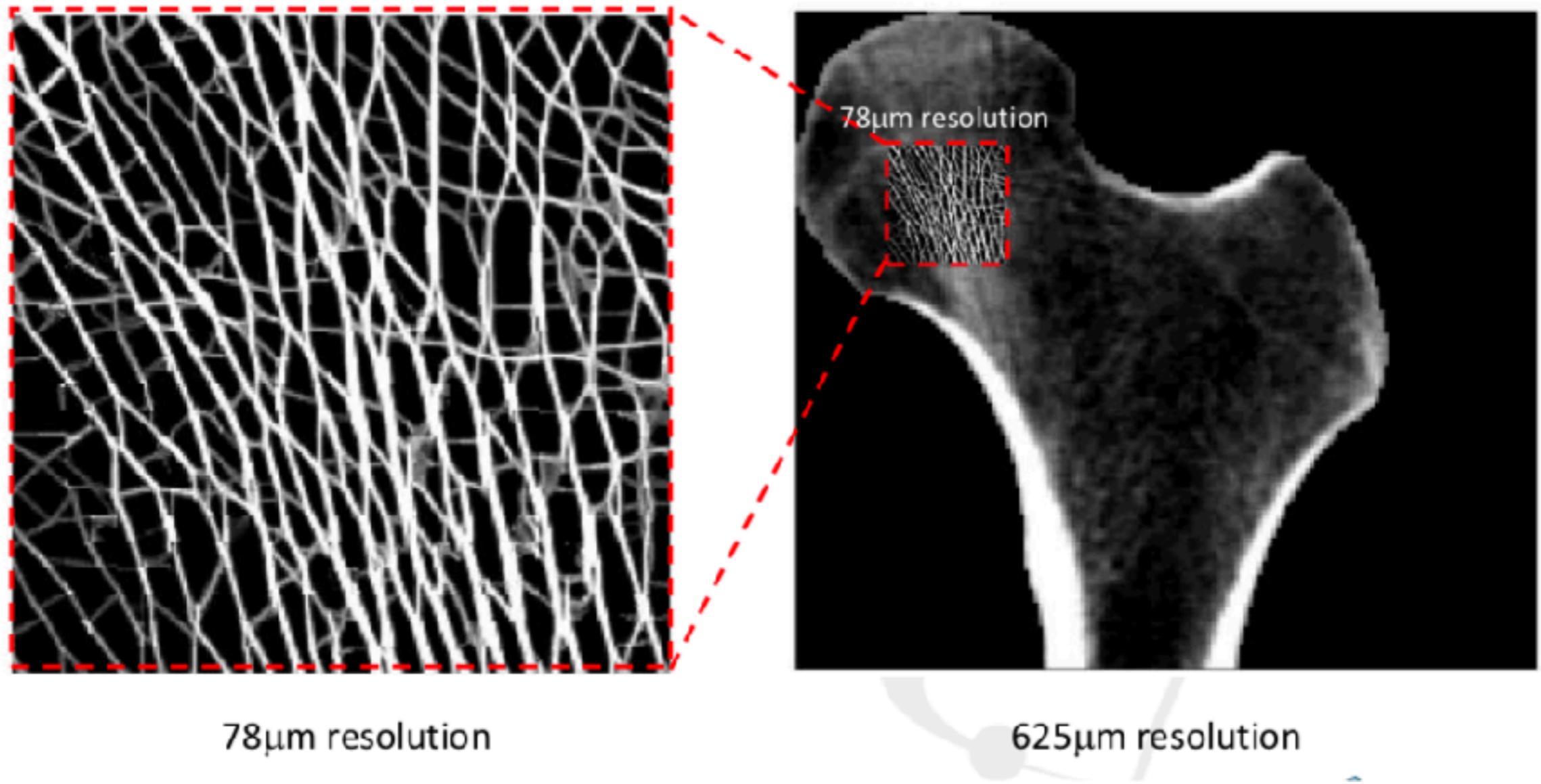


[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

한국 원자력 연구원



한국 원자력 연구원



농수산

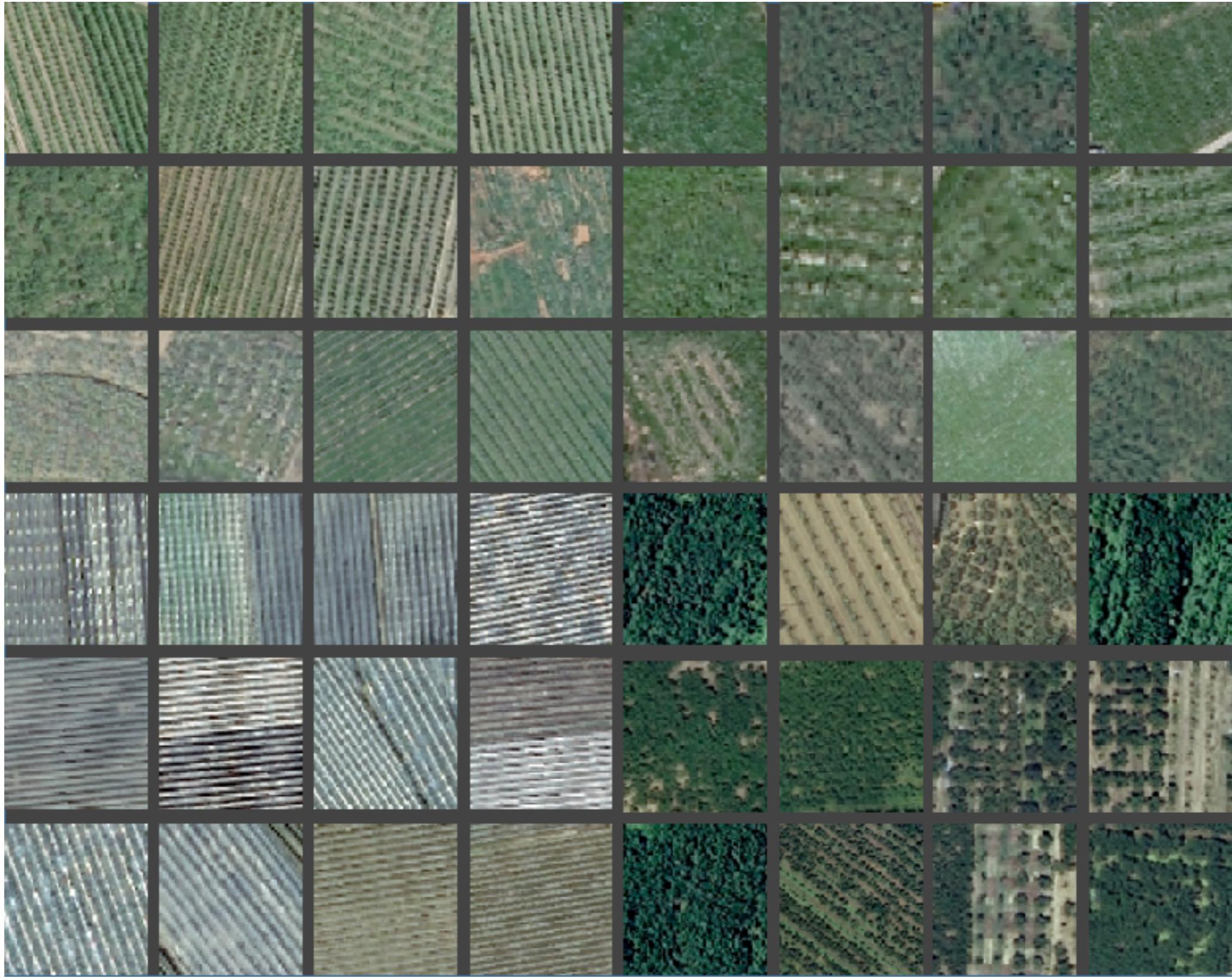
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

경작물 분류



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

사
과



포
도

배
복
숭
아

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

벼

옥수수

고구마

과수

비닐
하우스

고추

깻잎



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

[블록과 함께하는 파이썬 딥러닝 캐스](#), (주)인스페이스 기술이사 김태영

고객 서비스

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

맥도날드

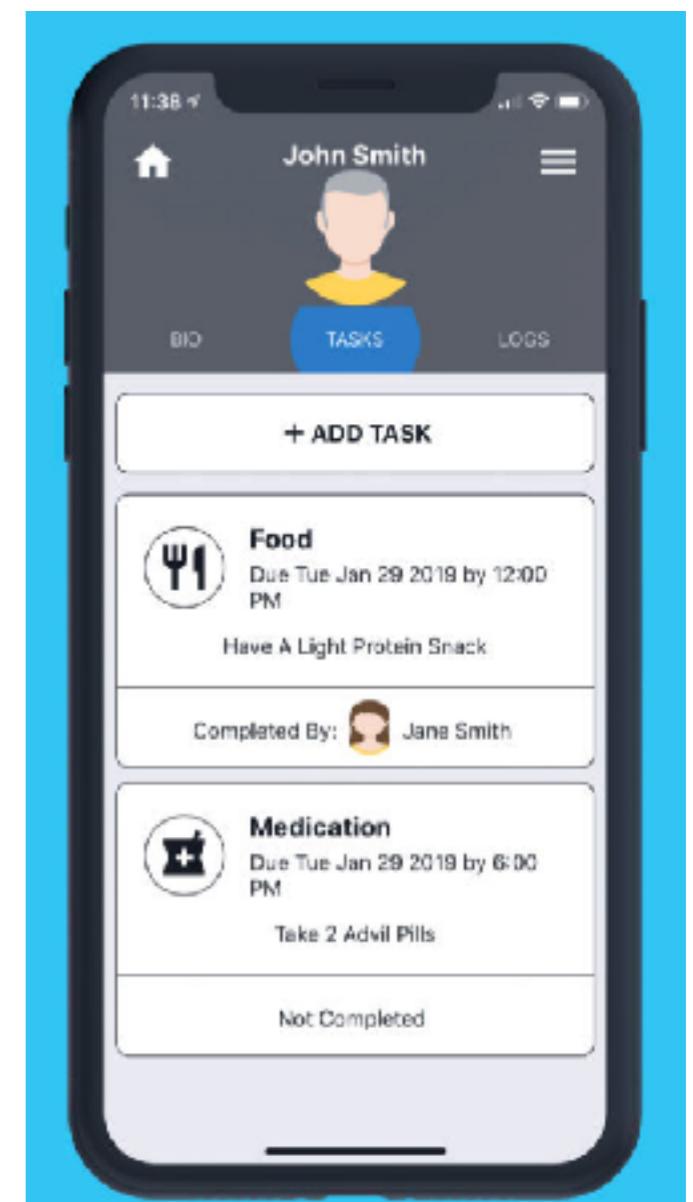


[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



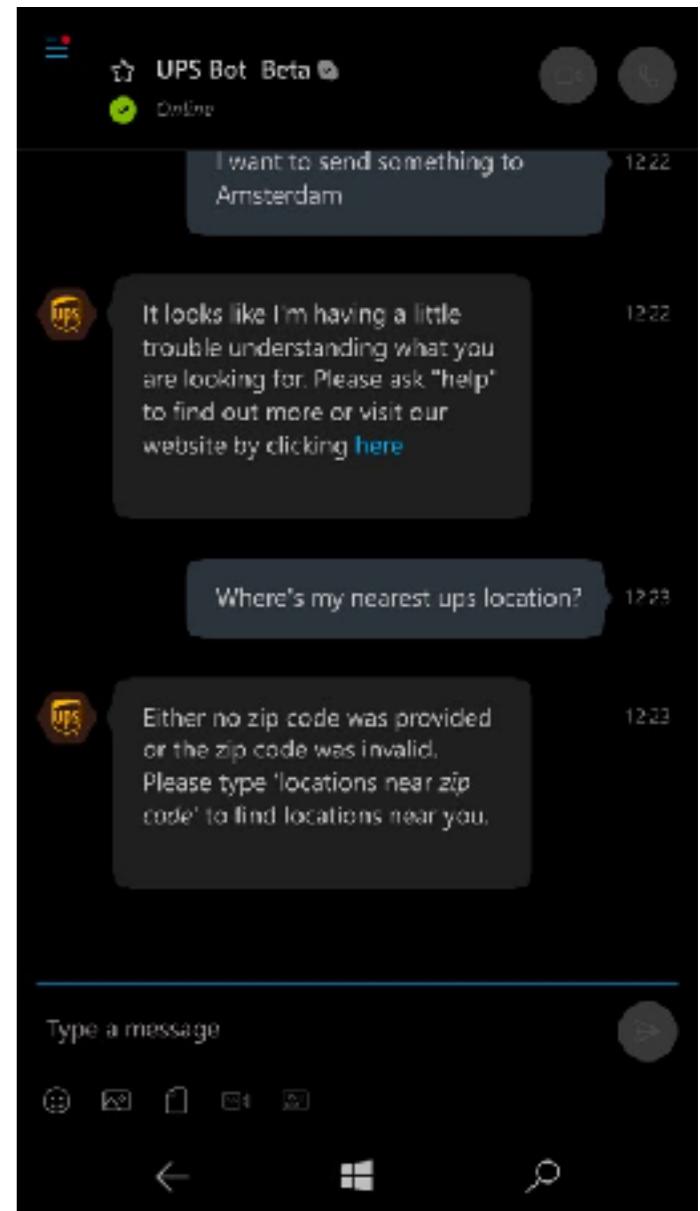
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

어린이 병원



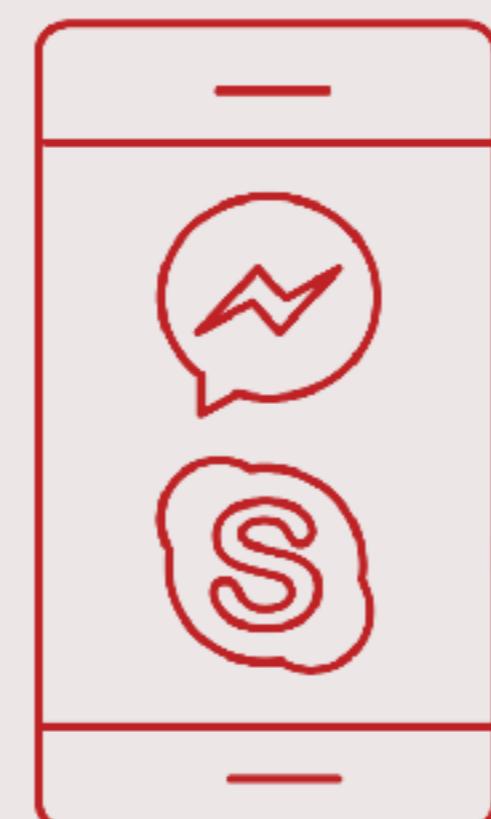
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

UPS



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

영국 철도



- Train Status**
Live arrivals and departures
- Journey Planner**
Where you need to be and by when
- Platform allocation**
When a platform is assigned
- Daily commute preferences**
Set your home and workplace
- Proactive delay and disruption**
Message notification alerts
- Alternative routes**
While you plan your journey

환경

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

동물 보호



멸종 위기에 놓인 코뿔소



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

동물 보호



영양, 얼룩말 센서 달린 목걸이 : 위치, 이동 방향, 속도 전송

제설 작업



<https://customers.microsoft.com/en-us/story/alaskadotpf-government-azure-iot>

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

식수 확보



UNIVERSITY OF
OXFORD



<https://customers.microsoft.com/en-us/story/reach>

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

토지



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

수중 생태계



<https://partner.microsoft.com/en-us/case-studies/Gramener>

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

의료

마케팅

보안

금융

스포츠

환경

예술

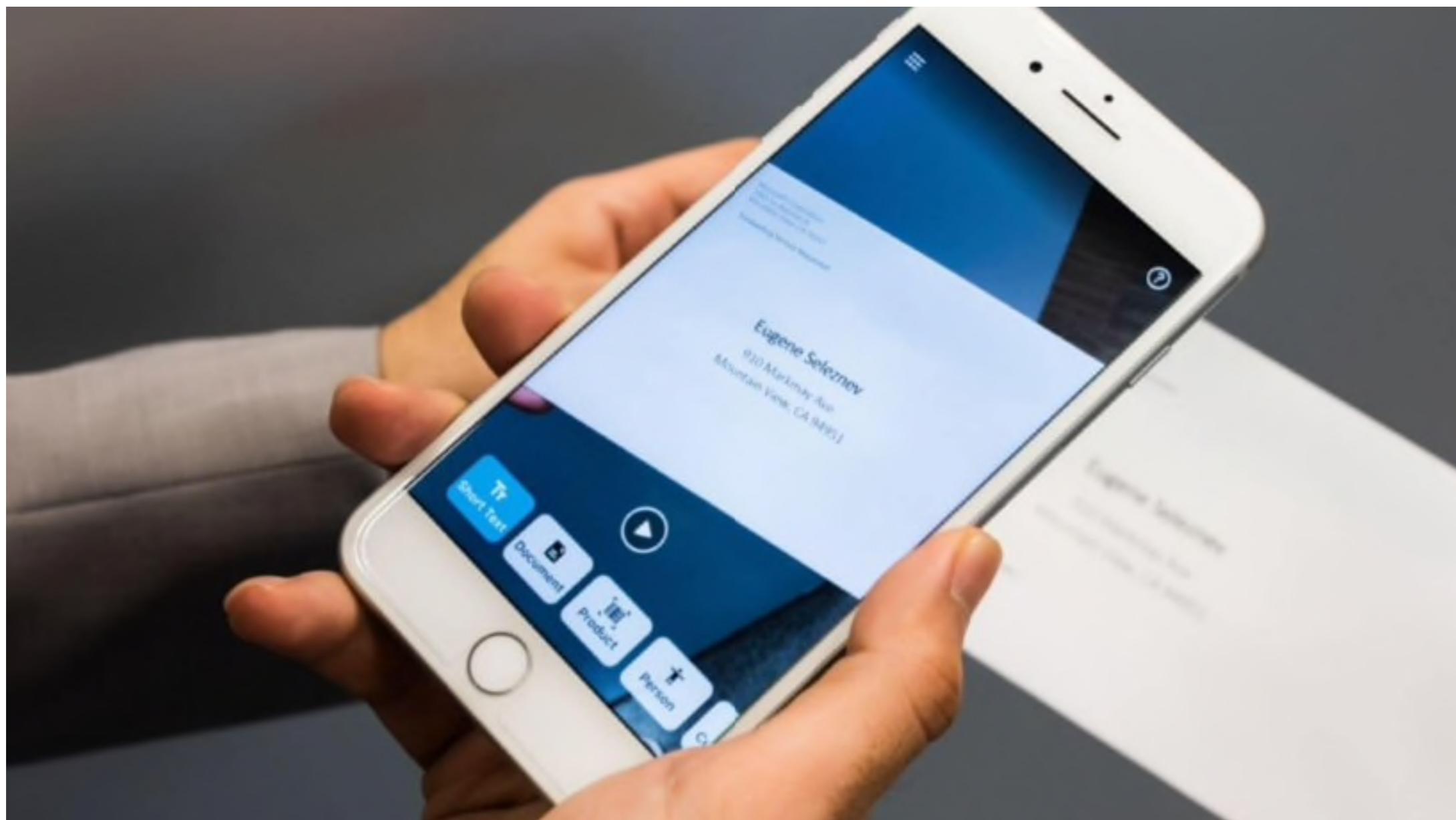
농수산

안전

법률

모바일

쇼핑



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

첫 번째 날

- 인공지능 개론
- 인공지능 사례
- 머신러닝 개론
- Classification 모델 생성

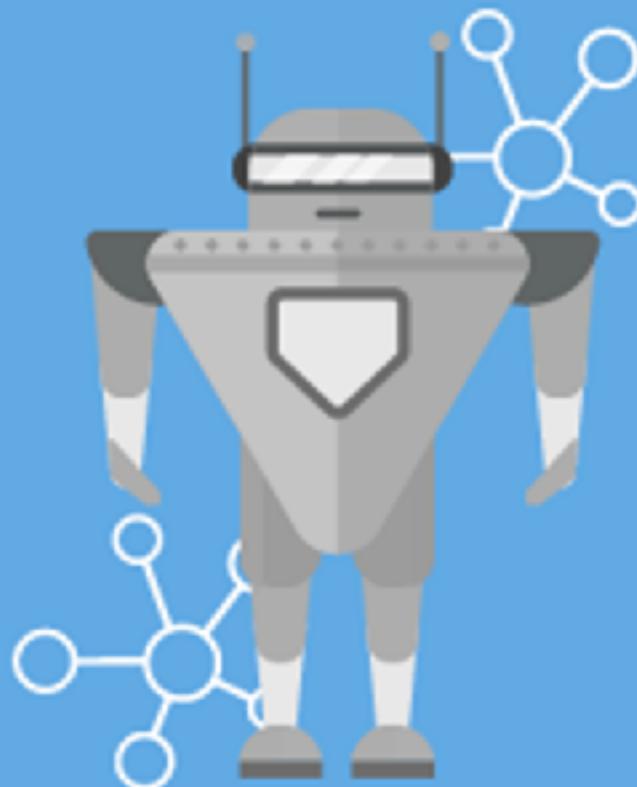
인공지능

머신러닝

딥러닝

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

인공지능



머신러닝



딥러닝



1950's 1960's 1970's

1980's 1990's 2000's 2010's

<http://bisintek.com/science/2017/12/27/knowing-basic-artificial-intelligence/>

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

인공지능

기계 혹은 컴퓨터가 인간의 지능을 모방해 인간과 비슷하게 동작하도록 만들어진 모든 기술
예) 딥블루, 로봇 청소기, 인공지능 에어컨, 자율주행 자동차

머신러닝

인공지능의 한 분야.

컴퓨터가 데이터를 이용해 학습하는 알고리즘 기술
인공신경망, 결정 트리, 서포트 벡터 머신 등 다양한 방법론

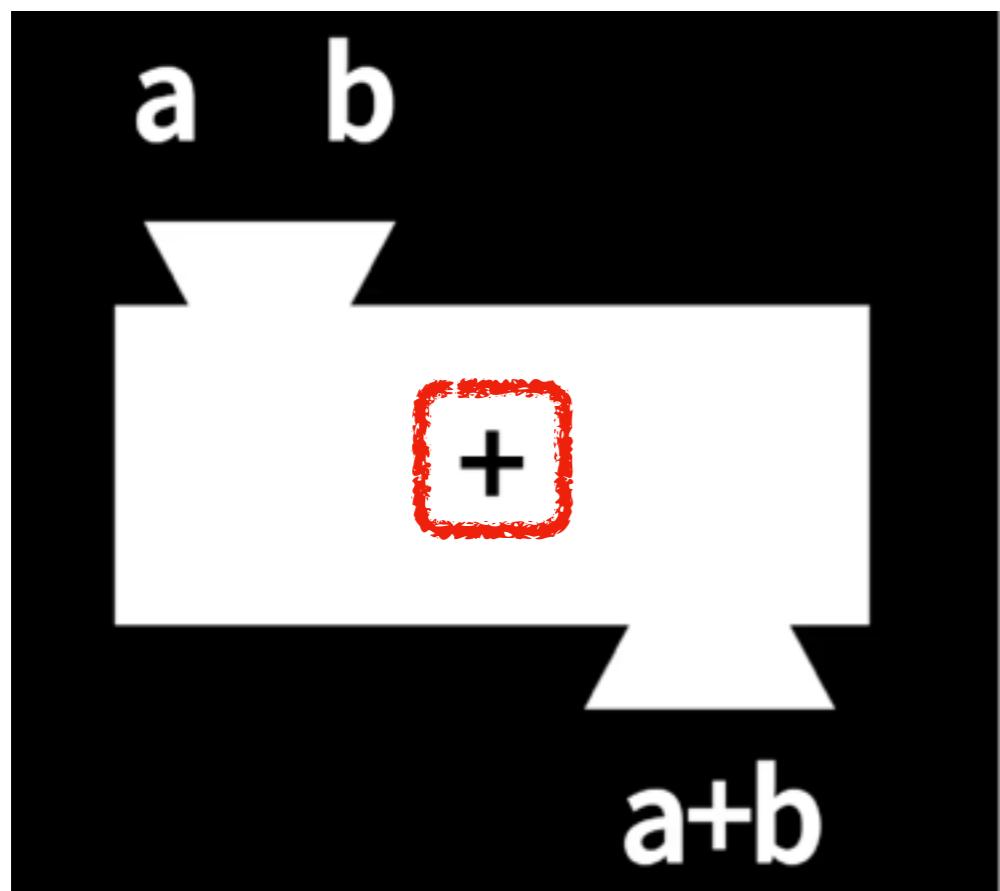
딥러닝

인공신경망을 사용하는 머신러닝 모델링 방법 중 하나
다층 인공신경망 구조를 사용하여 빅 데이터 학습
예) 사물인식, 감정분석, 필기체 인식, 음성인식

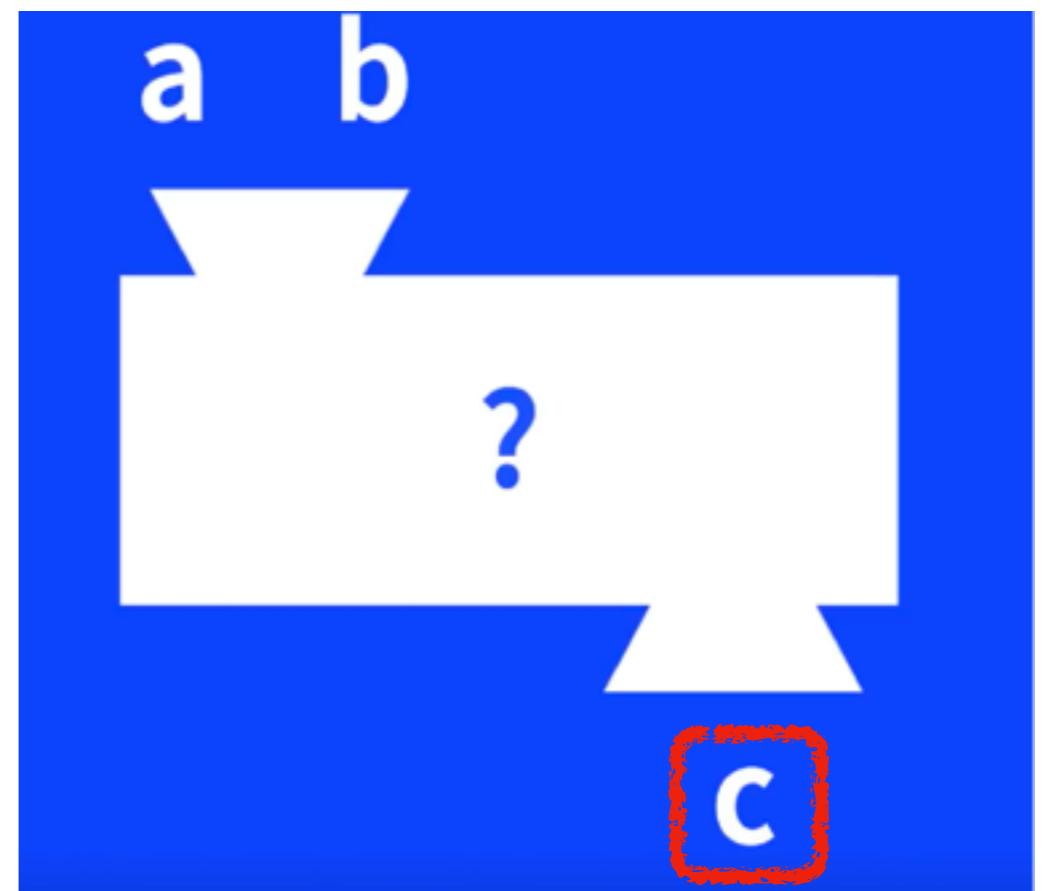
Machine Learning

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

일반적인 컴퓨터 사이언스

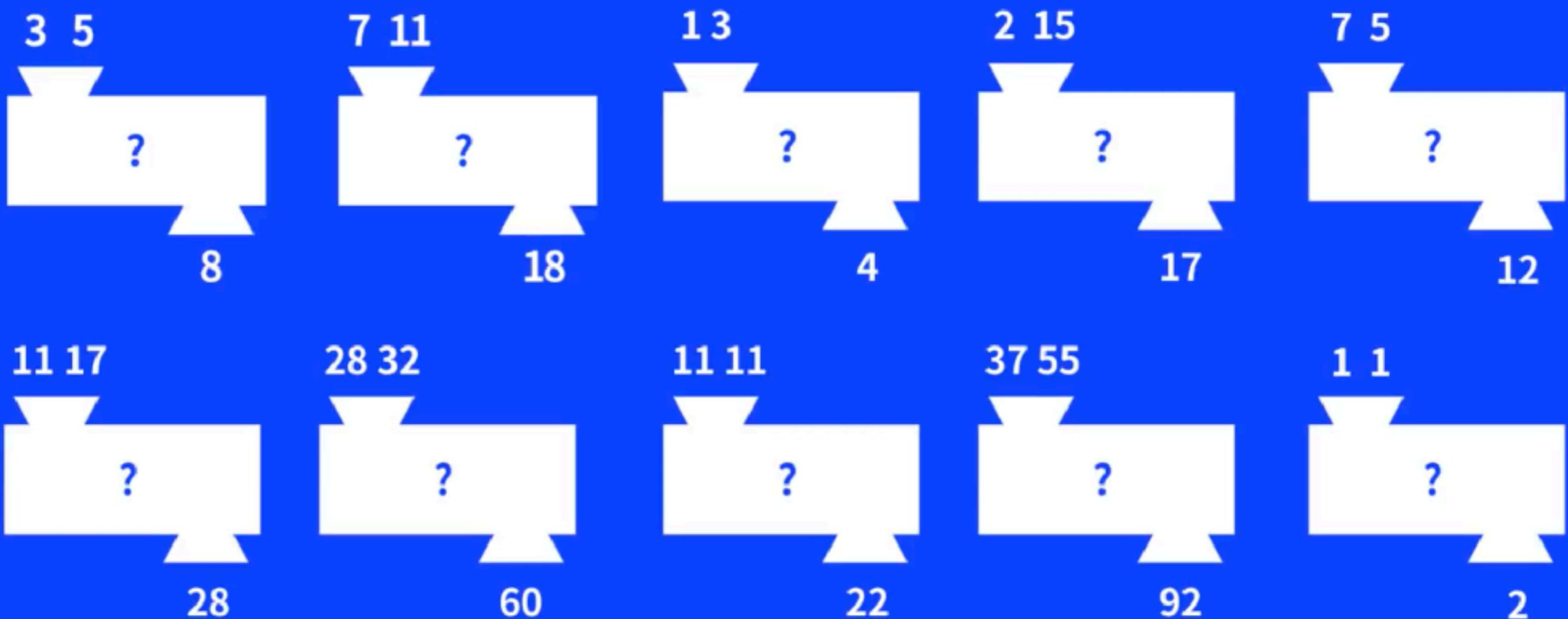


머신러닝



<https://youtu.be/aF03asAmQbY>

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

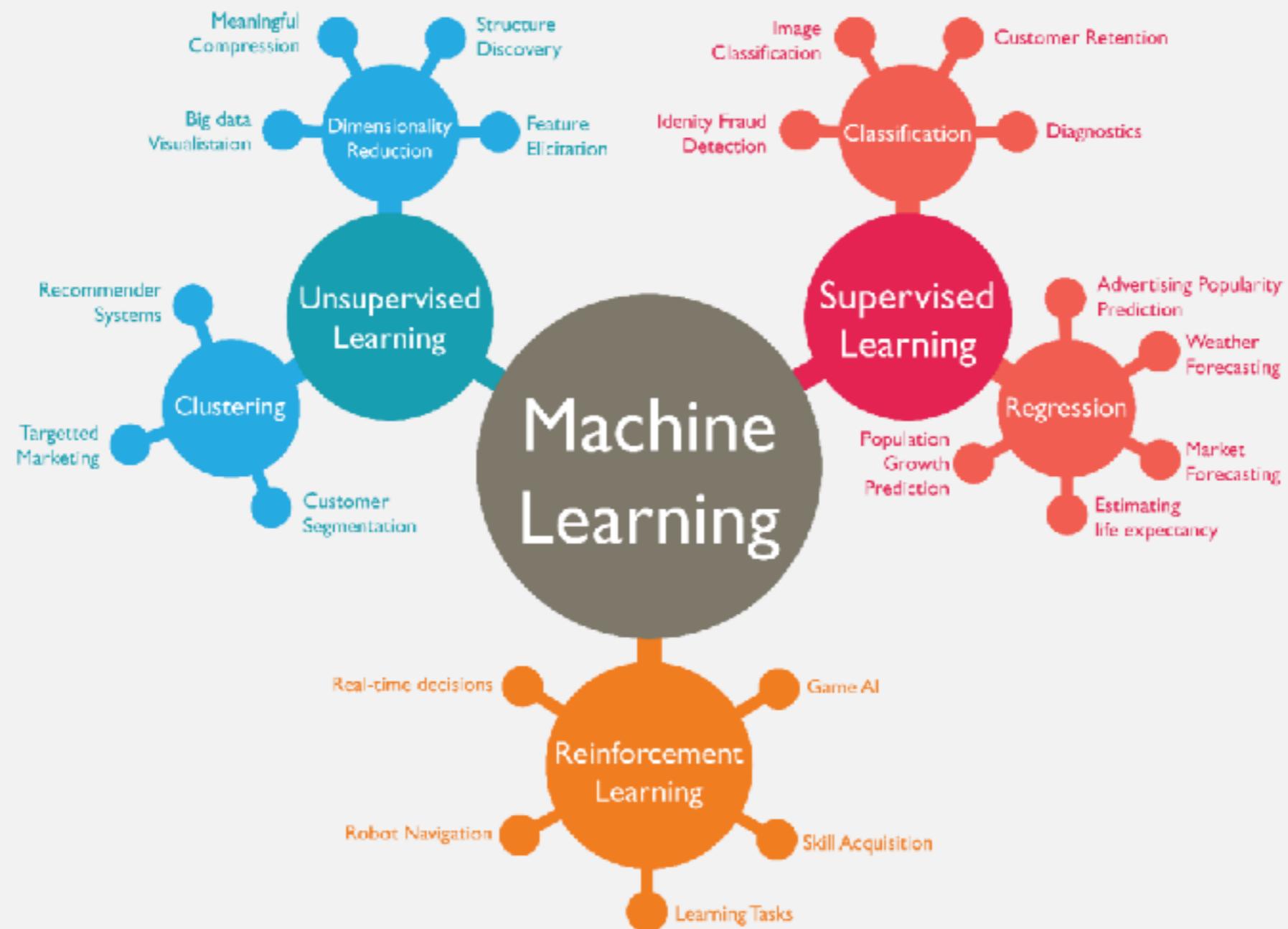


<https://youtu.be/aF03asAmQbY>

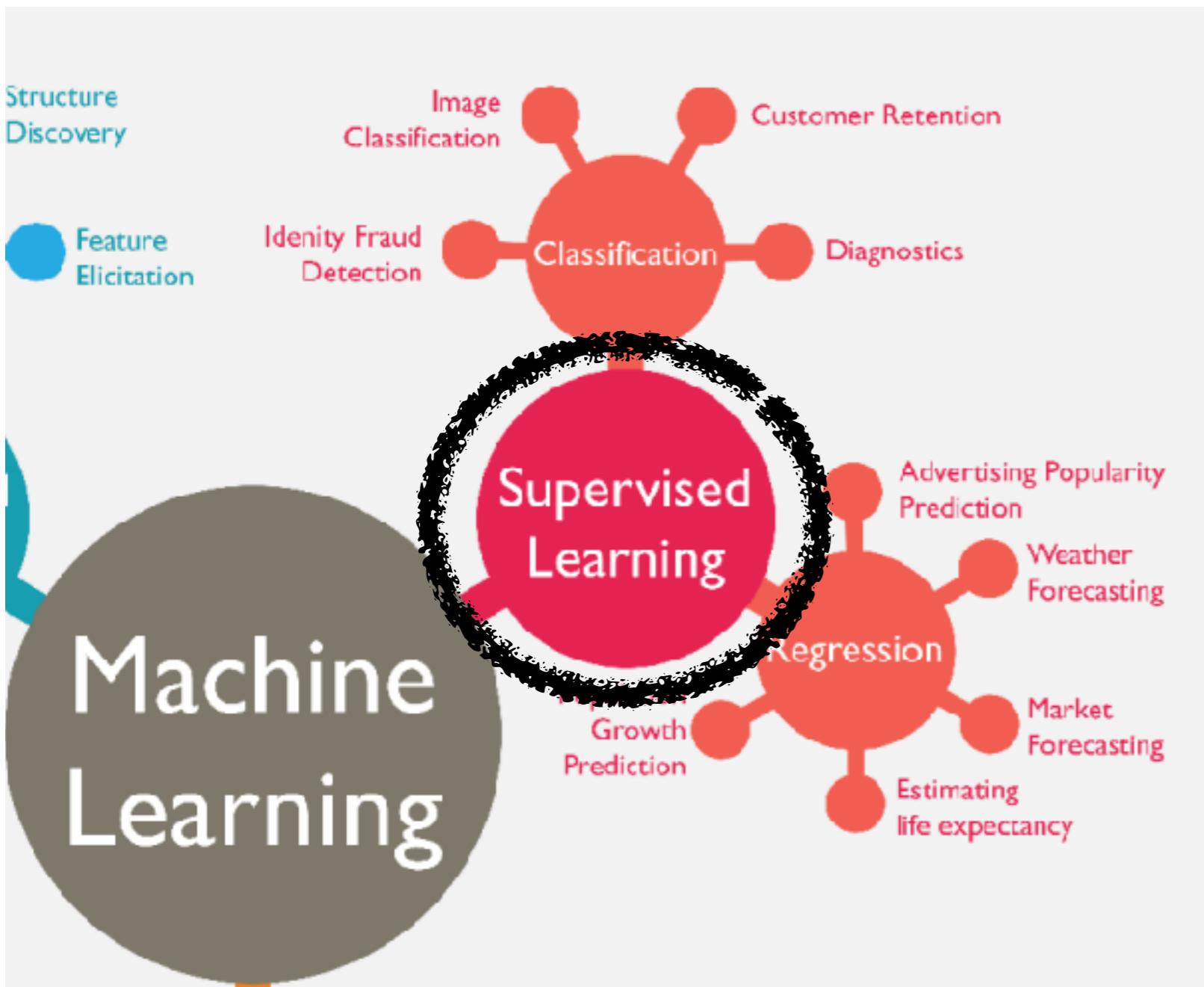
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

Machine Learning 종류

Machine Learning Bubble Chart

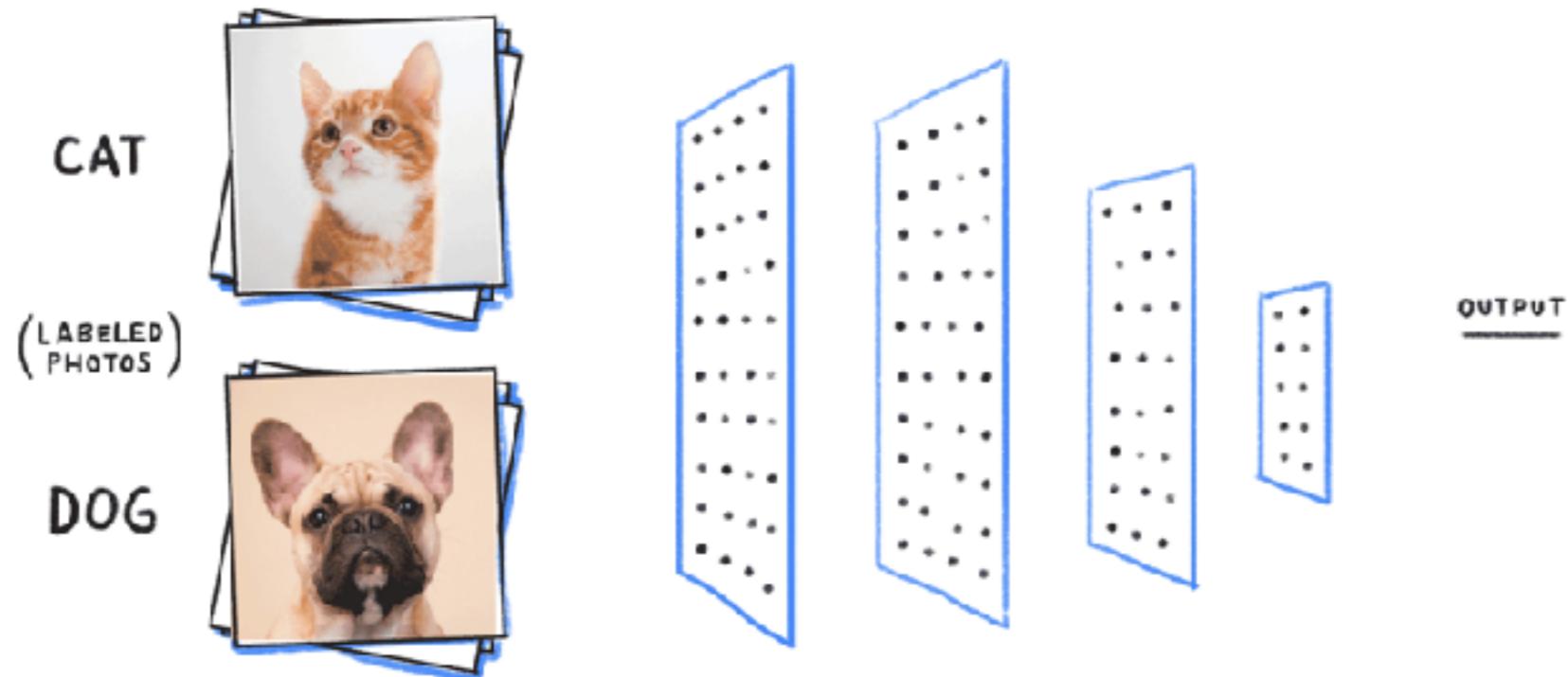


Supervised Learning (지도학습, 감독학습)

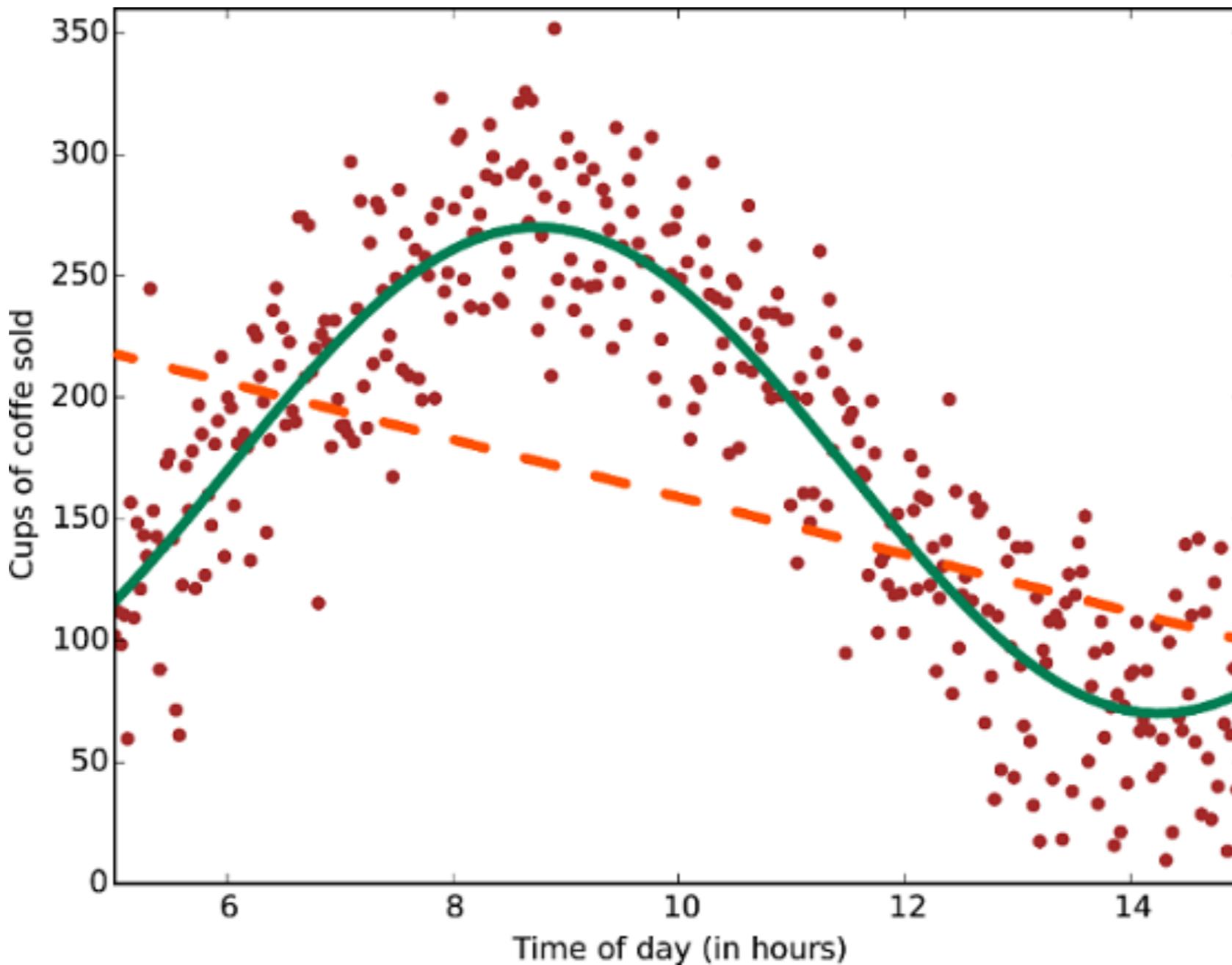


- 문제와 정답 제공
- 예측, 추정, 분류

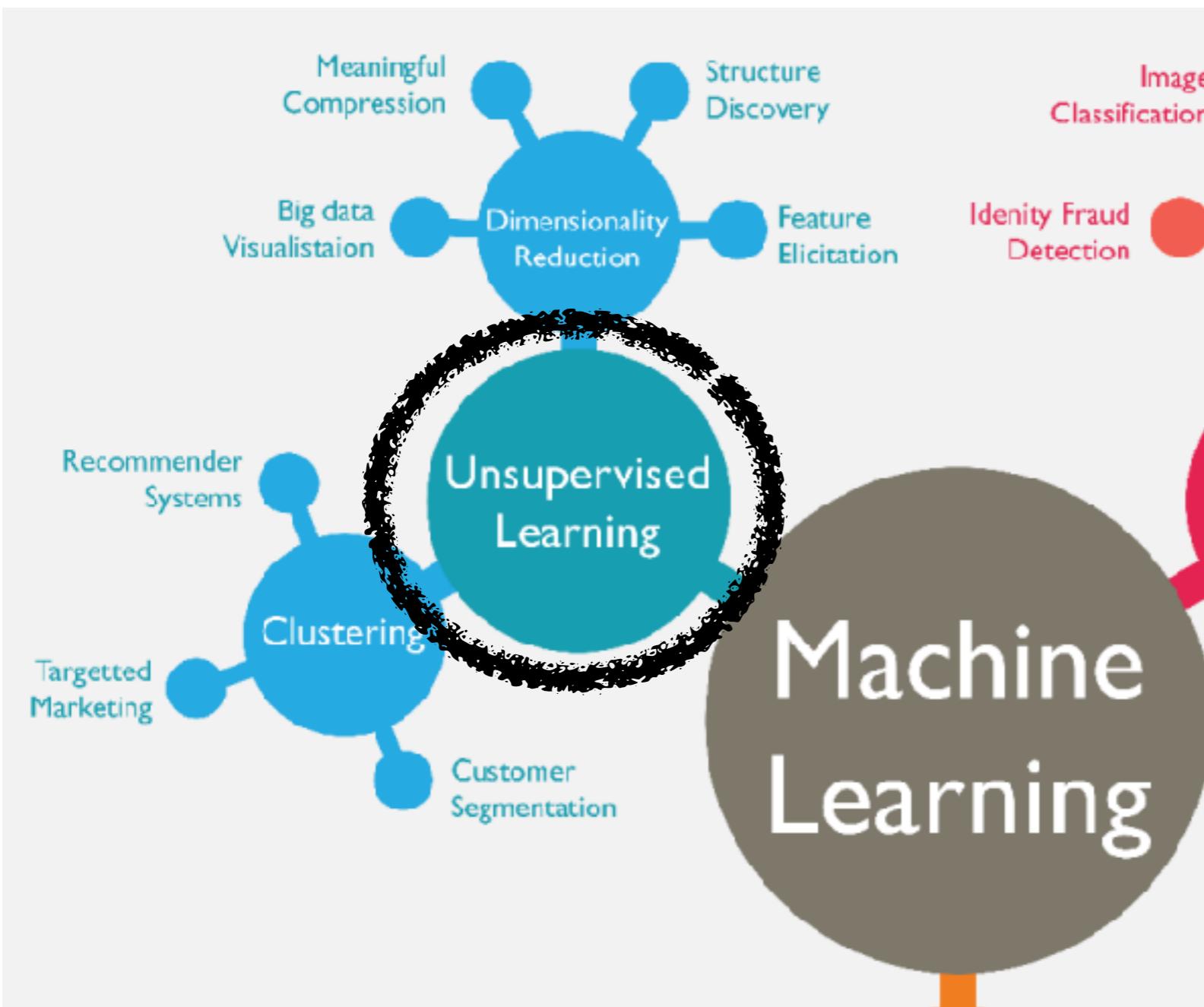
Classification



Regression

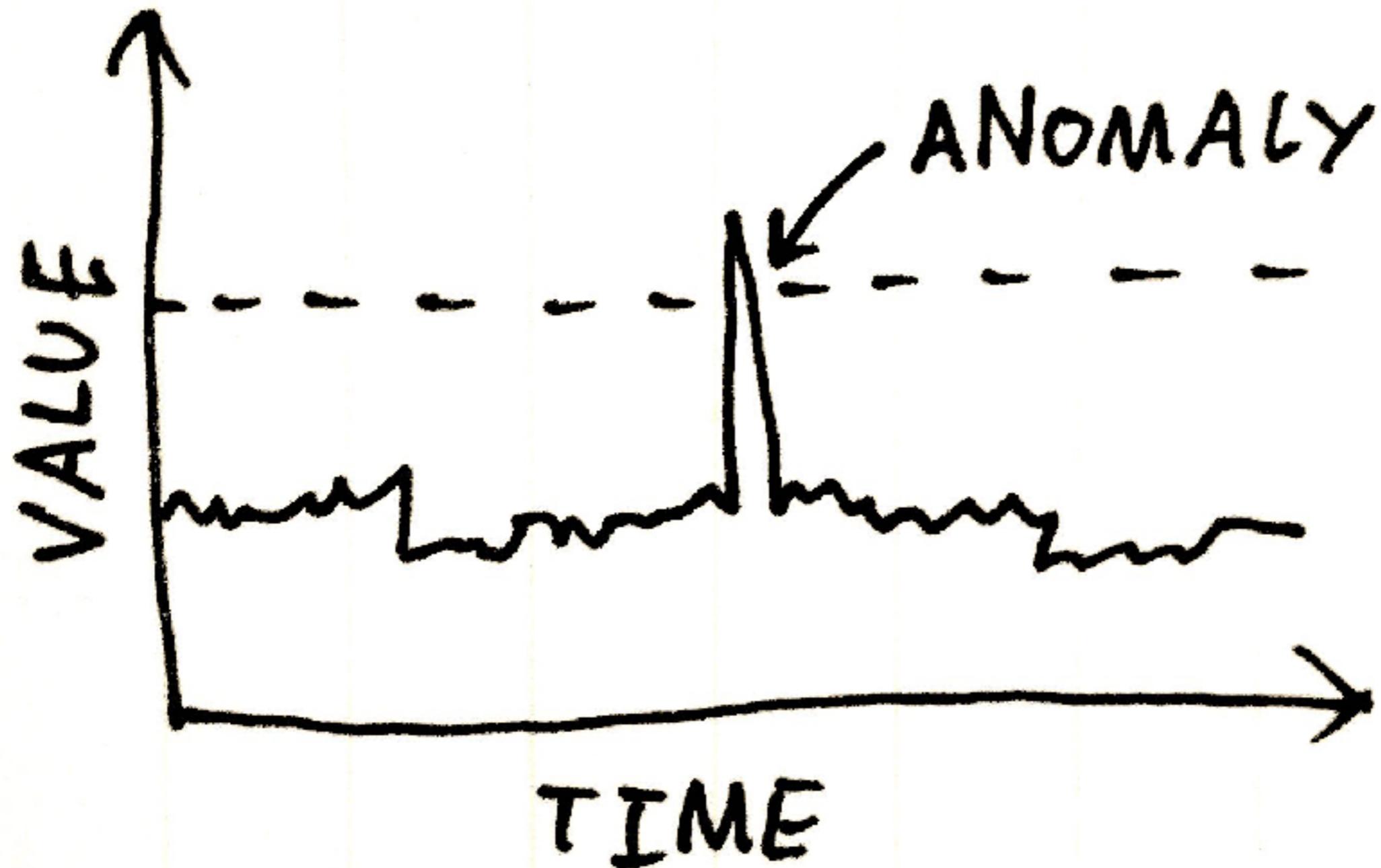


Unsupervised Learning (비지도학습)



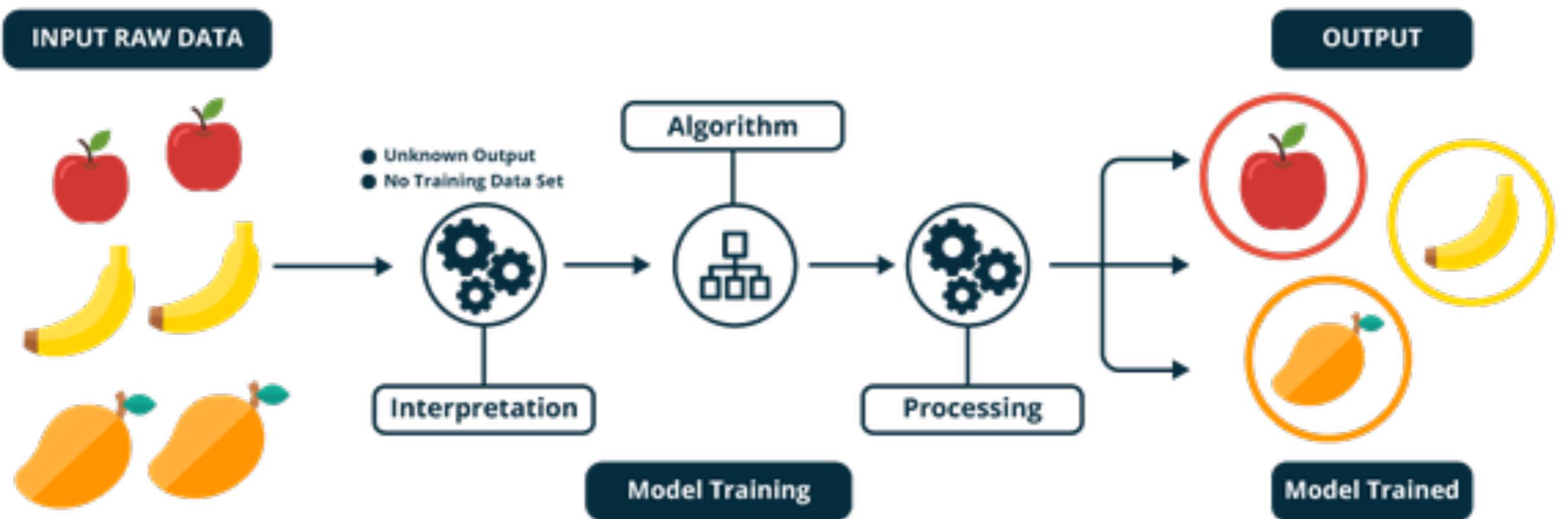
- 문제만 제공
- 패턴/구조 발견
- 그룹화

Anomaly Detection

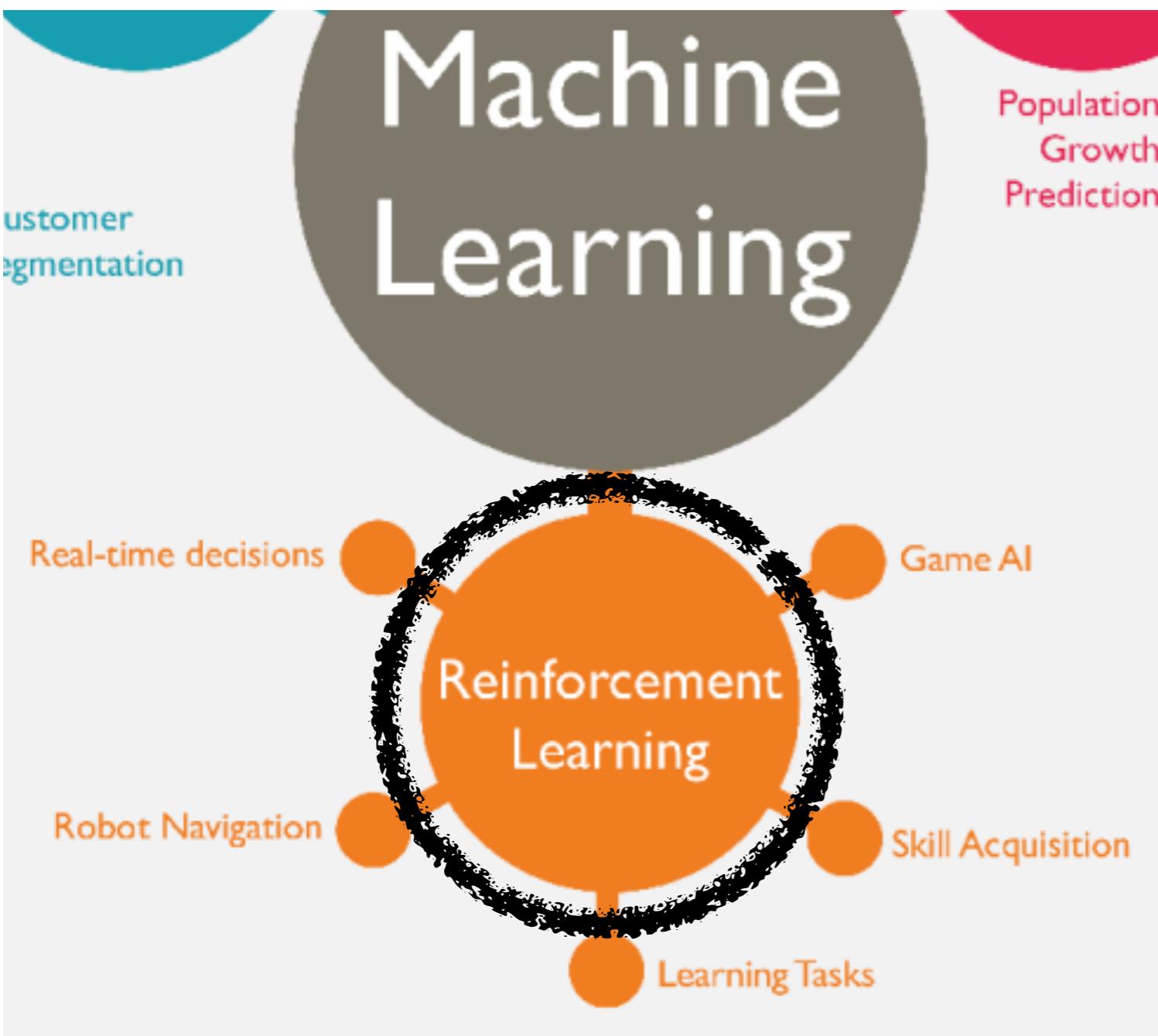


<https://towardsdatascience.com/fake-news-classification-via-anomaly-detection-765c4c71d539>

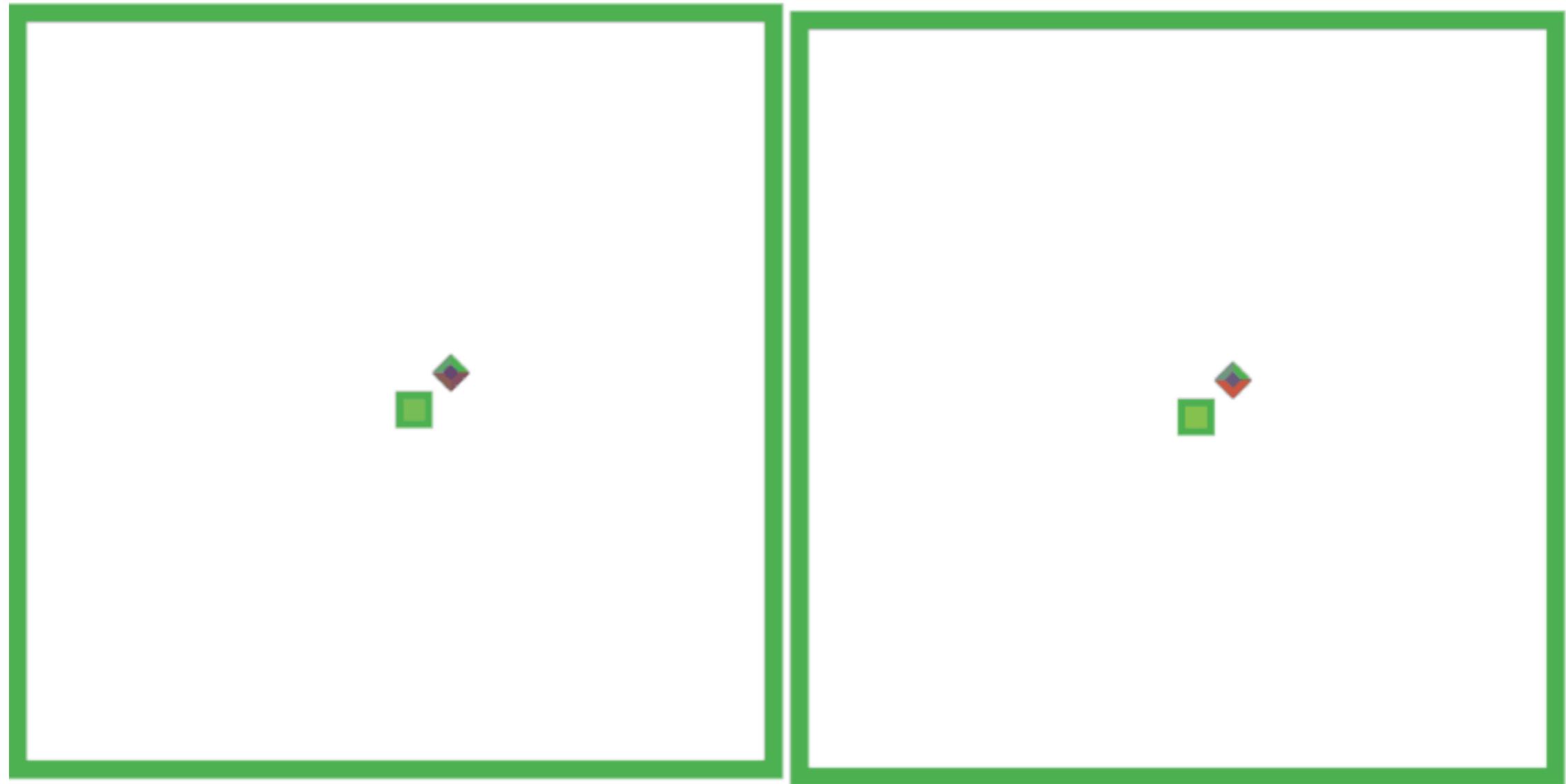
Clustering



Reinforcement Learning (강화학습)



- 정답이 아닌 보상 제공
- 인과관계가 중요
- 게임(알파고), 로봇



<https://towardsdatascience.com/how-to-teach-an-ai-to-play-games-deep-reinforcement-learning-28f9b920440a>

SFV: Reinforcement Learning of Physical Skills from Videos

(with audio)



Xue Bin Peng, Angjoo Kanazawa, Jitendra Malik,
Pieter Abbeel, Sergey Levine

UC Berkeley



<https://youtu.be/4Qg5I5vhX7Q>

Machine Learning 방법

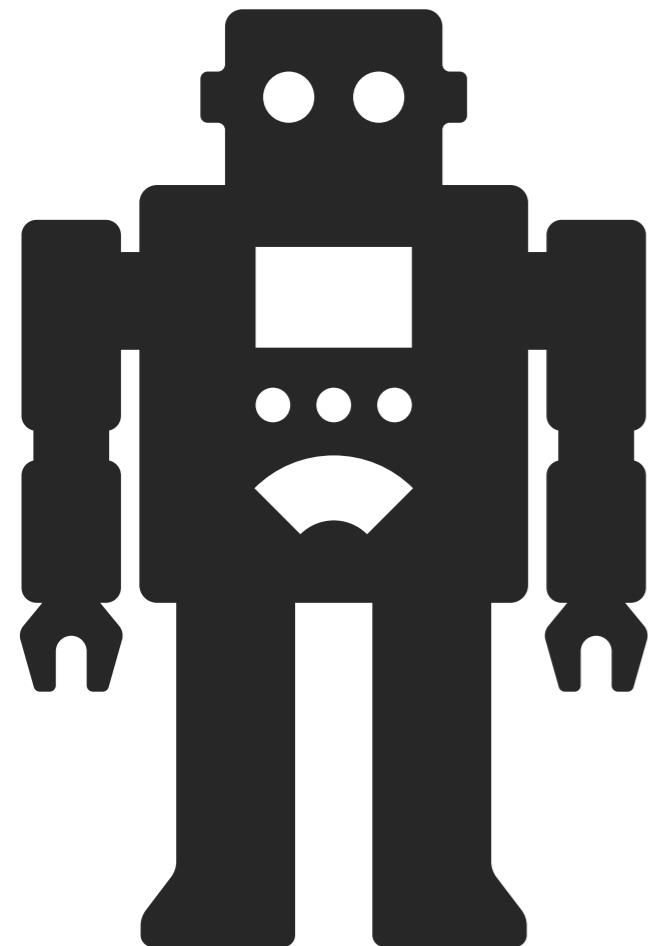
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

- Linear Regression
- Regression Trees
- SVM(Support Vector Machine)
- Decision Tree
- Ensemble
- Bayes Classification
- Logistic Regression
- SOM(Self-Organizing Map)
- k-Means Clustering
- Sequence Analysis
- Link Analysis
- Text mining
- Neural Network
-

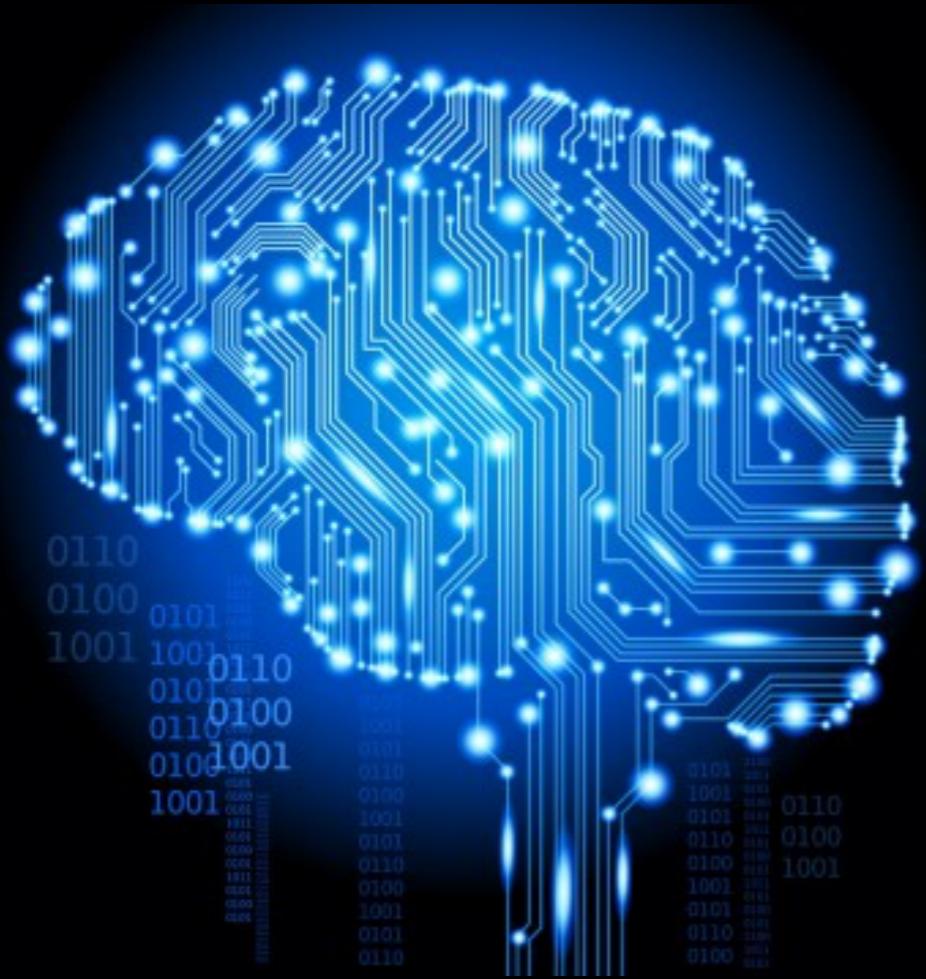
Deep Learning

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

- Linear Regression
- Regression Trees
- SVM(Support Vector Machine)
- Decision Tree
- Ensemble
- Bayes Classification
- Logistic Regression
- SOM(Self-Organizing Map)
- k-Means Clustering
- Sequence Analysis
- Link Analysis
- Text mining
- Neural Network
-

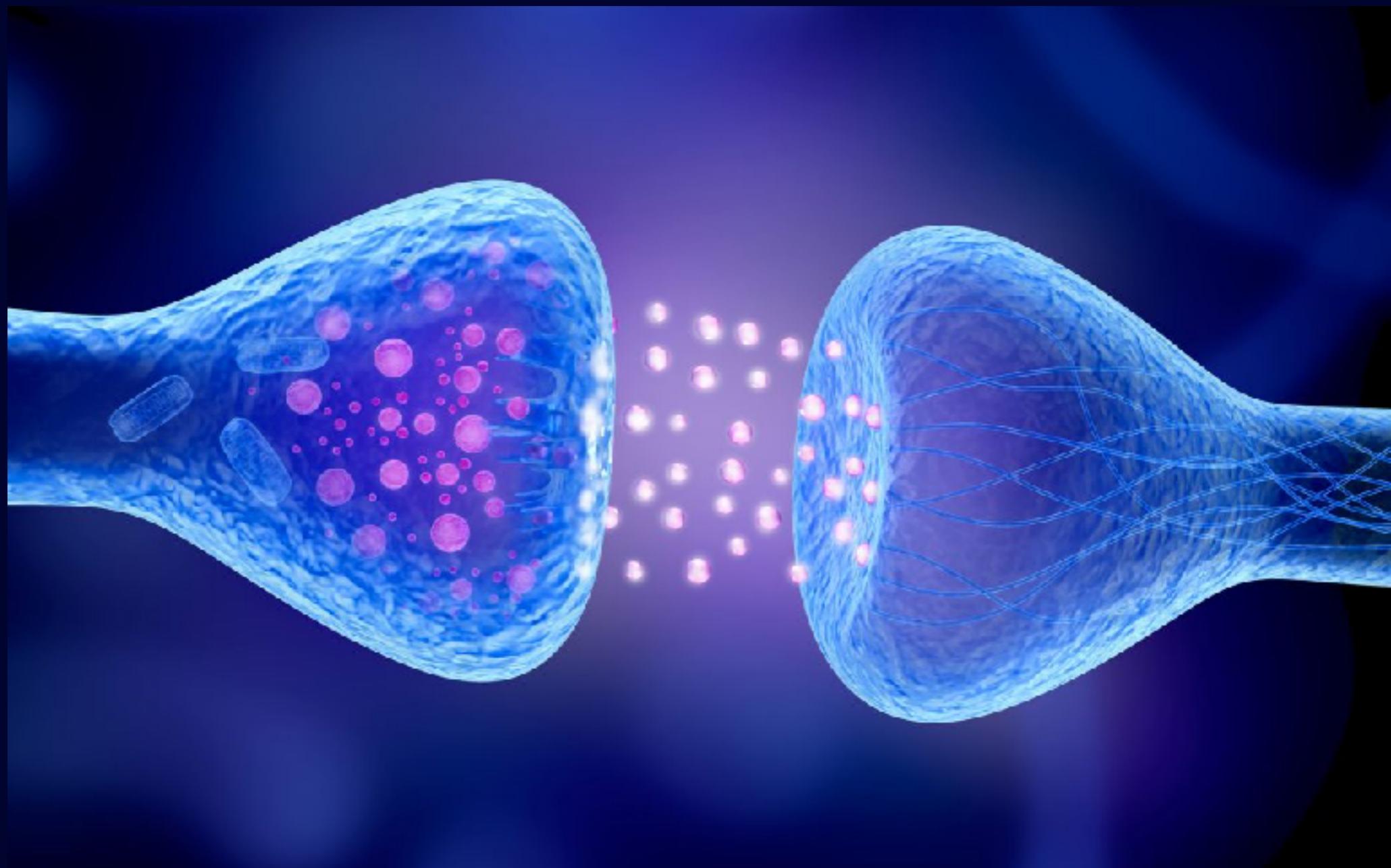


기계를 인간과 비슷하게 동작하게 하는 기술



<https://www.alzheimersreadingroom.com/2016/07/alzheimer-tau-protein-spreads-through-brain.html>

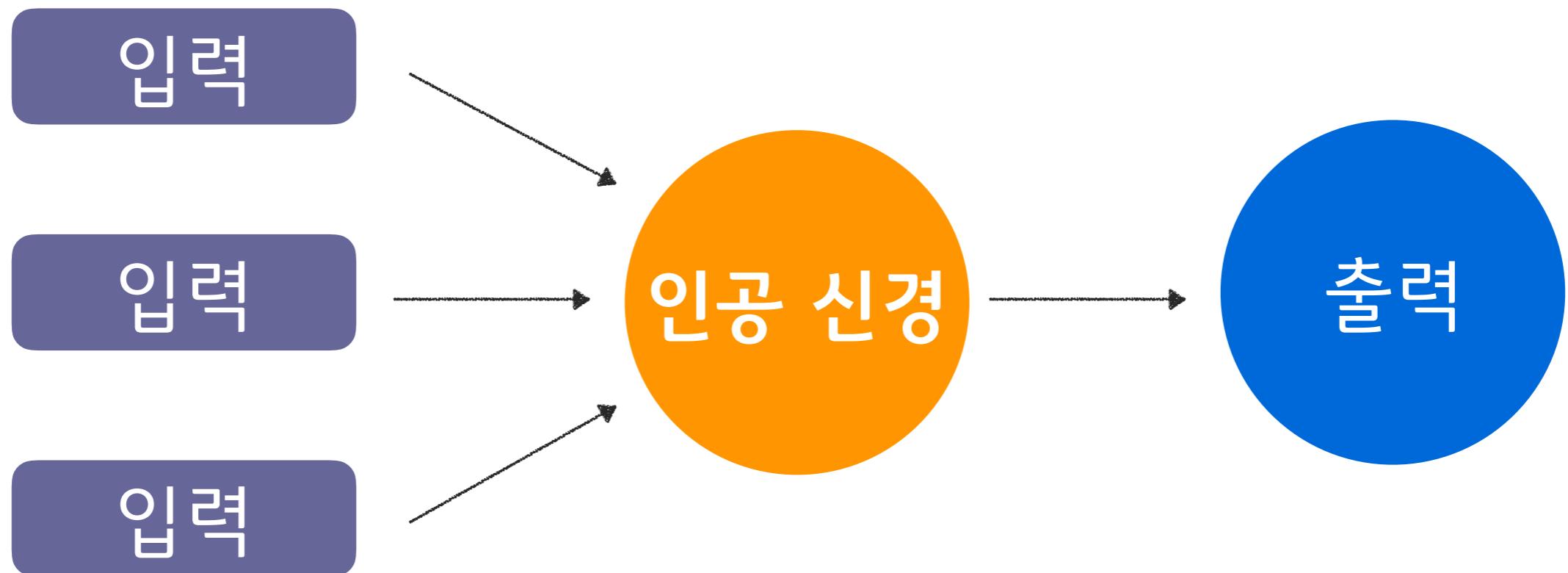
<https://www.extremetech.com/wp-content/uploads/2015/12/Brain-Machine-2-640x356.jpg>



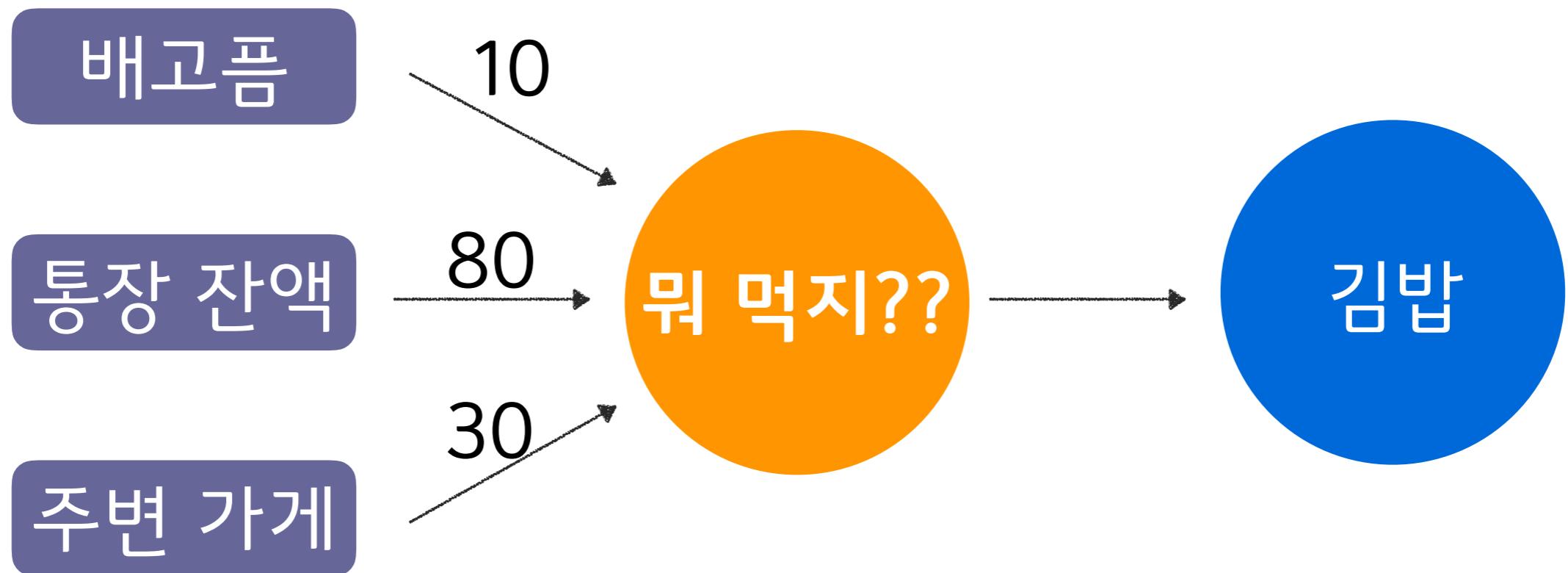
https://lh6.googleusercontent.com/4iAXI2yxBcbWVDXDaHOGGeAg-SDjcG0Akms48PpoocApztM9eXY6p9d-zPiztVlotO_SRl4Y9LZ2-dKCw-D_jl5DA_5pqYD6WPFo8NDMc0hPv8SiXYQ5bjx5ZMLBUghOVJv1zsleG

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

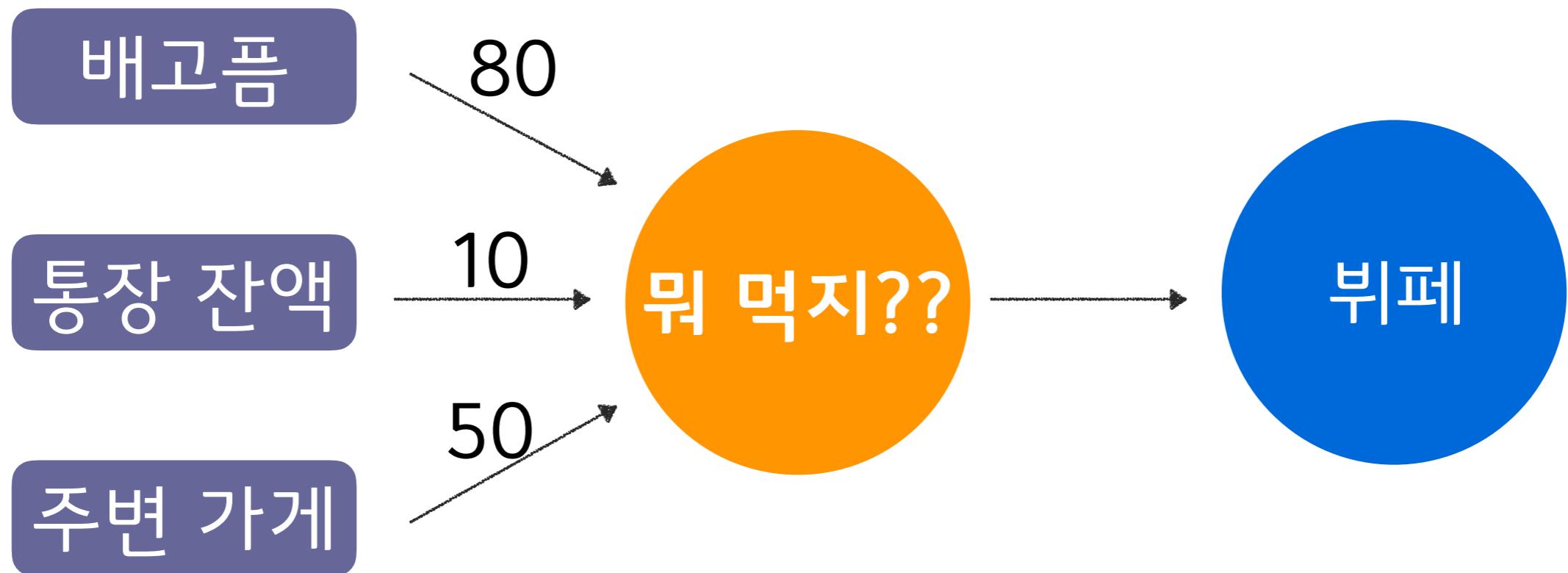
인공 신경



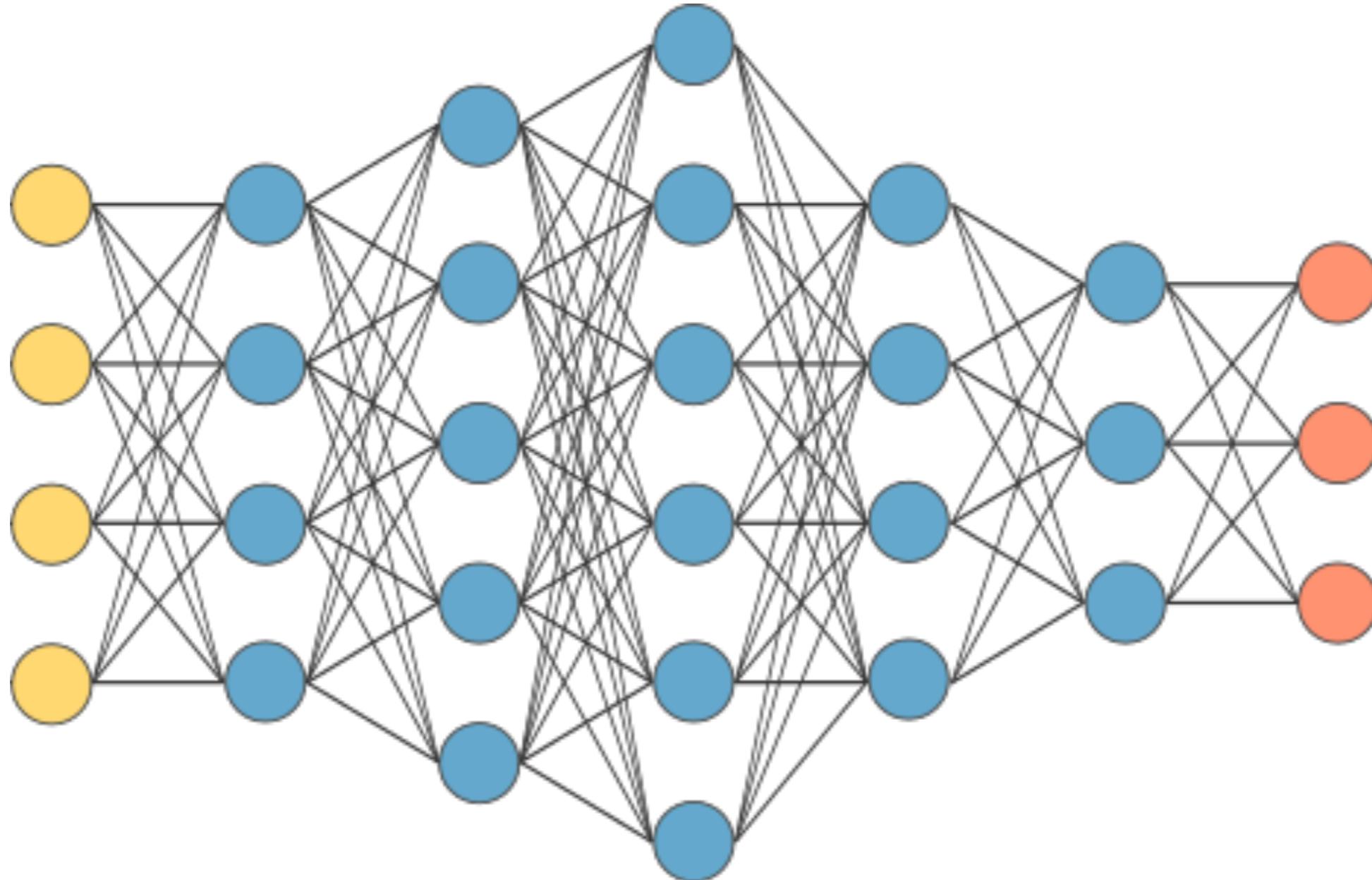
인공 신경



인공 신경

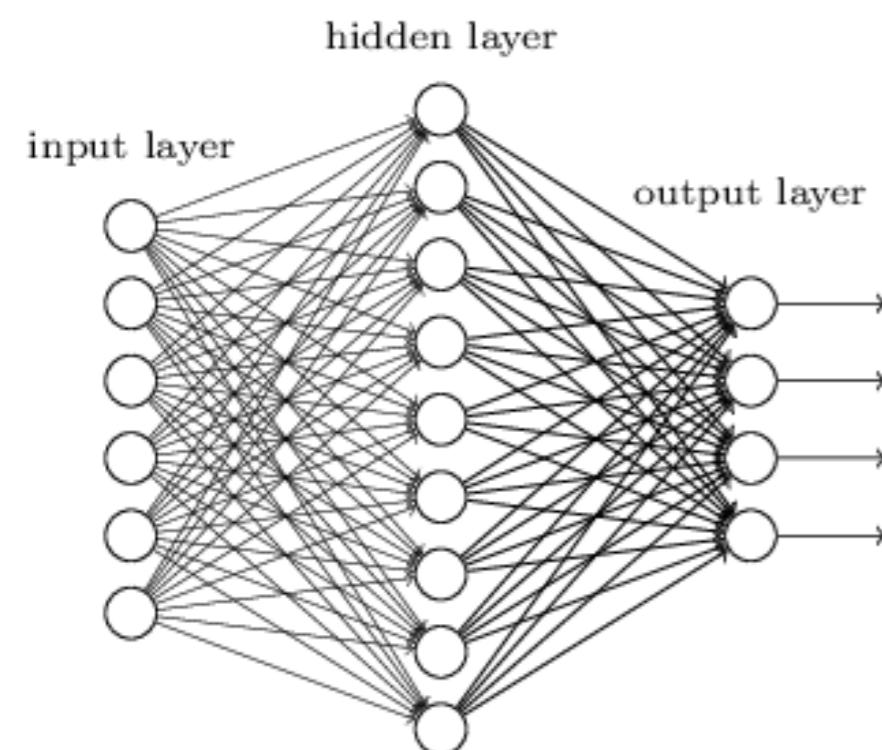


딥러닝

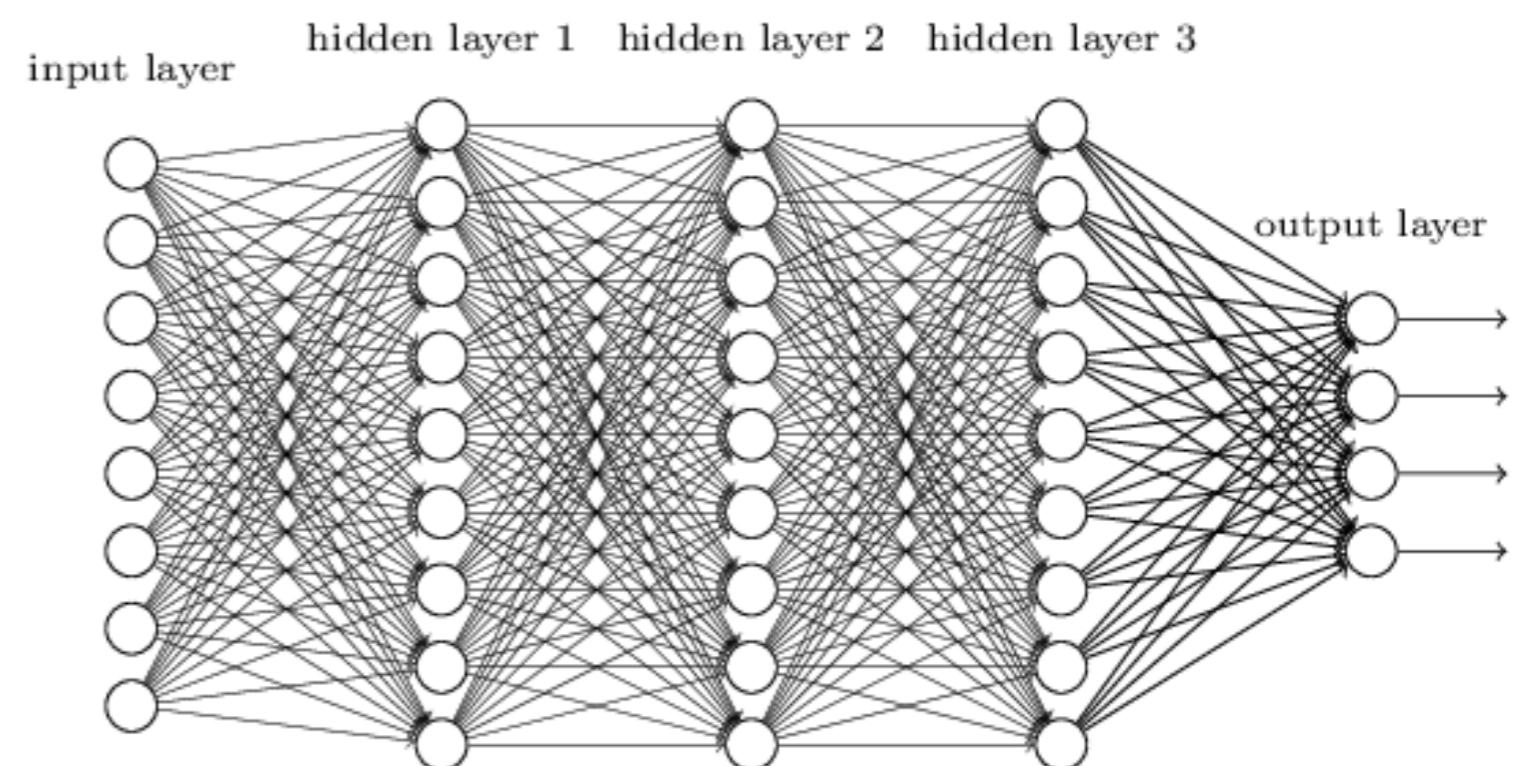


https://www.neuraldesigner.com/images/deep_neural_network_big.png

"Non-deep" feedforward neural network



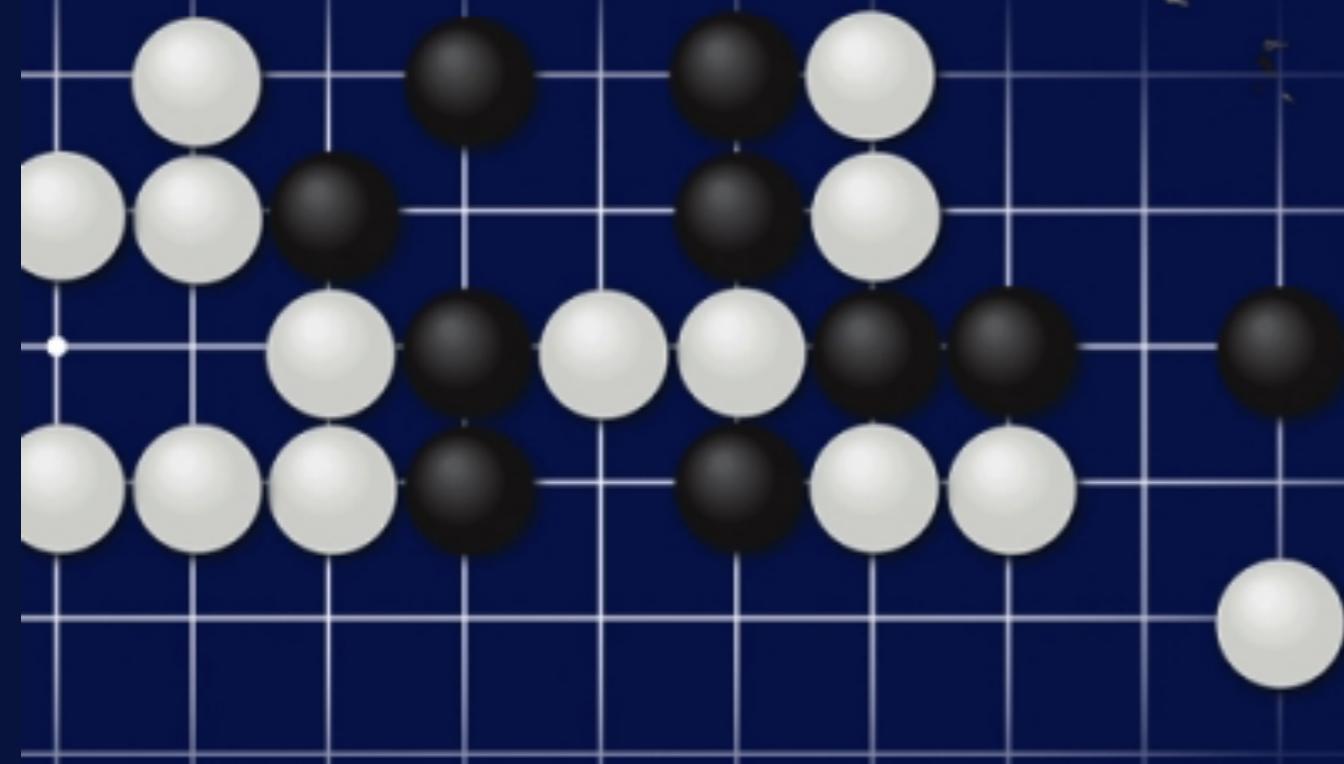
Deep neural network





AlphaGo

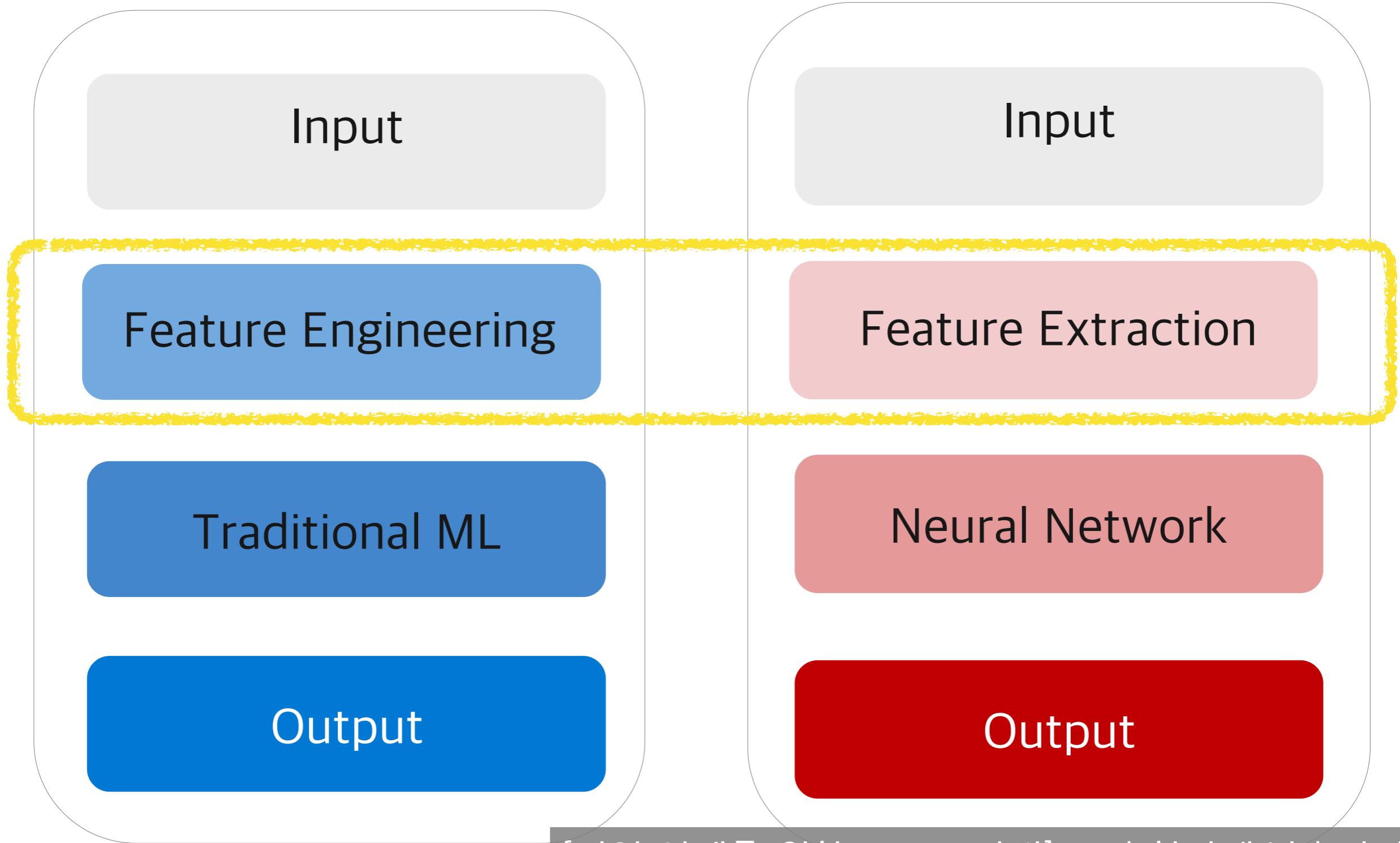
Deep Learning



????

머신러닝

딥러닝



Feature Engineering

이름	나이	성별	선실 등급	티켓넘버	티켓 요금	부모 자식	형제 자매	키	출항지	생존 여부
Rose	25	여	A	EA-1039	300	2명	1명	167	런던	Y
Jack	20	남	C	GE-3059	29	-	-	178	도버	N
Mark	57	남	B	BA-2031	89	4명	3명	167	뉴포트	N
Andy	48	남	B	NN-3928	102	5명	7명	182	런던	Y

Learning (학습)?

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



Sung Kim

게시일: 2016. 3. 21.

$$Y = w \times X + b$$

실제 값 $Y = w \times X + b$

예측 값 $\hat{Y} = \hat{w} \times \hat{X} + \hat{b}$

$$3 = w \times 1 + b$$

$$5 = w \times 2 + b$$

$$7 = w \times 3 + b$$

$$Y = \textcolor{red}{w} \times X + \textcolor{blue}{b}$$

$$3 = w \times 1 + b$$

$$5 = w \times 2 + b$$

$$7 = w \times 3 + b$$

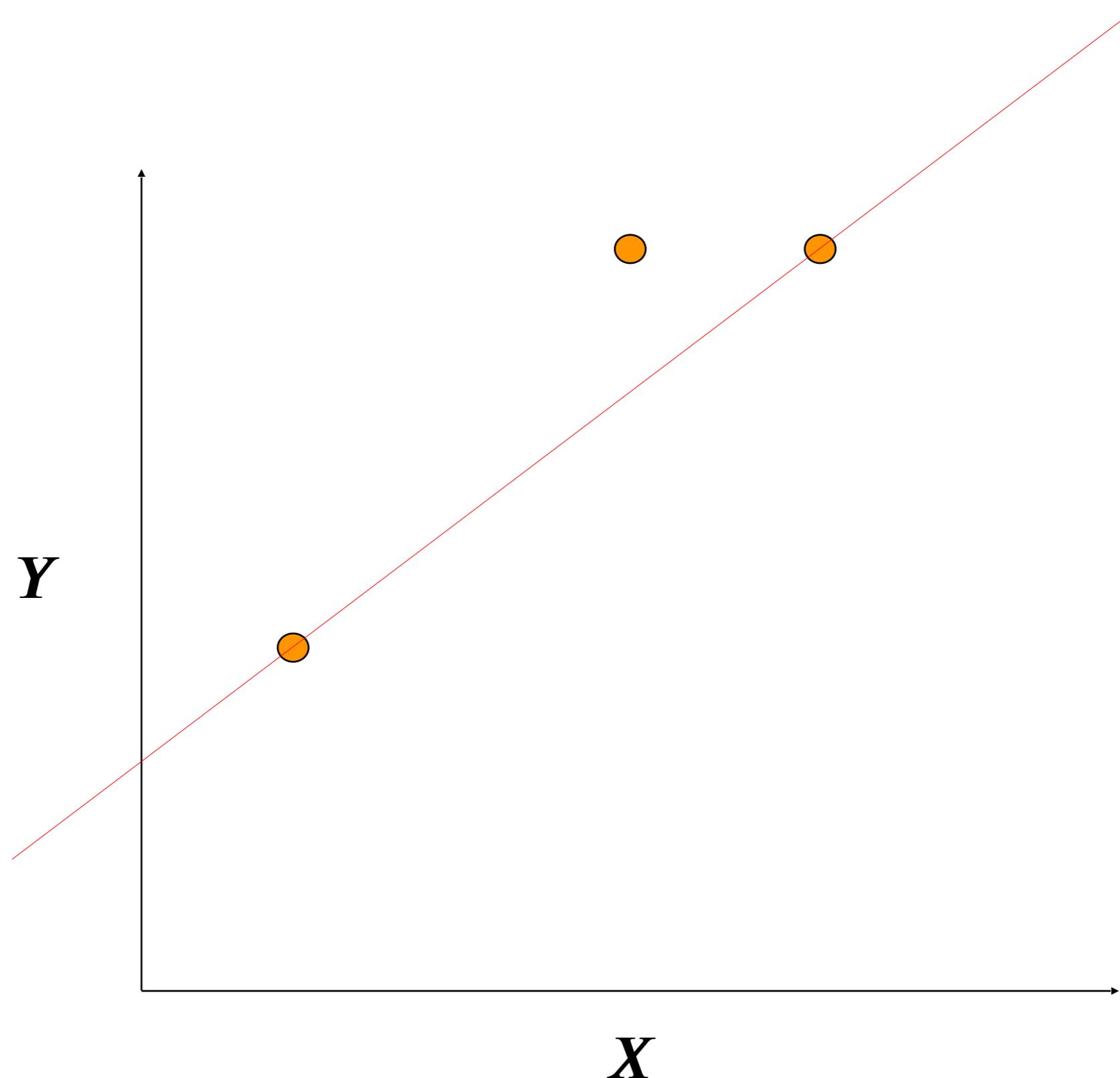
$$Y = 2 \times X + 1$$

$$3 = w \times 1 + b$$

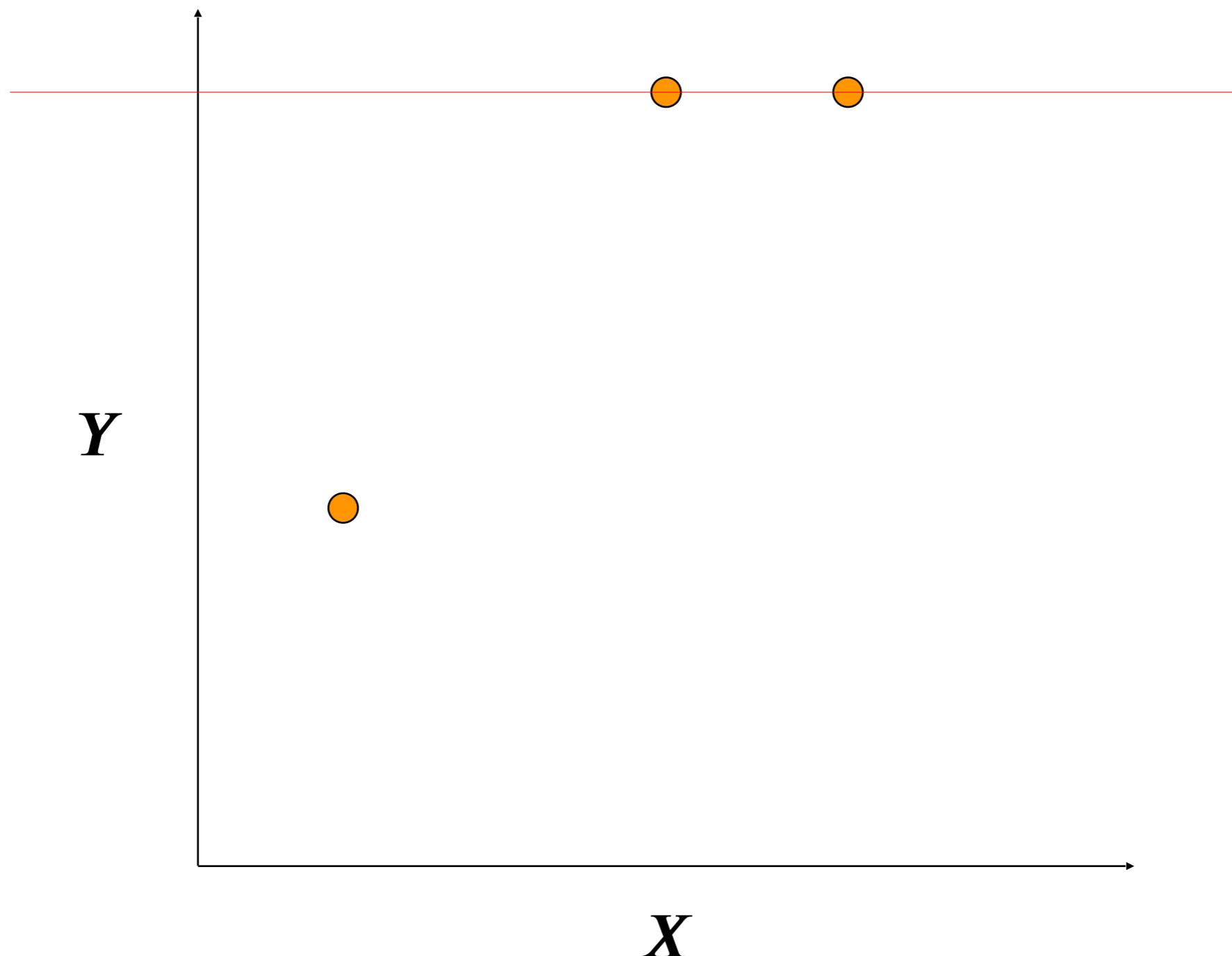
$$5 = w \times 2 + b$$

$$5 = w \times 3 + b$$

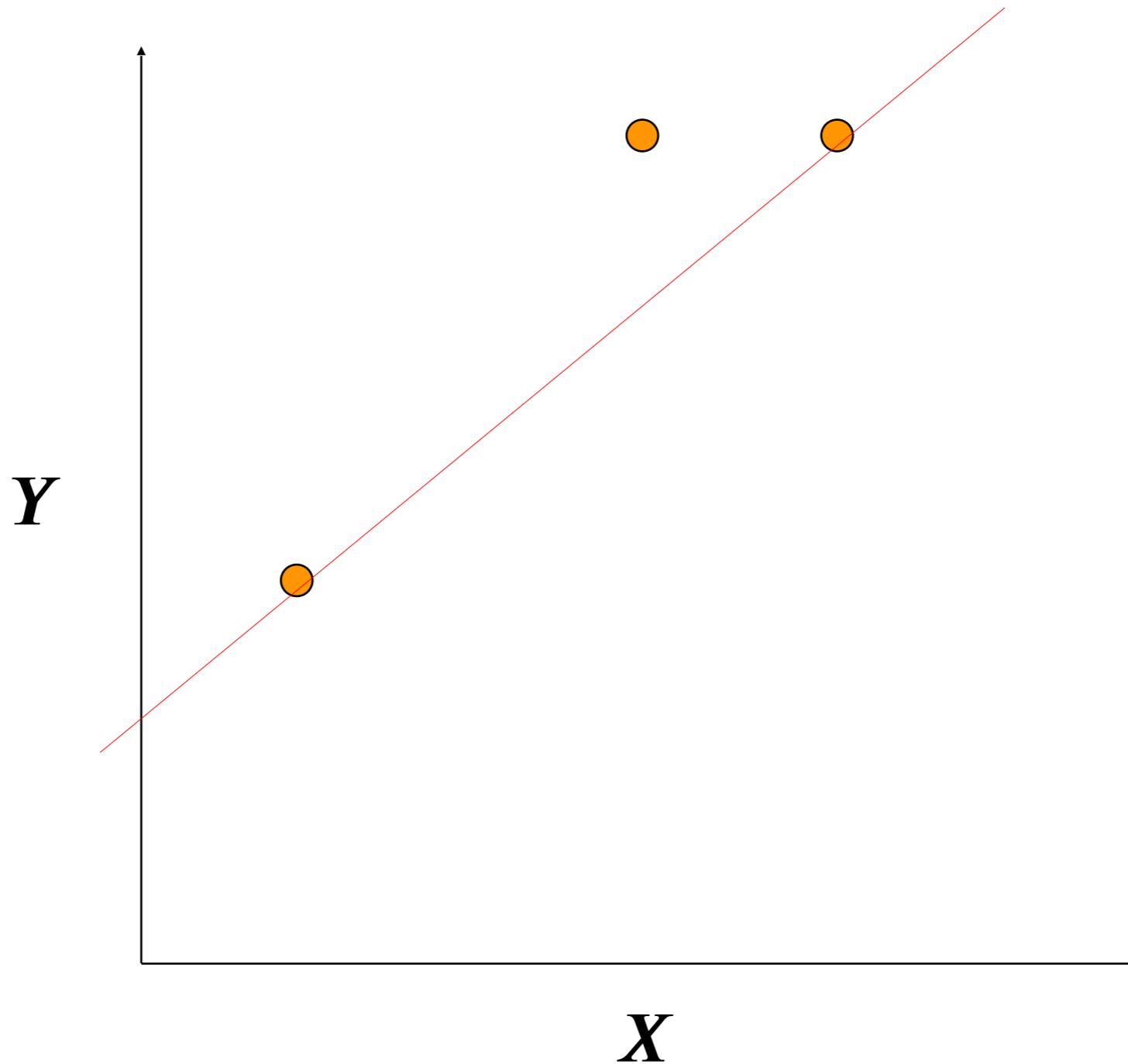
$$Y = ? \times X + ?$$

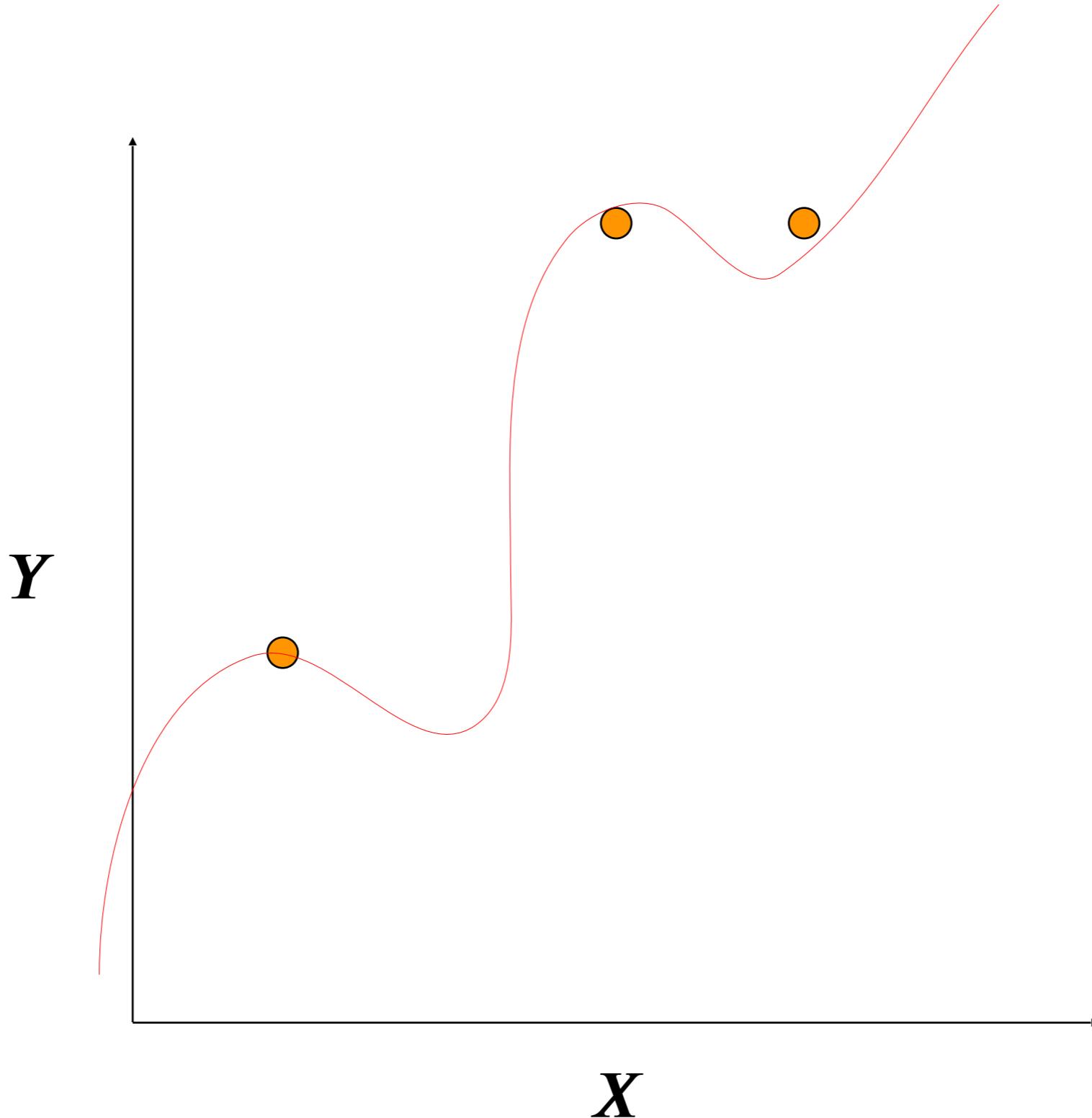


[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정





[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

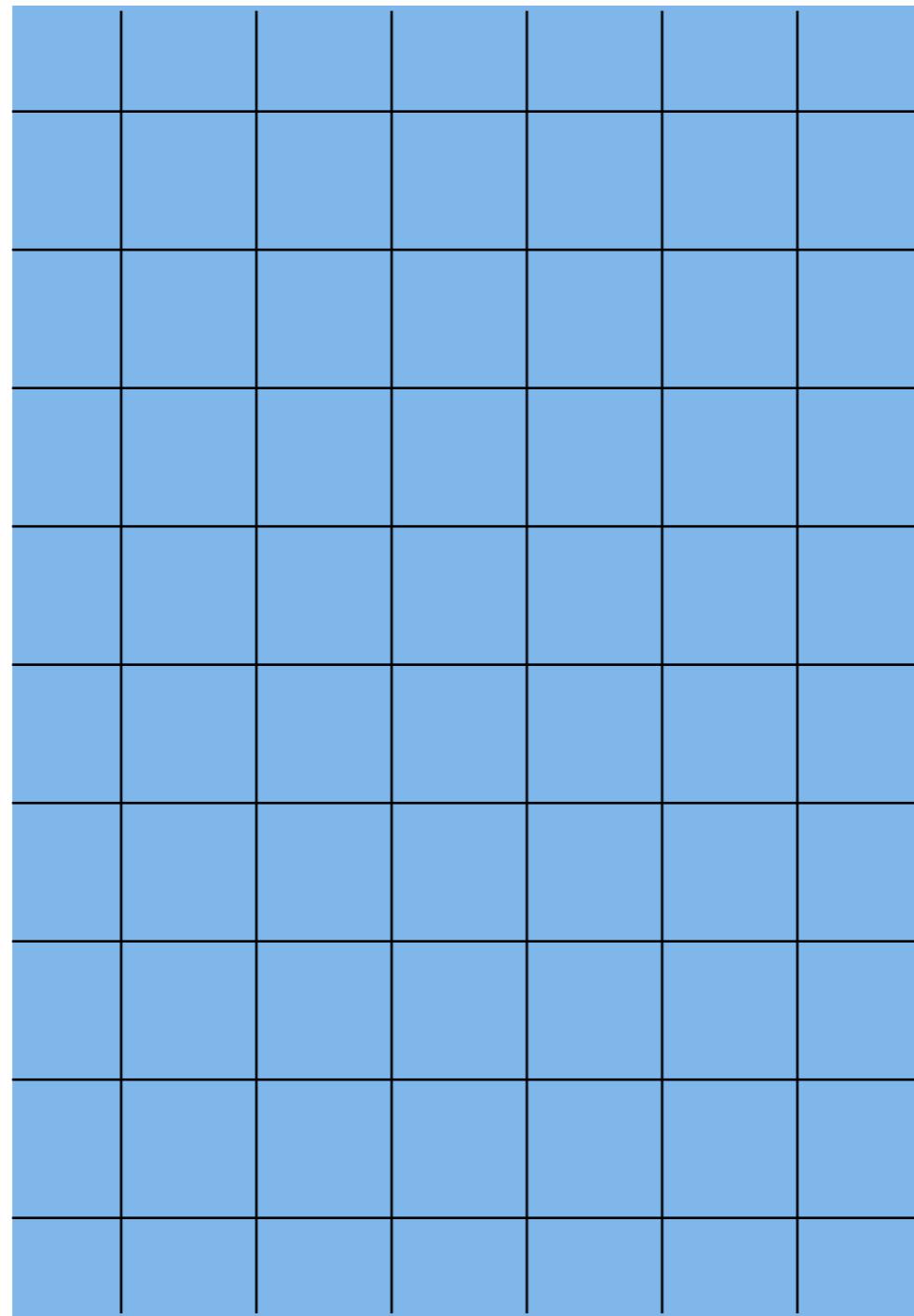
$$Y = w \times X + b$$

w와 b를 머신러닝 모델이
구해줍니다. 😊

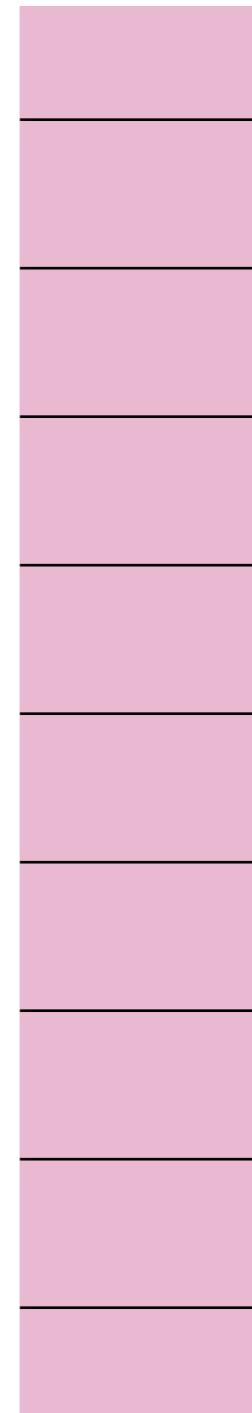
Data

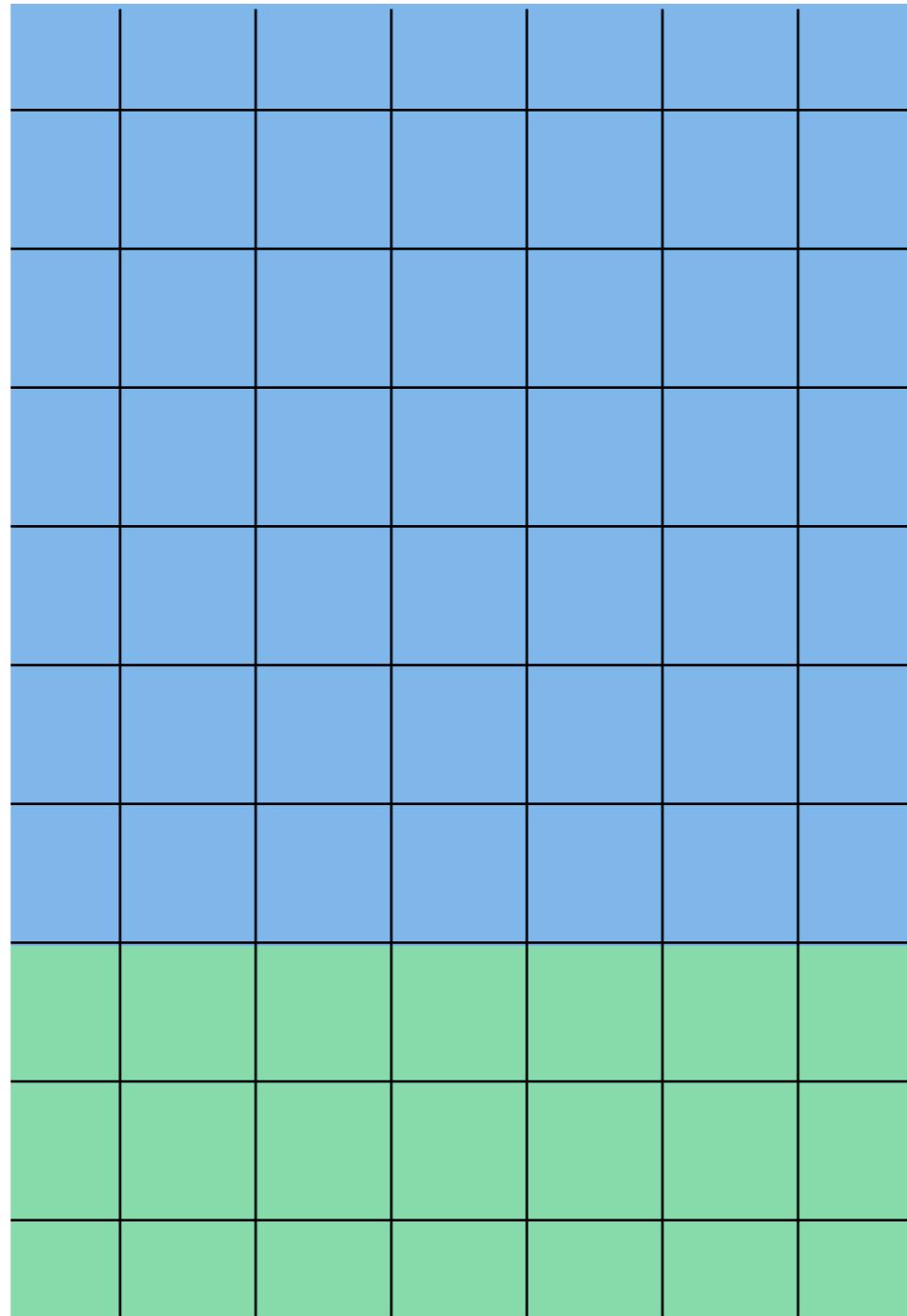
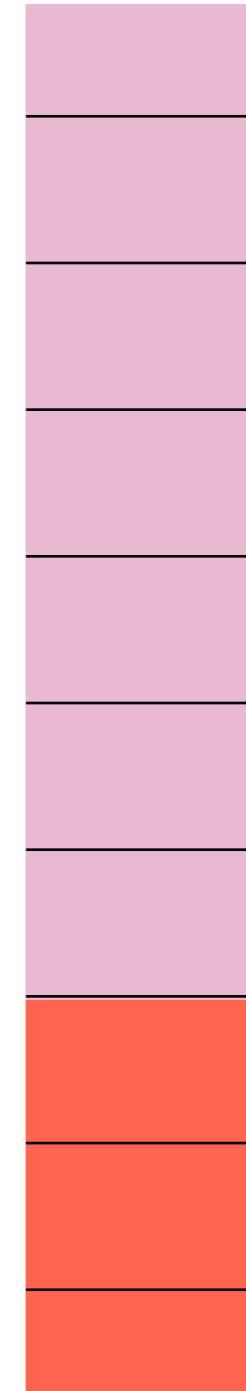
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

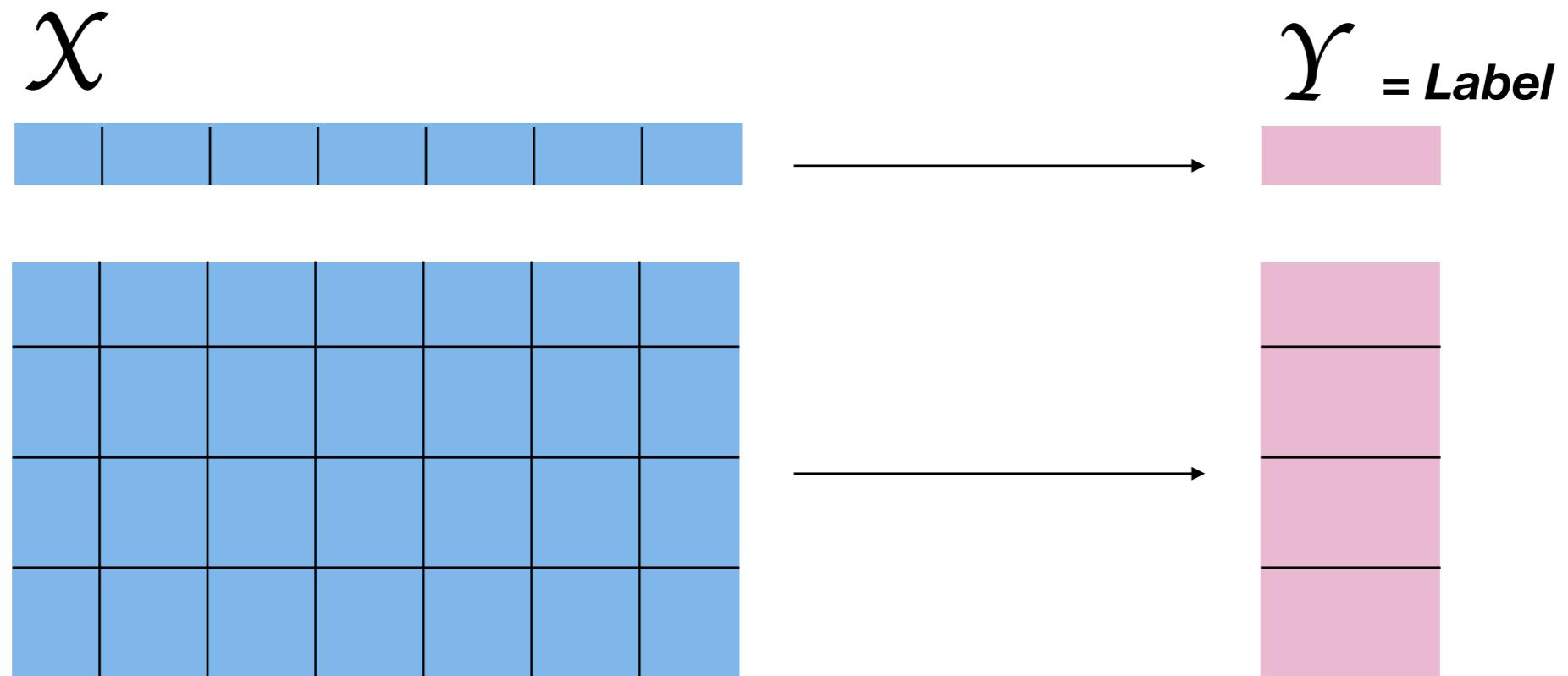
χ



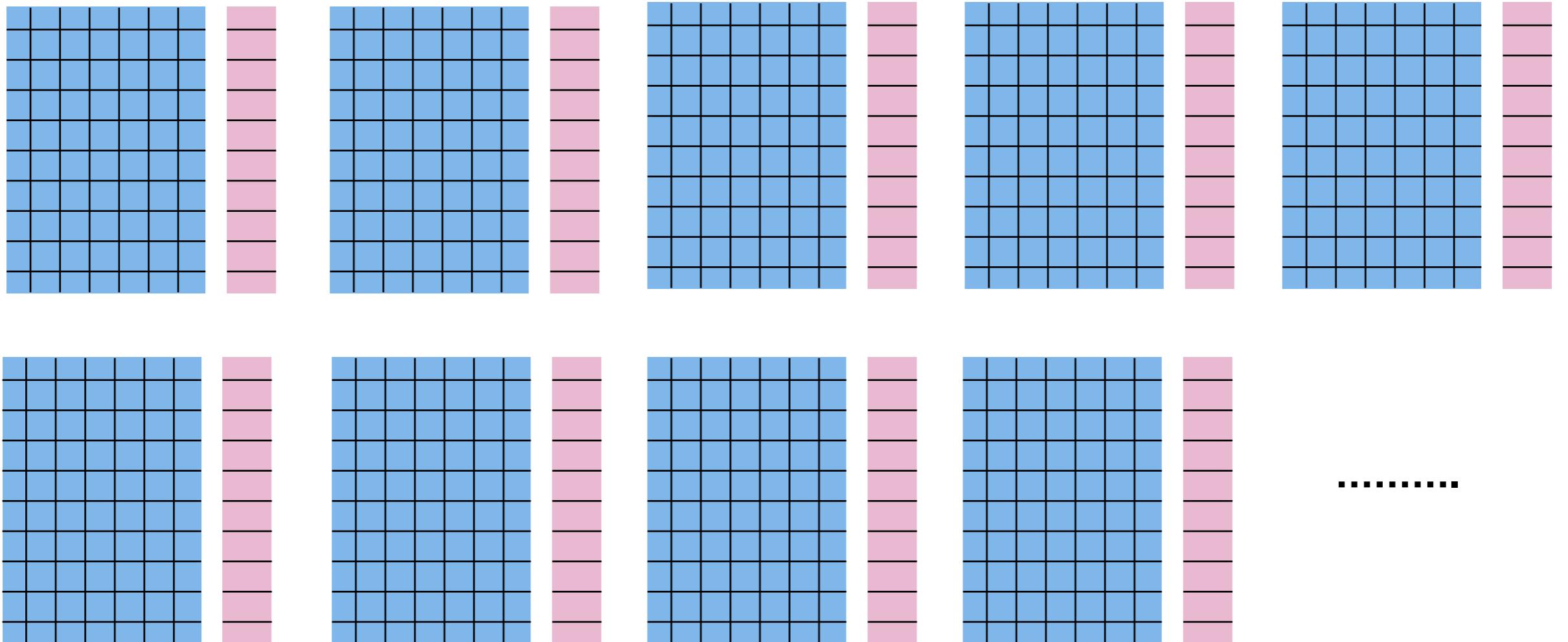
γ = Label



χ ***XTrain******XTest*** $\gamma = \text{Label}$ ***YTrain******YTest***



몇 문제를 풀고 답을 확인 할 것인가?

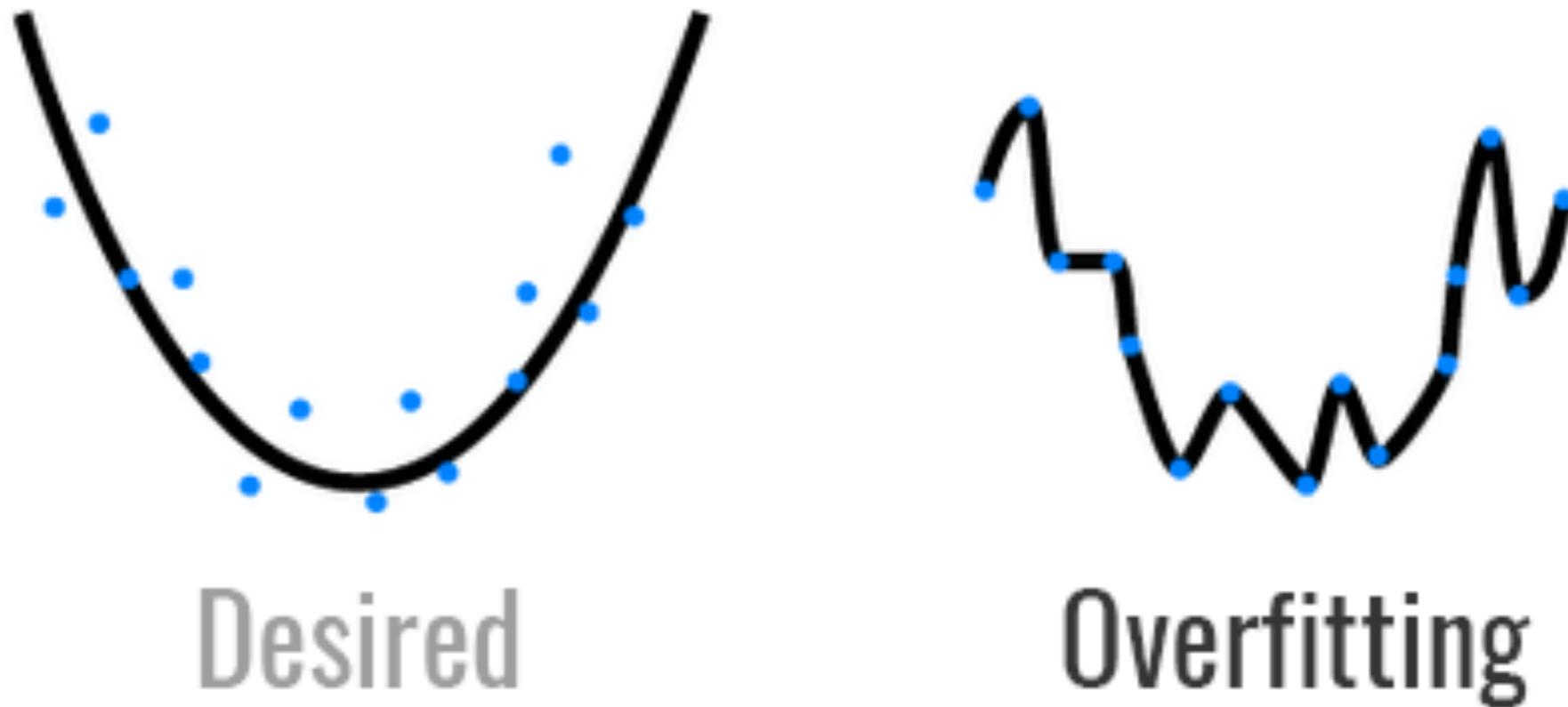


동일한 학습을 몇 번 반복할 것인가?



Desired

Overfitting



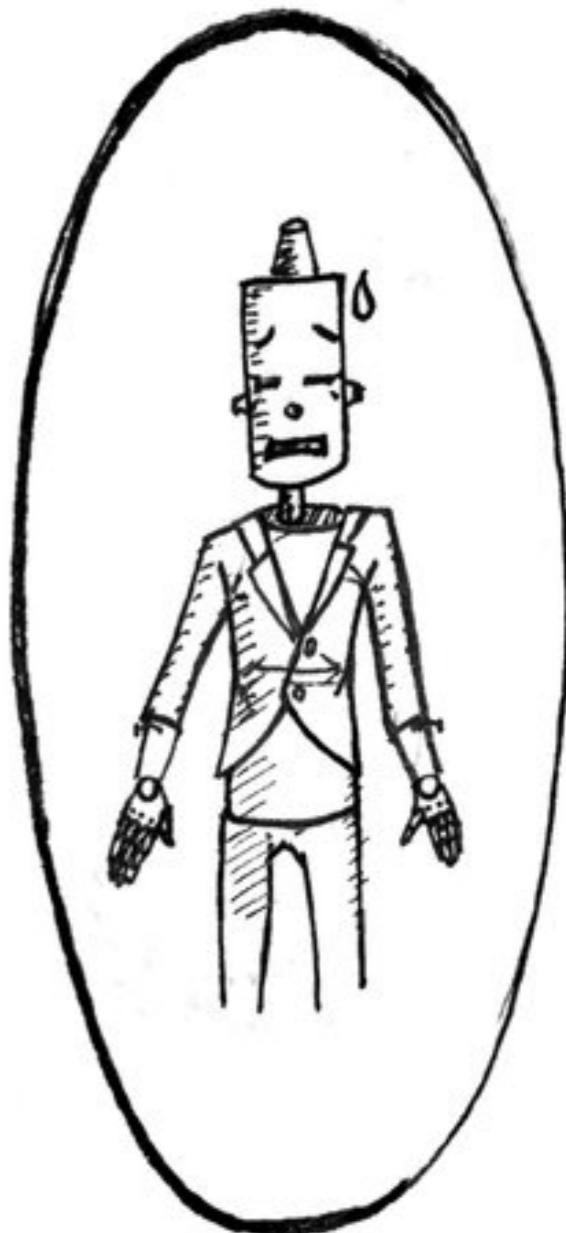
UNDERFIT



GOLDILOCKS ZONE

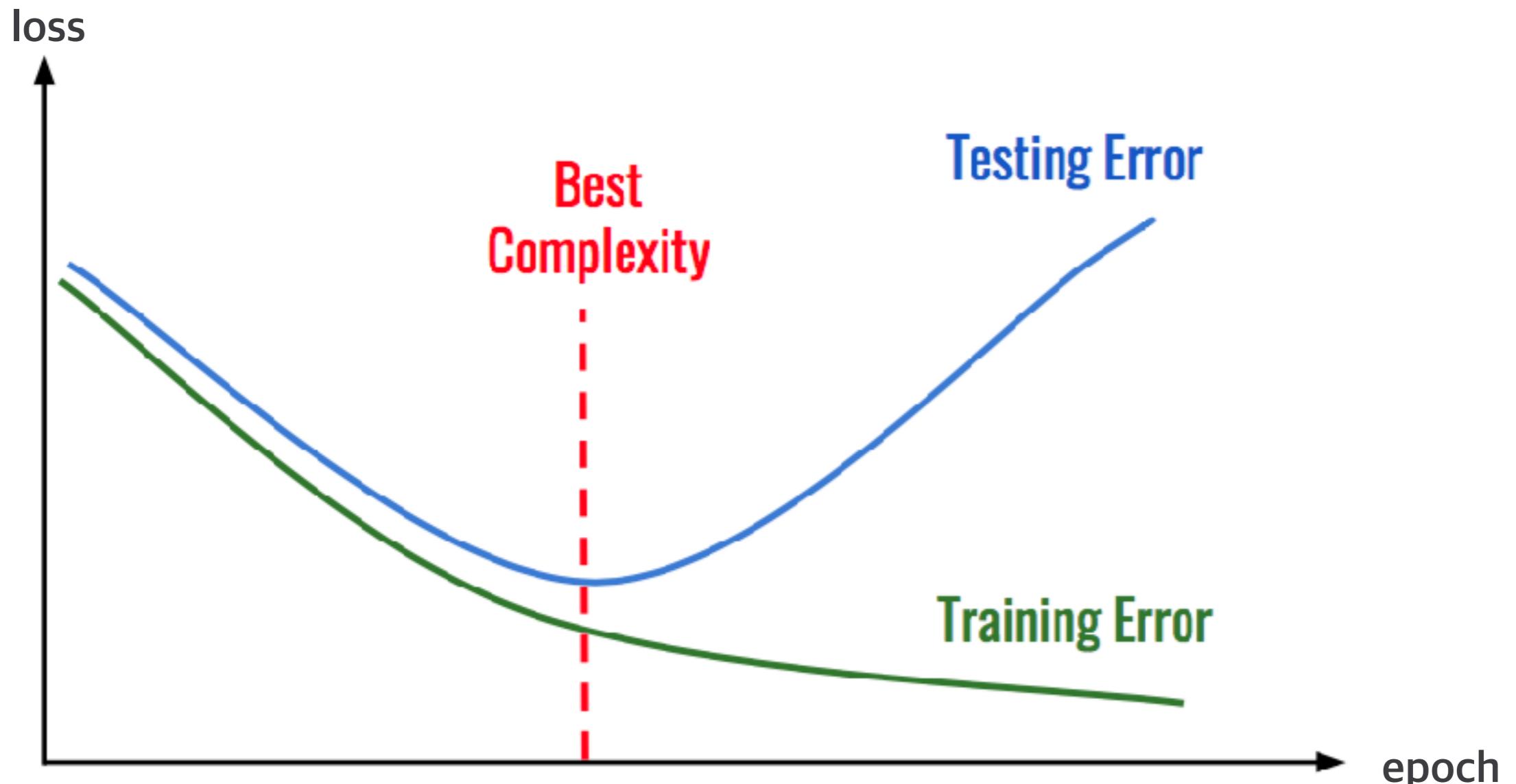


OVERFIT



[https://static1.squarespace.com/static/5213a664e4b01a5565dc90f1/t/5bc4e0c4e4966bc550291202/1546736393368/
Machine+Learning+Generalization](https://static1.squarespace.com/static/5213a664e4b01a5565dc90f1/t/5bc4e0c4e4966bc550291202/1546736393368/Machine+Learning+Generalization)

LOSS Function



<https://hackernoon.com/memorizing-is-not-learning-6-tricks-to-prevent-overfitting-in-machine-learning-820b091dc42>

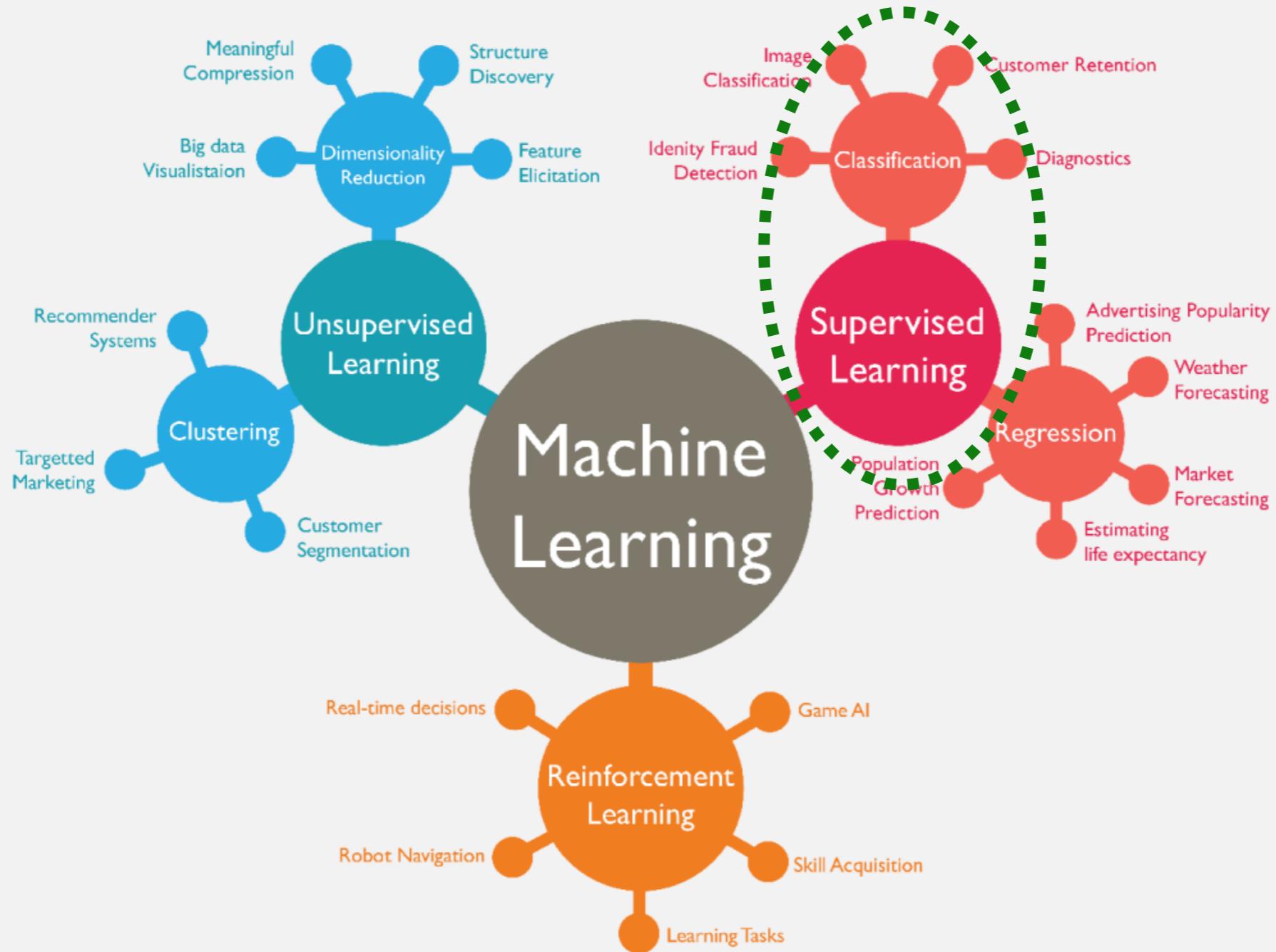
첫 번째 날

- 인공지능 개론
- 인공지능 사례
- 머신러닝 개론
- Classification 모델 생성

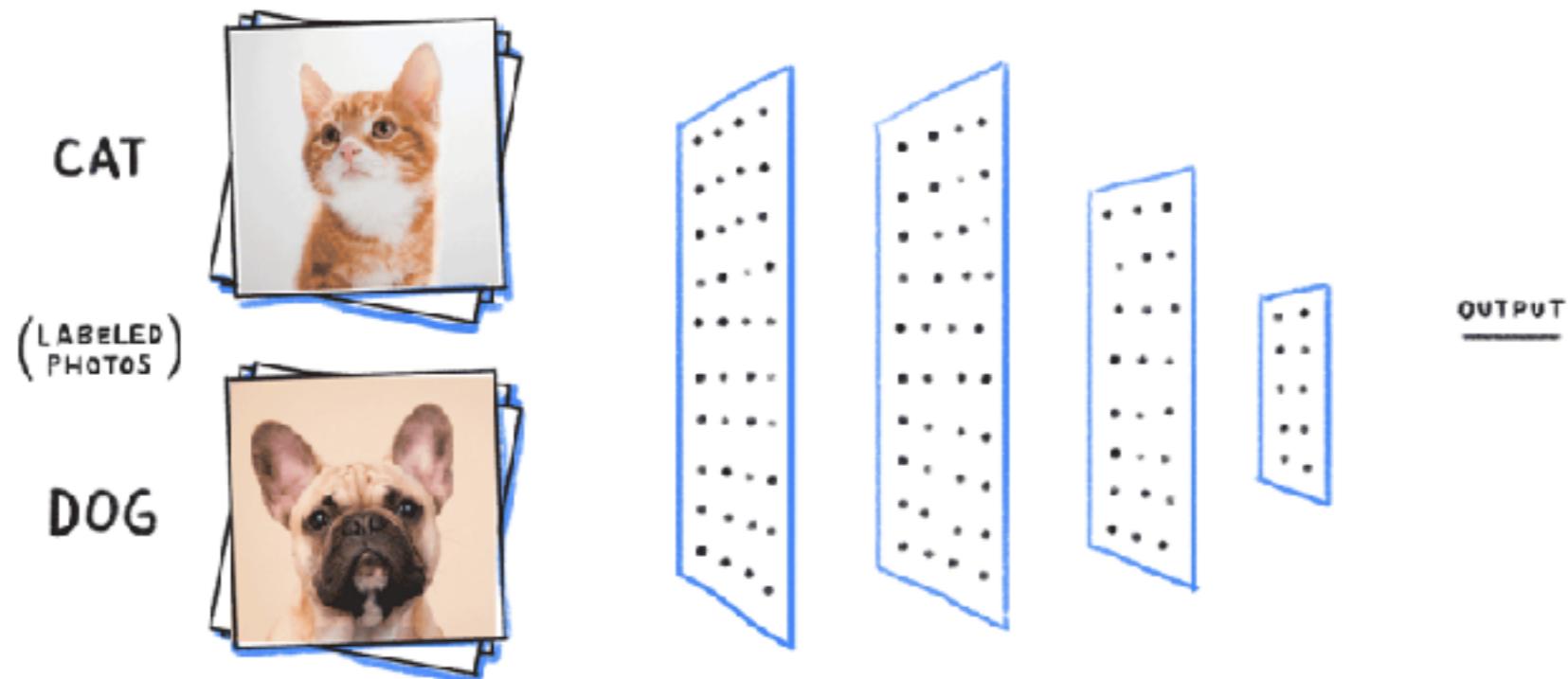
Azure ML Studio

Classification 모델 만들기

Machine Learning Bubble Chart



Classification



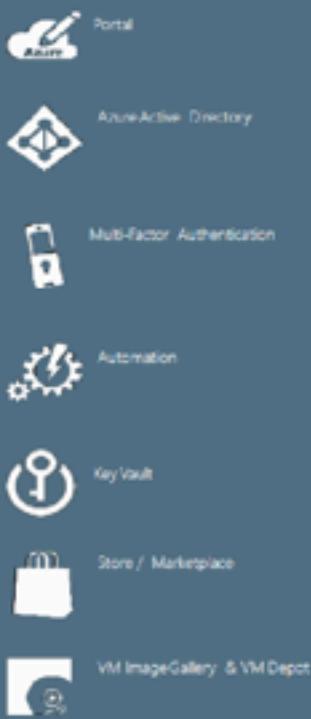
Azure ML Studio 소개

[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

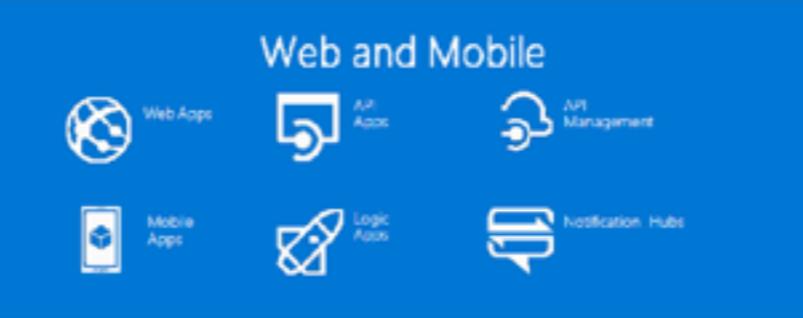


[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

Security & Management



Platform Services



Hybrid Operations



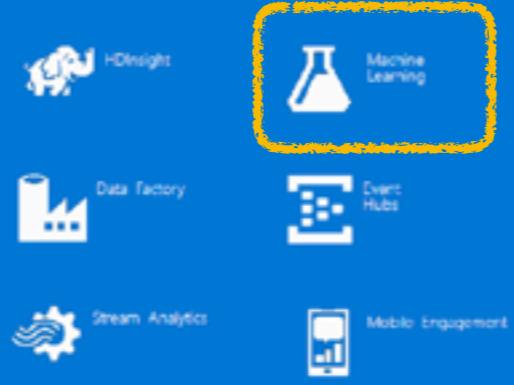
Integration



Media & CDN



Analytics & IoT



Data



Infrastructure Services

Compute



Storage



Networking



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

Azure AI Service



Machine Learning

쉽고 빠르게 모델을 만들어
배포, 관리하세요



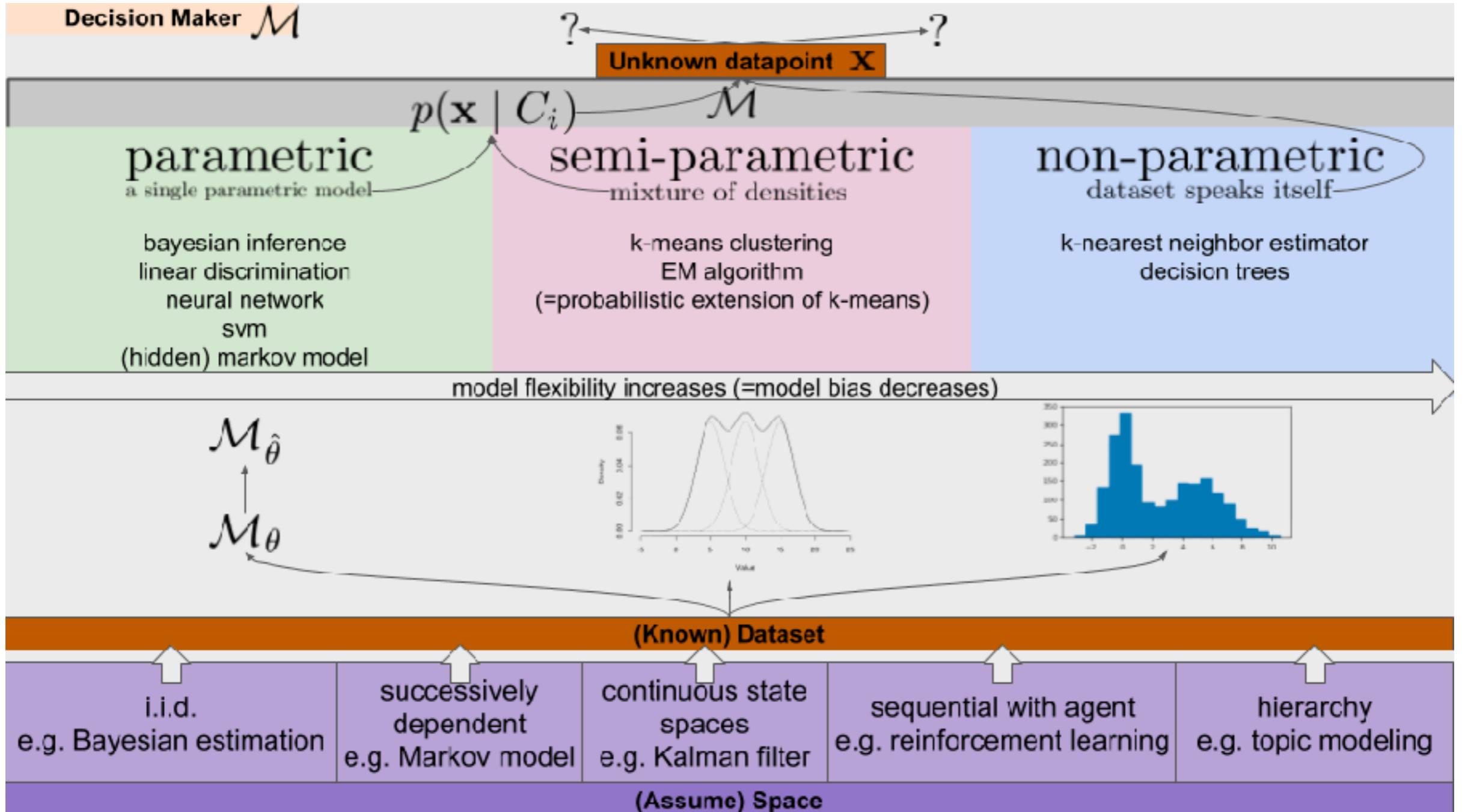
Azure Machine Learning Studio

이 공동 작업을 위한 끌어서 놓기 도구를 사용하여 데이터에 대한 예측 분석 솔루션을 빌드, 테스트 및 배포하는 방법을 알아봅니다. 자세히 알아보기...



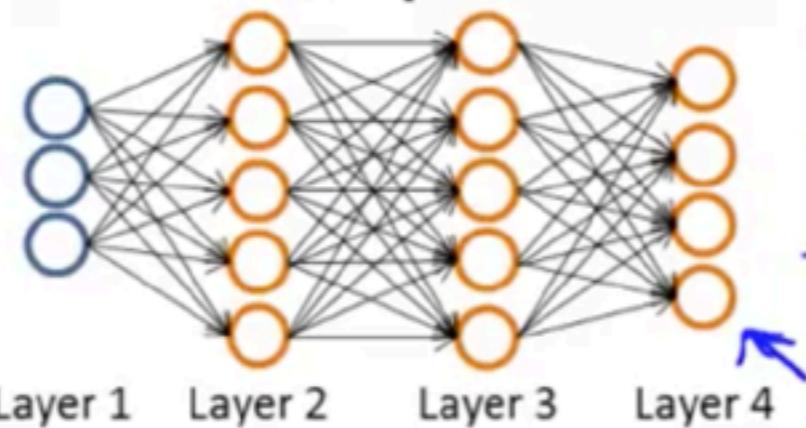
Azure Machine Learning 서비스

이 서비스에서 데이터 과학자가 CLI/Python 도구 및 라이브러리와 다양한 Azure 데이터 및 계산 서비스를 함께 사용하여 AI 솔루션을 개발하고 관리하도록 지원하는 방법을 알아봅니다. 자세히 알아보기...



<https://s3-ap-northeast-2.amazonaws.com/opentutorials-user-file/module/3653/9826.png>

Neural Network (Classification)



$\rightarrow \{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$

$\rightarrow L =$ total no. of layers in network $L = 4$

$\rightarrow s_l =$ no. of units (not counting bias unit) in layer l $s_1 = 3, s_2 = 5, s_3 = s_4 = 4$

Binary classification

$y = 0$ or 1

$$\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ h_{\Theta}(x) \end{array}$$

1 output unit

$$h_{\Theta}(x) \in \mathbb{R}$$

$$s_L = 1, \quad k = 1$$

Multi-class classification (K classes)

$$y \in \mathbb{R}^K \quad \text{E.g. } \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \leftarrow \begin{array}{l} \text{pedestrian} \\ \text{car} \\ \text{motorcycle} \\ \text{truck} \end{array}$$

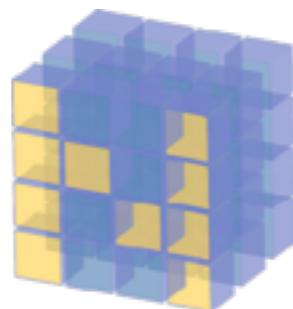
K output units

$$h_{\Theta}(x) \in \mathbb{R}^k$$

$$s_L = k \quad (k \geq 3)$$



matplotlib

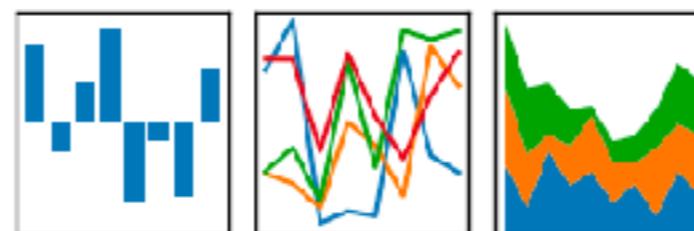


python

K Keras

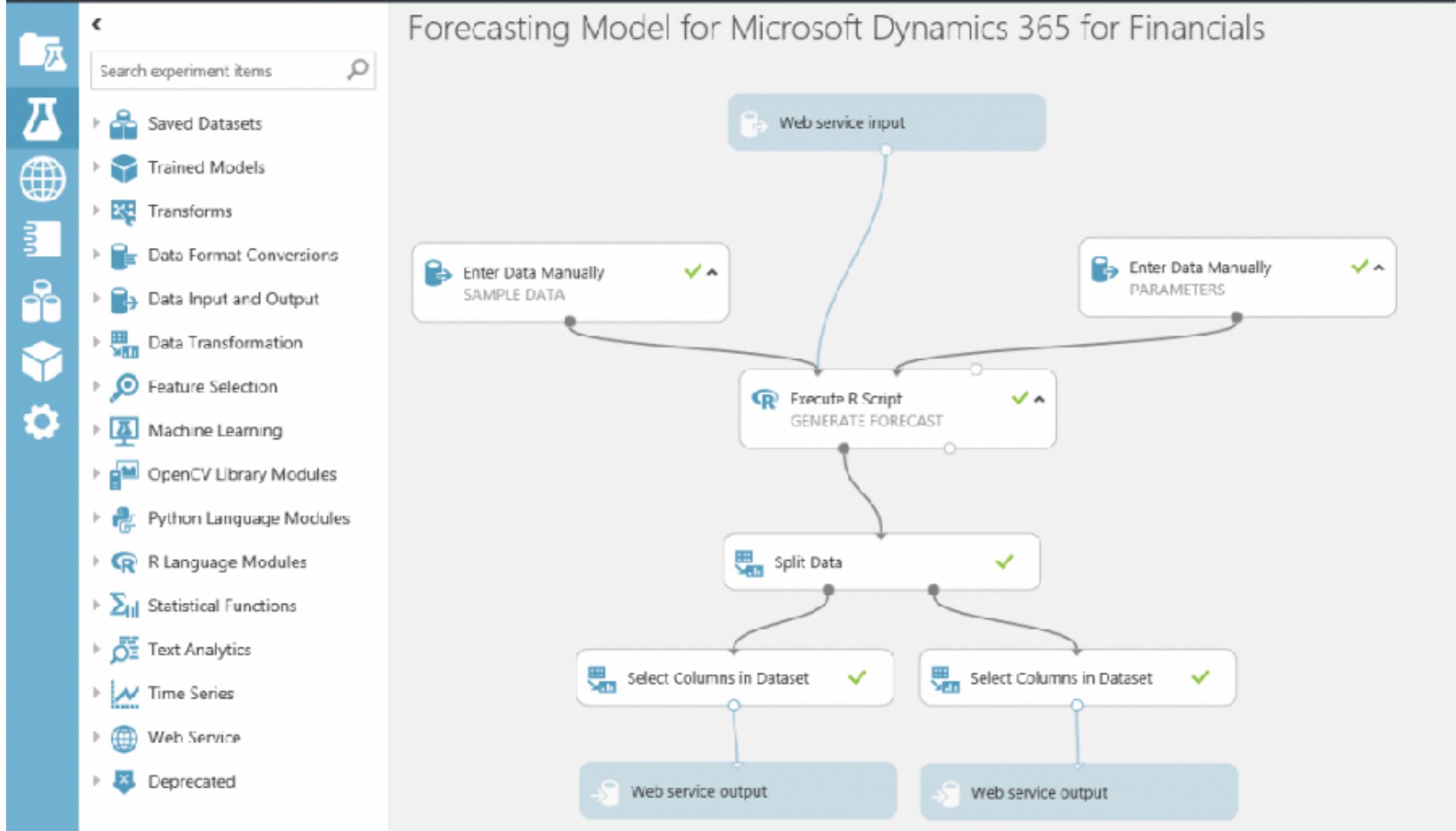
pandas

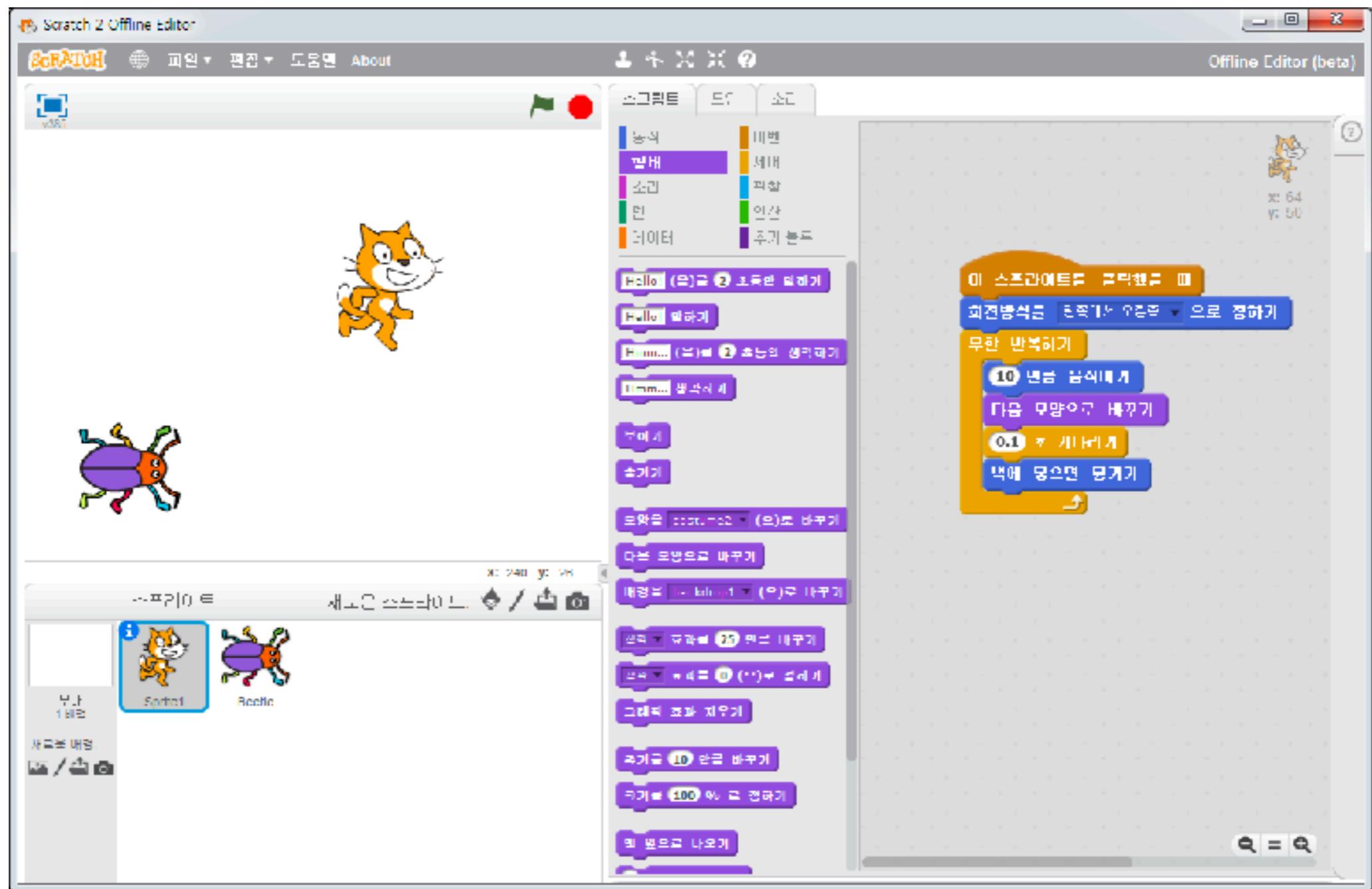
$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



TensorFlow

Microsoft Azure Machine Learning Studio





<https://t1.daumcdn.net/cfile/tistory/260FDC33530C309D2C>

Azure ML Studio

Visual Interface

Drag & Drop

No Terminal

No Code

R, Python, SQL

Tensorflow

	무료	STANDARD
가격	무료	₩11,235.254/매월 실제 사용자 수 ₩1,124.65/스튜디오 실험 시간
Azure 구독	필요 없음	필수
실험당 최대 모듈 수	100	제한 없음
최대 실험 기간	실험당 1시간	실험당 최대 7일, 모듈당 최대 24시간
최대 저장 공간	10GB	제한 없음 - BYO
온-프레미스 SQL에서 데이터 읽기 <small>미리보기</small>	아닙니다.	예
실행/성능	단일 노드	다중 노드
프로덕션 웹 API	아닙니다.	예
SLA	아닙니다.	예

문제 정의

데이터 셋 준비

모델 설정

모델 훈련 / 평가

모델 활용

Machine Learning in ML Studio

Anomaly Detection

- One-class Support Vector Machine
- Principal Component Analysis-based Anomaly Detection
- Time Series Anomaly Detection*

Classification

Two-class Classification

- Averaged Perceptron
- Bayes Point Machine
- Boosted Decision Tree
- Decision Forest
- Decision Jungle
- Logistic Regression
- Neural Network
- Support Vector Machine

Multi-class Classification

- Decision Forest
- Decision Jungle
- Logistic Regression
- Neural Network
- One-vs-all

Clustering

- K-means Clustering

Recommendation

- Matchbox Recommender

Regression

- Bayesian Linear Regression
- Boosted Decision Tree
- Decision Forest
- Fast Forest Quantile Regression
- Linear Regression
- Neural Network Regression
- Ordinal Regression
- Poisson Regression

Statistical Functions

- Descriptive Statistics
- Hypothesis Testing T-Test
- Linear Correlation
- Probability Function Evaluation

Text Analytics

- Feature Hashing
- Named Entity Recognition
- Vowpal Wabbit

Computer Vision

- OpenCV Library

Data/Model Visualization

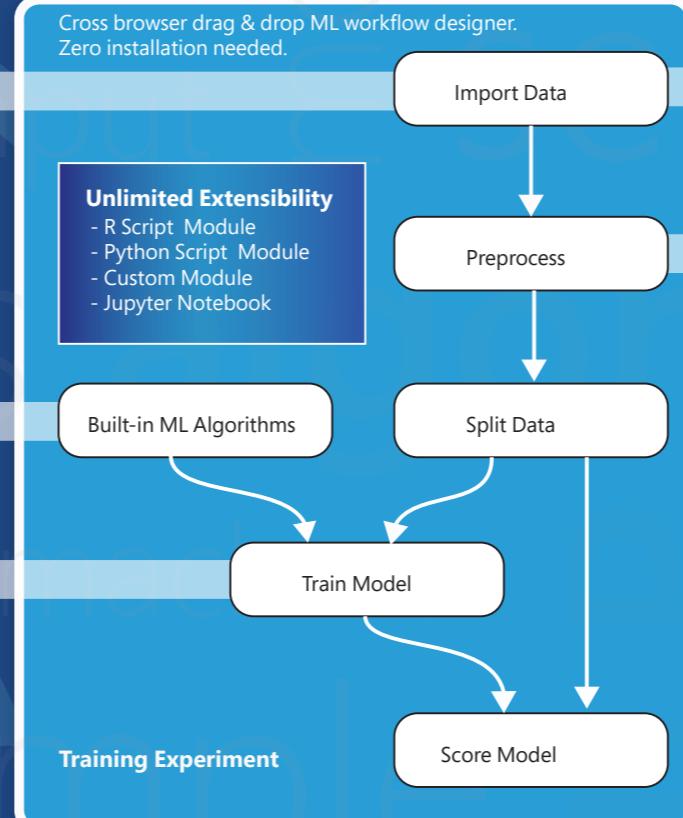
- Scatterplots
- Bar Charts
- Box plots
- Histogram
- R and Python Plotting Libraries
- REPL with Jupyter Notebook
- ROC, Precision/Recall, Lift
- Confusion Matrix
- Decision Tree*

Training

- Cross Validation
- Retraining
- Parameter Sweep

<https://studio.azureml.net>

Guest Access Workspace: Free trial access without logging in.
Free Workspace: Free persisted access, no Azure subscription needed.
Standard Workspace: Full access with SLA under an Azure subscription.



Data Source

- Azure Blob Storage
- Azure SQL DB
- Azure SQL DW*
- Azure Table
- Desktop Direct Upload
- Hadoop Hive Query
- Manual Data Entry
- OData Feed
- On-prem SQL Server*
- Web URL (HTTP)

Data Format

- ARFF
- CSV
- SVMLight
- TSV
- Excel
- ZIP

Data Preparation

- Clean Missing Data
- Clip Outliers
- Edit Metadata
- Feature Selection
- Filter
- Learning with Counts
- Normalize Data
- Partition and Sample
- Principal Component Analysis
- Quantize Data
- SQLite Transformation
- Synthetic Minority Oversampling Technique

Enterprise Grade Cloud Service

- SLA: 99.95% Guaranteed Up-time
- Azure AD Authentication
- Compute at Large Scale
- Multi-geo Availability
- Regulatory Compliance*

Community

- Gallery (<http://gallery.azureml.net>)
- Samples & Templates
- Workspace Sharing and Collaboration
- Live Chat & MSDN Forum Support

* Feature Coming Soon



Azure Machine Learning Studio Capabilities Overview

© 2015 Microsoft Corporation. All rights reserved.

Created by the Azure Machine Learning Team

Email: AzurePoster@microsoft.com

Download this poster: <http://aka.ms/MLStudioOverview>



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

Cross browser drag & drop ML workflow designer.
Zero installation needed.

Unlimited Extensibility

- R Script Module
- Python Script Module
- Custom Module
- Jupyter Notebook

Import Data

Preprocess

Built-in ML Algorithms

Split Data

Train Model

Training Experiment

Score Model

문제 정의

데이터 셋 준비

모델 설정

모델 훈련 / 평가

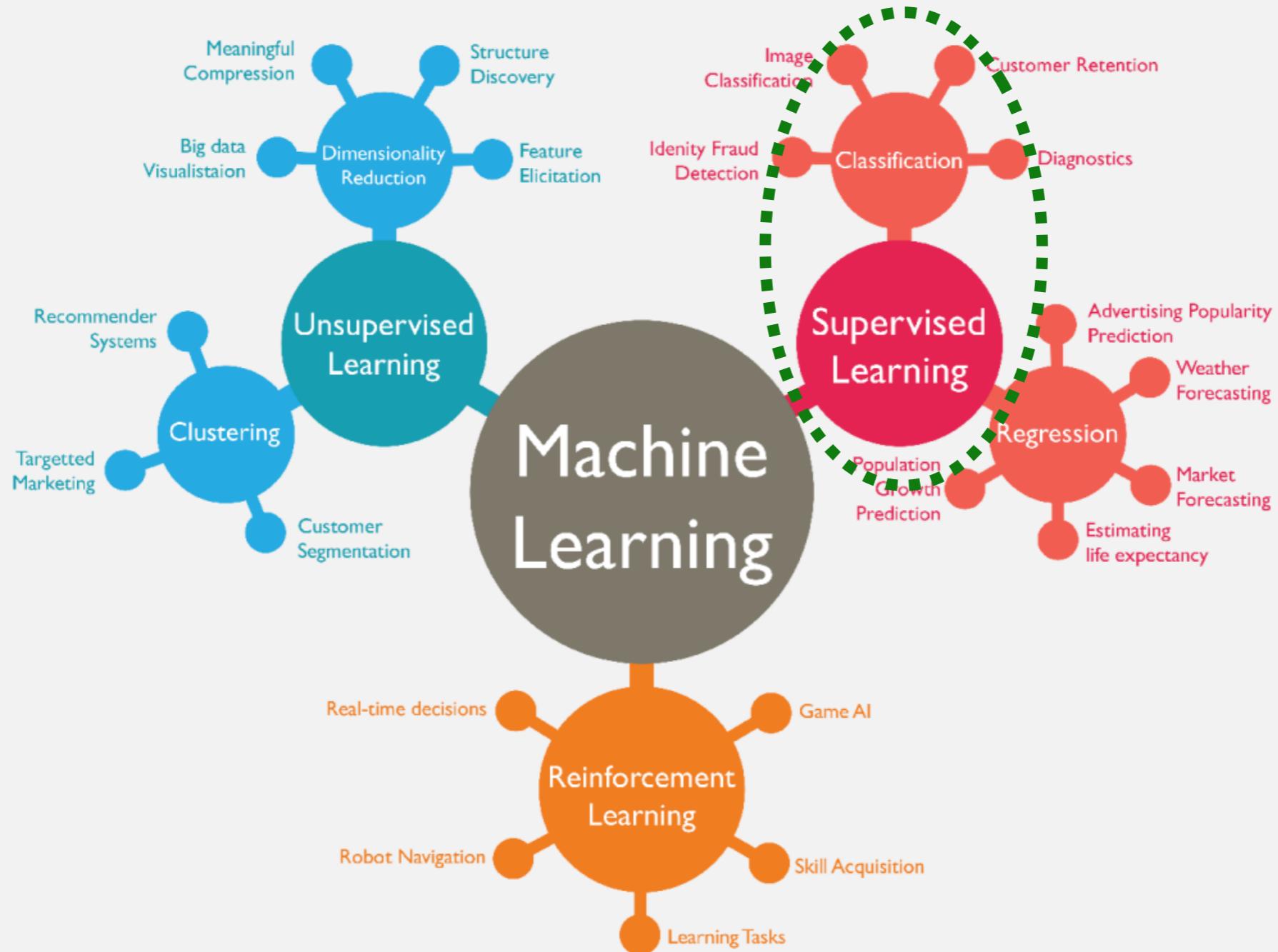
모델 활용

총 111개 모듈

문제 정의

어떤 모델을 만들 것인가?

Machine Learning Bubble Chart



타이타닉 탑승객 생존 여부 예측



[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

타이타닉 탑승객 생존 여부 예측

Alive or Dead?

Let's start HOL



part1. 데이터 전처리

part2. 모델 학습 / 예측

part1. 데이터 전처리

- 데이터 지원 형식
- 데이터 업로드
- 데이터 합치기
- 데이터 전처리

데이터 지원 형식

- 헤더가 있거나(.csv) 없는(.nh.csv) 쉼표로 구분된 값(CSV)
- 헤더가 있거나(.tsv) 없는(.nh.tsv) 탭으로 구분된 값(TSV)
- 일반 텍스트(.txt)
- Excel 파일, Azure 테이블, Hive 테이블
- SQL 데이터베이스 테이블
- SVMLight 데이터(.svmlight)
- 특성 관계 파일 형식(ARFF) 데이터(.arff)
- Zip 파일(.zip)
- R 개체 또는 작업 영역 파일(. RData)

데이터 지원 유형

- 문자열
- 정수
- double
- BOOLEAN
- Datetime
- timespan

데이터 지원 용량

1.98 GB/file

Total 10 GB

Cross browser drag & drop ML workflow designer.
Zero installation needed.

Unlimited Extensibility

- R Script Module
- Python Script Module
- Custom Module
- Jupyter Notebook

Built-in ML Algorithms

Train Model

Import Data

Preprocess

Split Data

Score Model

Training Experiment

http://bit.ly/0812_data1

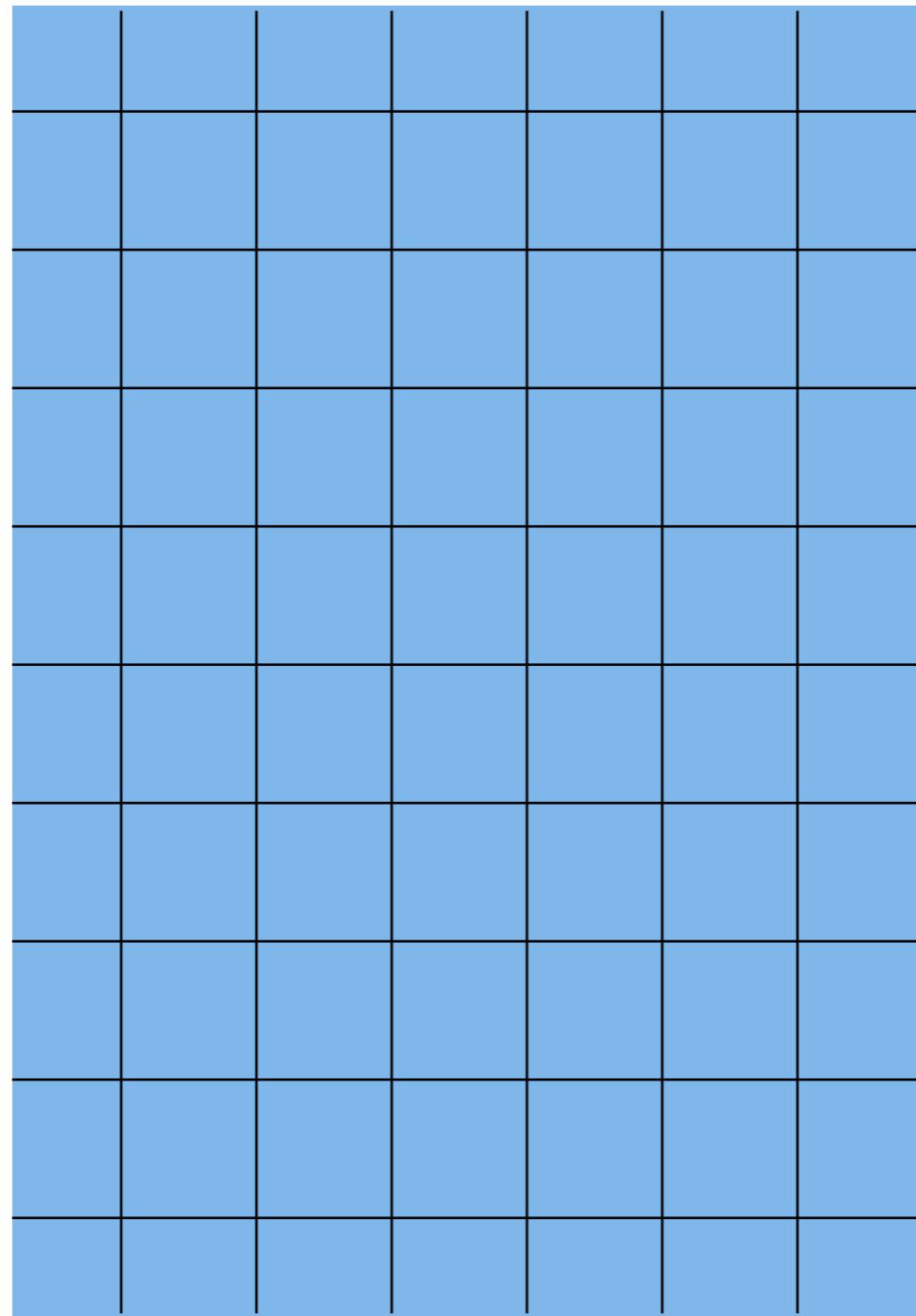


AzureHOL_Titanic_dataset.csv

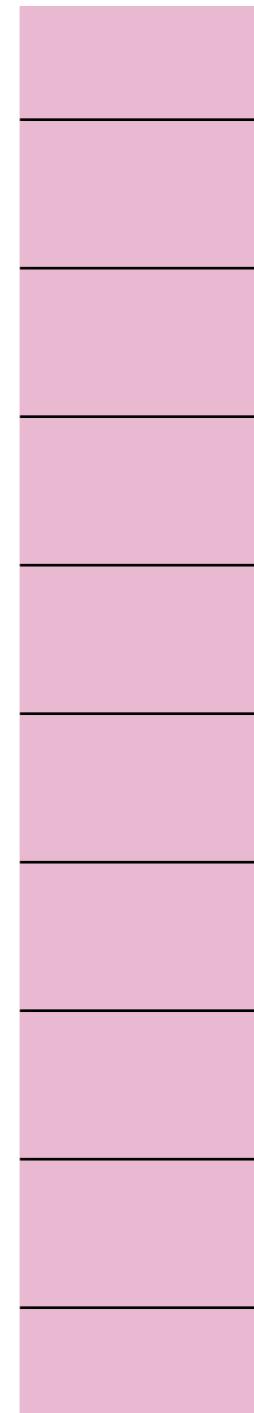
login

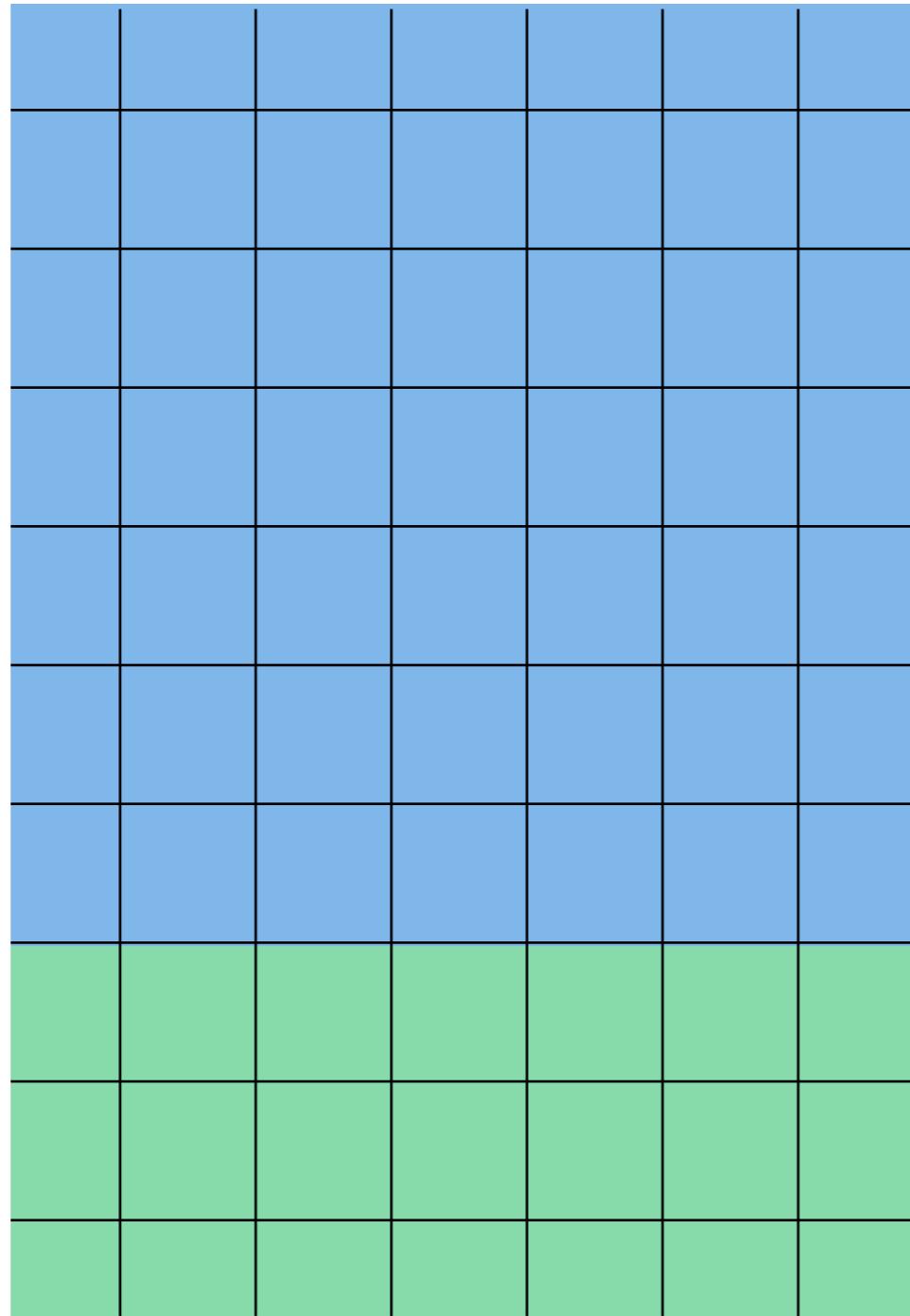
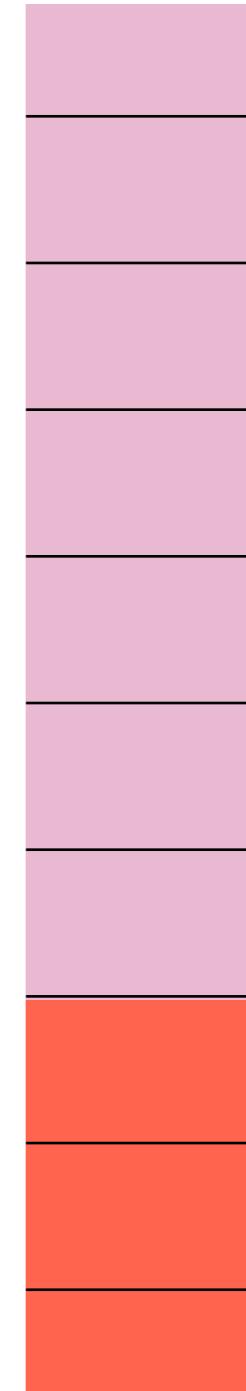
<https://studio.azureml.net>

χ

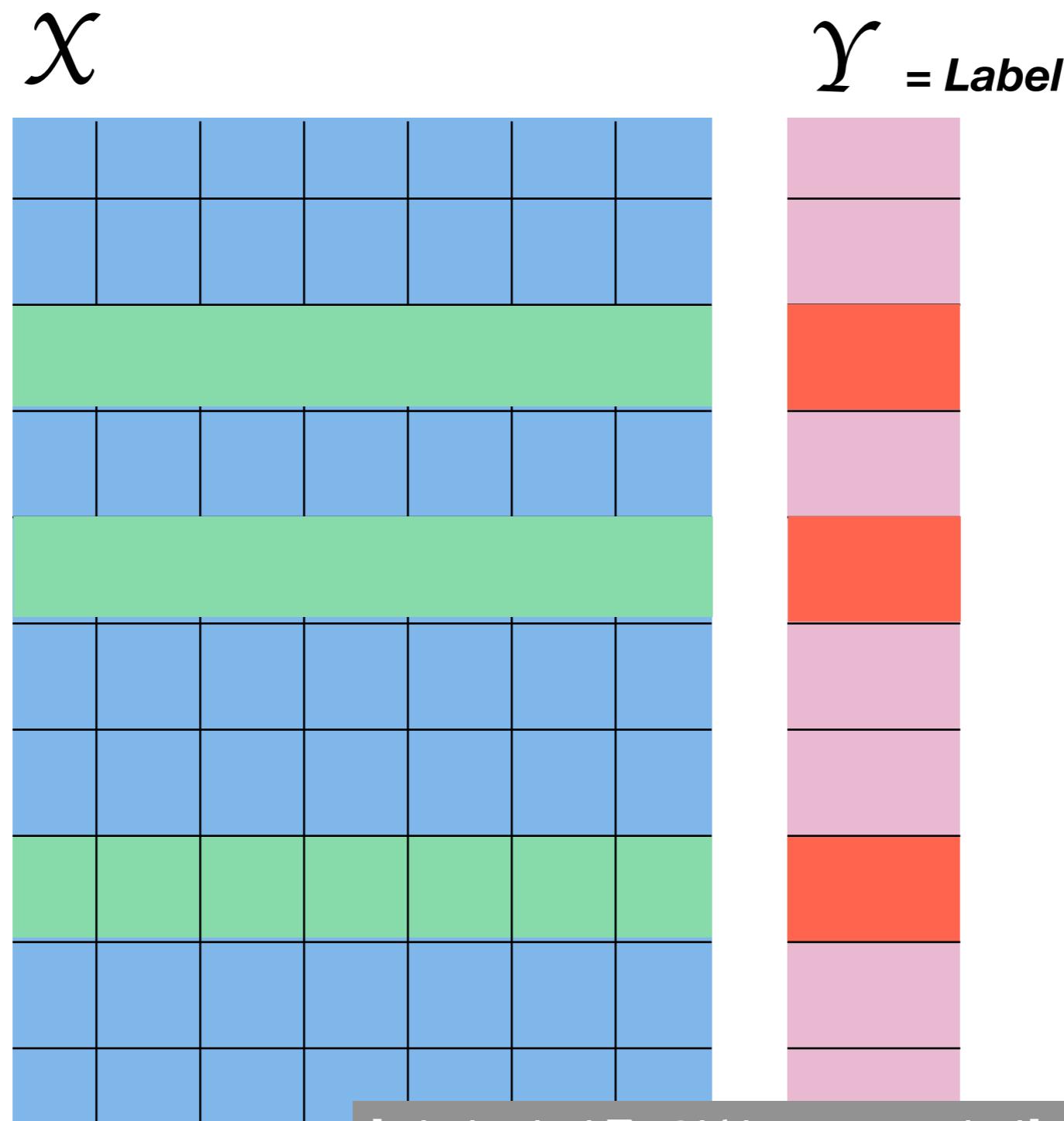


γ = Label

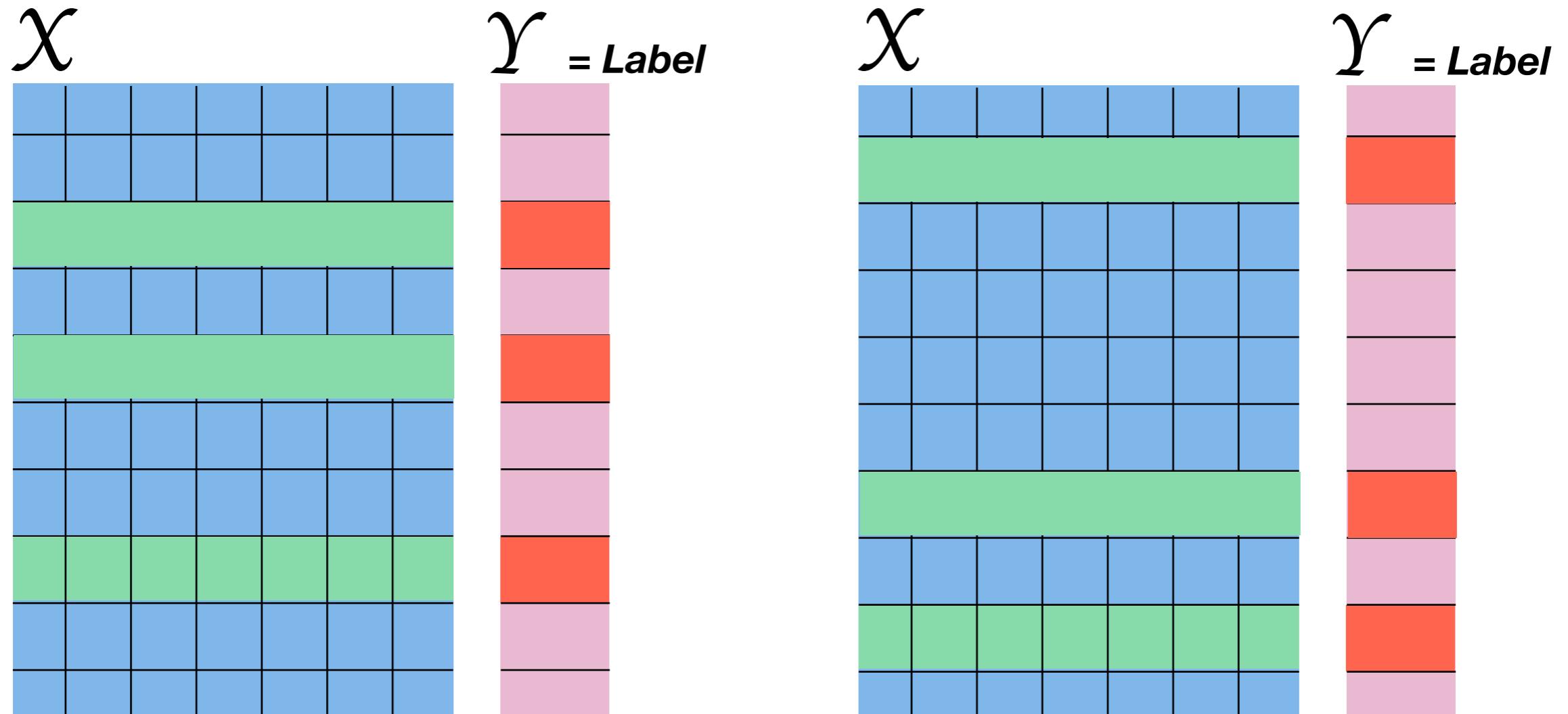


χ ***XTrain******XTest*** $\gamma = \text{Label}$ ***YTrain******YTest***

Randomized split



Random Seed

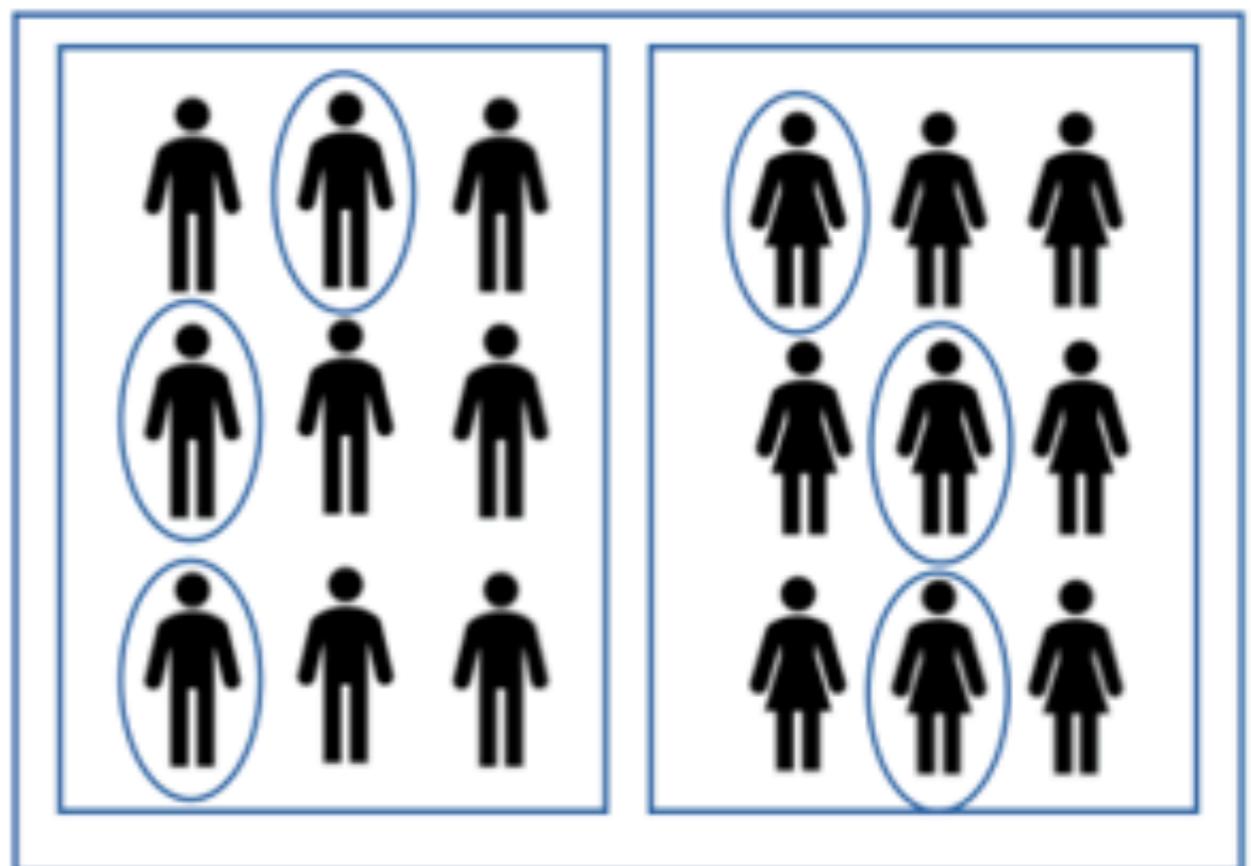


Stratified Split

Normal Sampling 6 of 18



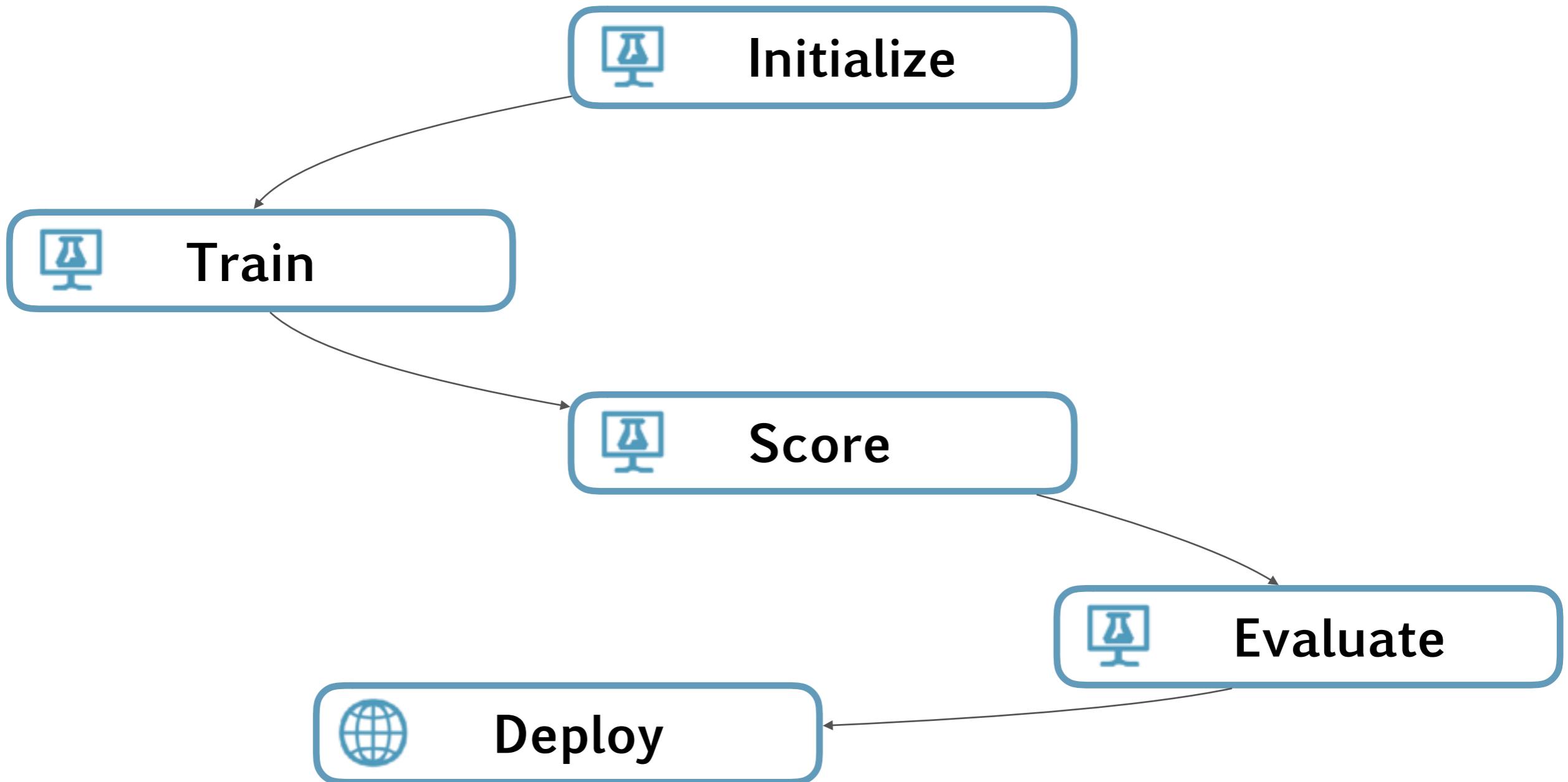
Stratified Sampling 6 of 18



<https://dlbjbjzgnk95t.cloudfront.net/1106000/1106385/graphic%201%20-%20srs.png>

part2. 모델 학습 / 예측

- 다양한 모델 살펴보기
- 올바른 모델 선택
- 모델 학습, 평가
- 모델 배포, 활용



***“What machine learning algorithm
should I use?”***

“It depends.”

*On size & quality of data, train time, computer,
application*

Supervised Learning

**Classification
Regression**

Unsupervised Learning

Clustering

**Reinforcement
Learning**

Accuracy

Parameter

Training Time

Feature

Linearity

Etc.

<https://docs.microsoft.com/ko-kr/azure/machine-learning/studio/algorithm-choice>



Microsoft Azure Machine Learning: Algorithm Cheat Sheet

This cheat sheet helps you choose the best Azure Machine Learning Studio algorithm for your predictive analytics solution. Your decision is driven by both the nature of your data and the question you're trying to answer.

ANOMALY DETECTION

One-class SVM

>100 features,
aggressive boundary

PCA-based anomaly detection

Fast training

CLUSTERING

K-means

Discovering
structure

REGRESSION

Ordinal regression

Data in rank ordered categories

Poisson regression

Predicting event counts

Fast forest quantile regression

Predicting a distribution

Linear regression

Fast training, linear model

Bayesian linear regression

Linear model, small data sets

Neural network regression

Accuracy, long training time

Decision forest regression

Accuracy, fast training

Boosted decision tree regression

Accuracy, fast training,
large memory footprint

MULTI-CLASS CLASSIFICATION

MULTI-CLASS CLASSIFICATION

Fast training, linear model

Multiclass logistic regression

Accuracy, long training times

Multiclass neural network

Accuracy, fast training

Multiclass decision forest

Accuracy, small memory footprint

Multiclass decision jungle

Depends on the two-class
classifier, see notes below

One-v-all multiclass

START

TWO-CLASS CLASSIFICATION

Two-class SVM

>100 features,
linear model

Two-class decision forest

Two-class averaged perceptron

Fast training,
linear model

Two-class boosted decision tree

Two-class logistic regression

Fast training,
linear model

Two-class decision jungle

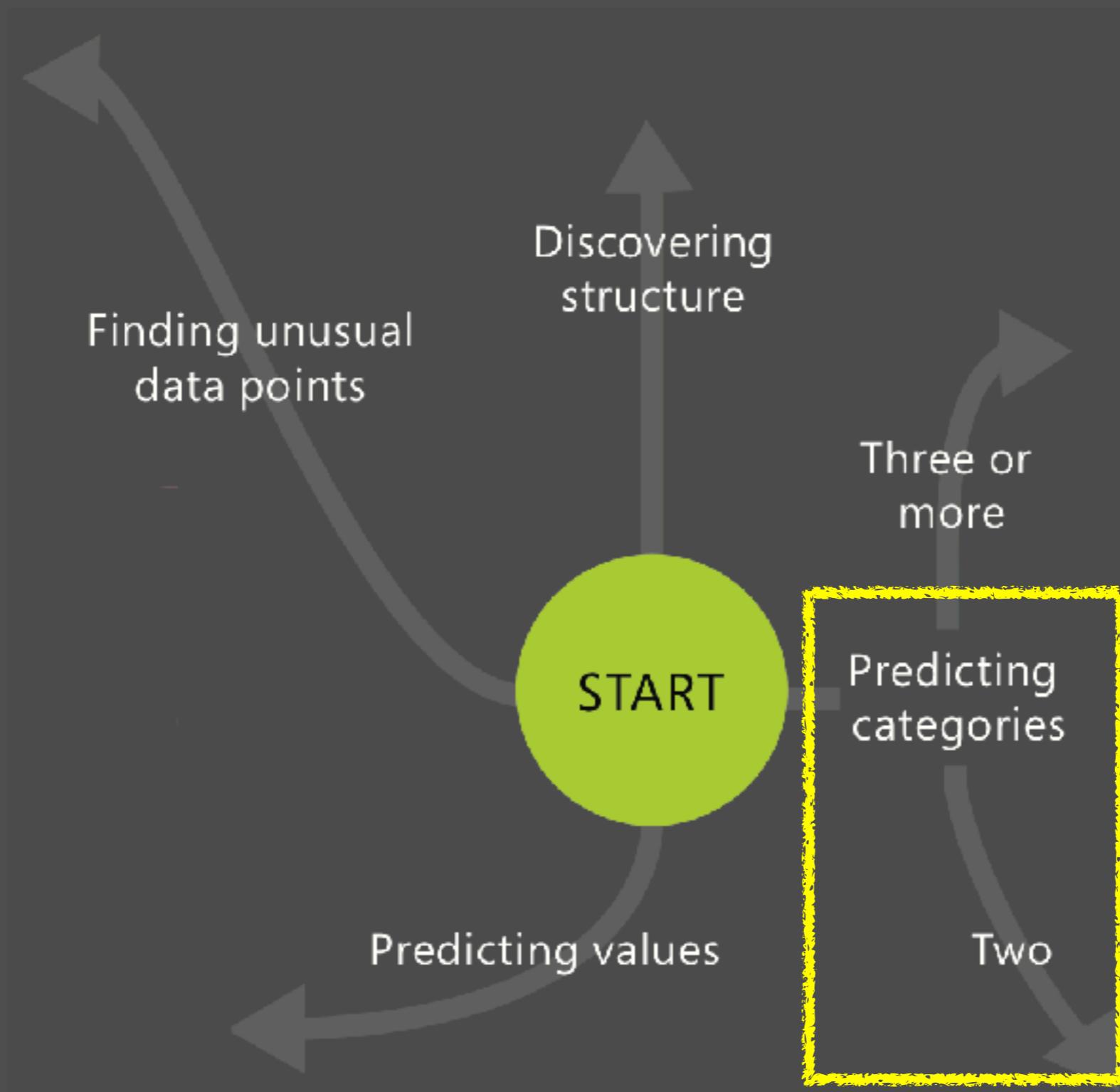
Two-class Bayes point machine

Fast training,
linear model

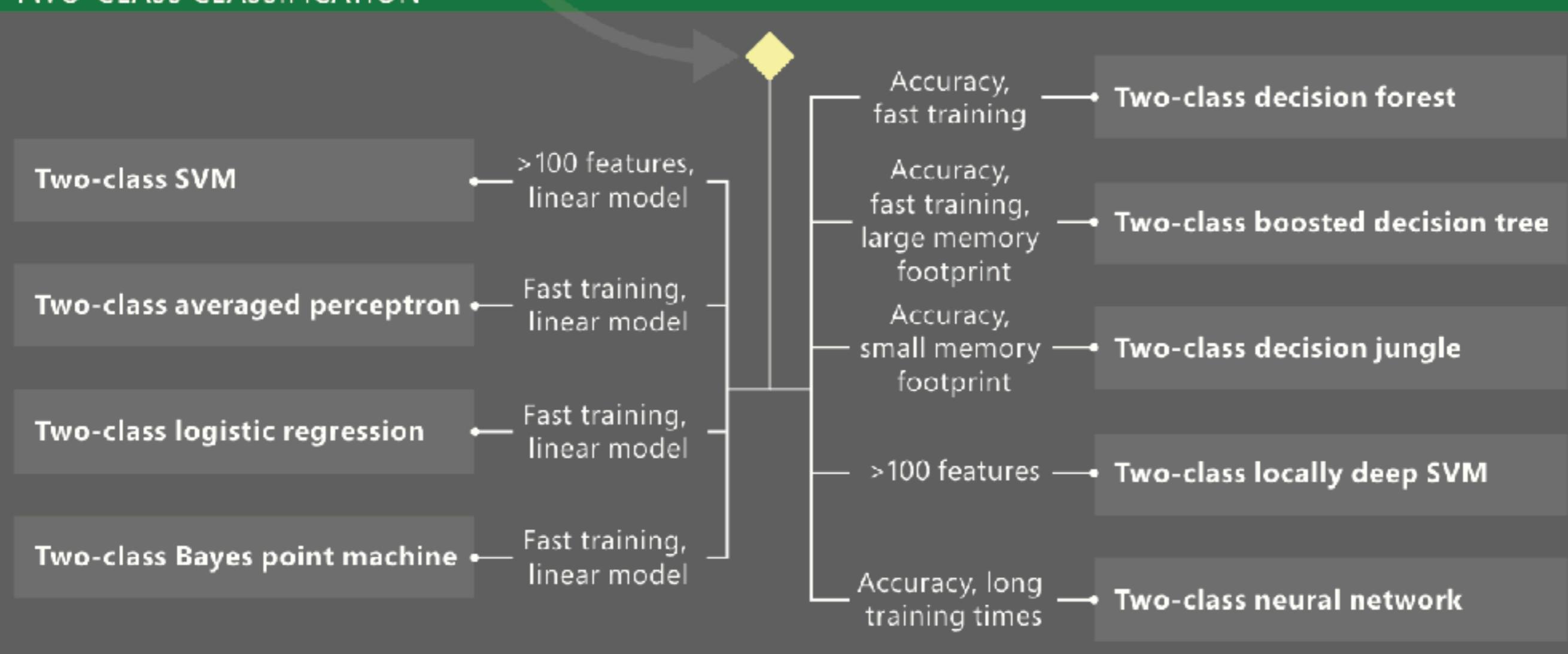
Two-class locally deep SVM

Accuracy, long
training times





TWO-CLASS CLASSIFICATION



Classification

label = 5



label = 0



label = 2



label = 1



- 개 / 고양이 이미지 분류
- 0 ~ 9 숫자 이미지 분류(MNIST)
- 쇼핑몰 고객 등급 분류
- 장바구니 물품 구매 여부 예측
- 입시 합격 여부 예측
- 뉴스 기사 분야 분류

<https://corochann.com/mnist-dataset-introduction-1138.html>

Classification

Averaged Perceptron	데이터 셋이 간단한 경우 정확도 < 속도	4
Bayes Point Machine	오버피팅이 우려되는 경우 간단한 셋팅	3
Boosted Decision Tree	높은 정확도가 필요한 경우 메모리 사용률 높음	6
Decision Forest	양상을 모델에 적합 효율적인 메모리 사용	6
Decision Jungle	뛰어난 일반화 성능 효율적인 메모리 사용 다소 긴 훈련 시간	6

SVM(Support Vector Machine)	Feature가 많은 경우 속도 > 정확도	5
Locally Deep SVM	Feature가 많은 경우 SVM 보다 빠른 속도 작은 모델 용량	8
Logistic Regression	통계적 접근법 숫자 변수 활용	5
Neural Network	인공신경망 레이어 다양한 parameter 조절이 가능	9

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/machine-learning-initialize-model-classification>

**“All models are wrong
but some are useful”**

- George Box



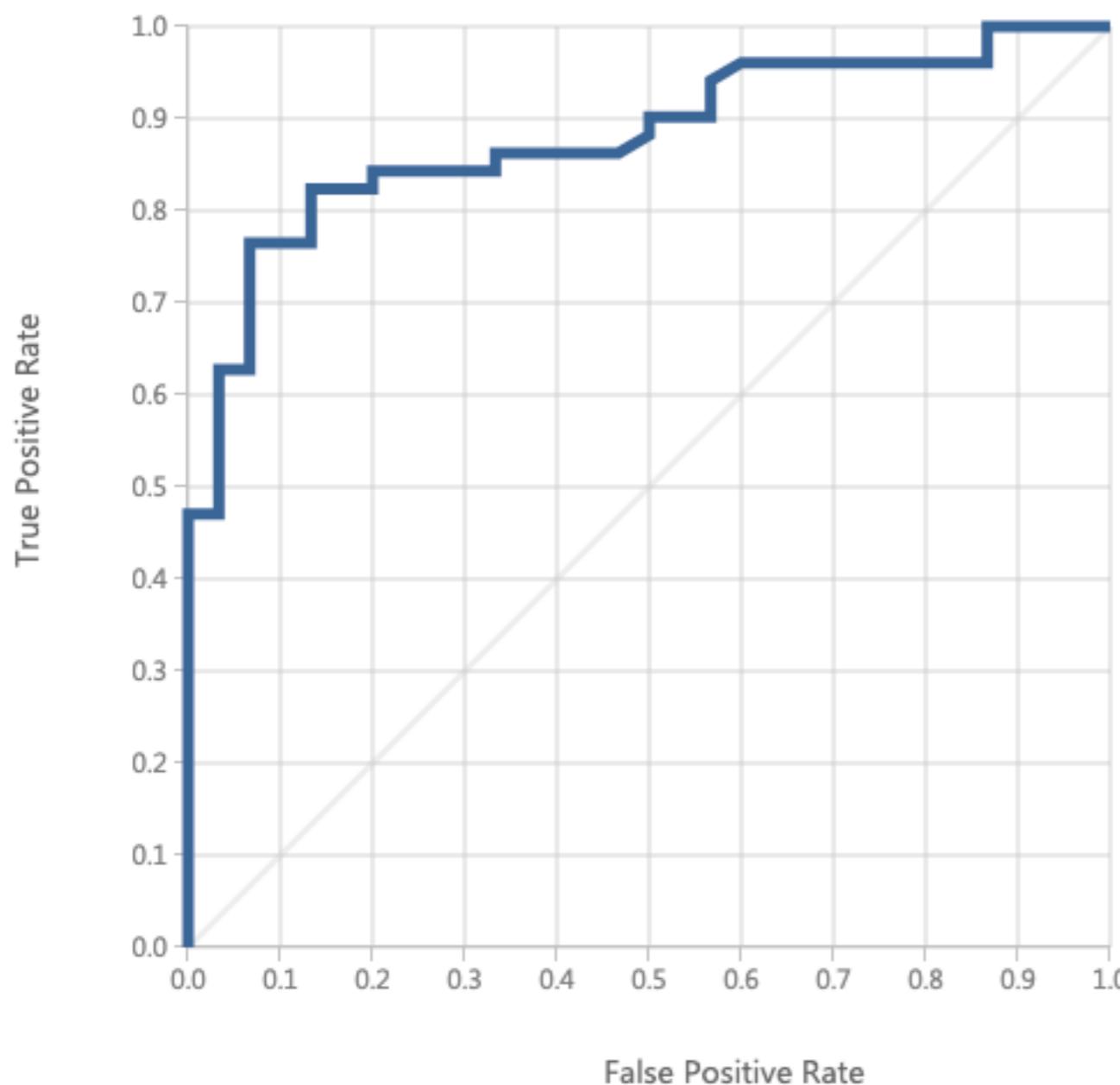
▲ Evaluate

Cross Validate Model

Evaluate Model

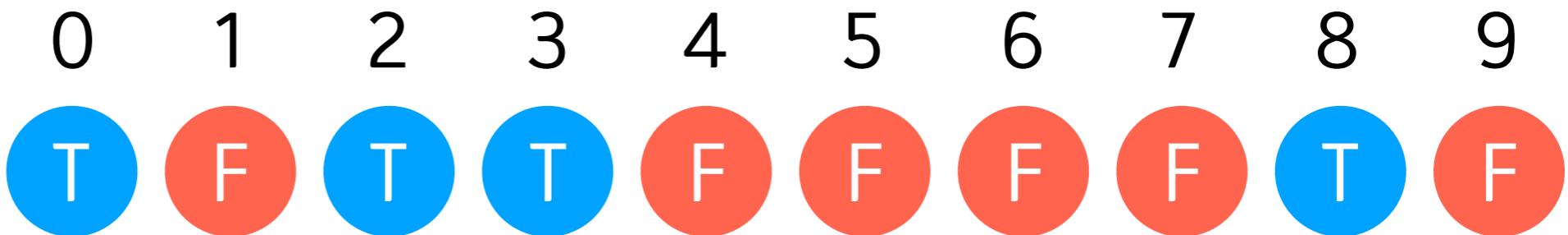
Evaluate Recommender

ROC PRECISION/RECALL LIFT



True Positive	False Negative	Accuracy	Precision
43	8	0.827	0.878
False Positive	True Negative	Recall	F1 Score
6	24	0.843	0.860
Positive Label		Negative Label	
True		False	

실측 값



예측 값



예측 값

실측 값	T	True Positive(TP)	3	False Negative(FN)	1
	F	False Positive(FP)	3	True Negative(TN)	3

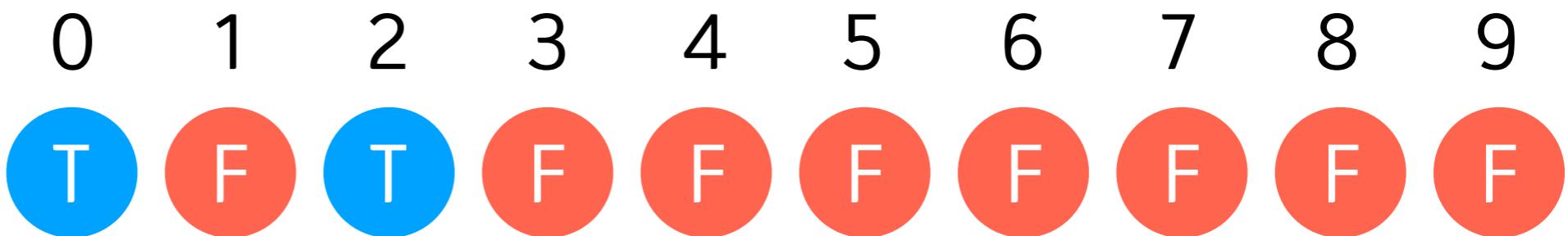
[기업 연계를 위한 AI-IoT 과정] AI | 첫번째 날 | 전미정

예측 값

	T	F
T	True Positive(TP) 3	False Negative(FN) 1
F	False Positive(FP) 3	True Negative(TN) 3

Accuracy	$TP+TN / TP+FP+FN+TN$	$(3 + 3) / 10 = 0.60$
Precision	$TP / TP+FP$	$3 / (3 + 3) = 0.50$
Recall	$TP / TP+FN$	$3 / (3 + 1) = 0.75$
F1 Score	$2 \times (precision * recall) / (precision + recall)$	$2 \times (0.5 * 0.75) / (0.5 + 0.75) = 0.6$

실측 값



예측 값



예측 값

실측 값	T	True Positive(TP)	2	False Negative(FN)	0
	F	False Positive(FP)	3	True Negative(TN)	5

예측 값

	T	F
T	True Positive(TP) 2	False Negative(FN) 0
F	False Positive(FP) 3	True Negative(TN) 5

Accuracy	$TP+TN / TP+FP+FN+TN$	$(2 + 5) / 10 = 0.70$
Precision	$TP / TP+FP$	$2 / (2 + 3) = 0.40$
Recall	$TP / TP+FN$	$2 / (2 + 0) = 1.00$
F1 Score	$2 \times (pre * recall) / (pre + recall)$	$2 * (0.4 * 1.0) / (0.4 + 1.0) = 0.57$

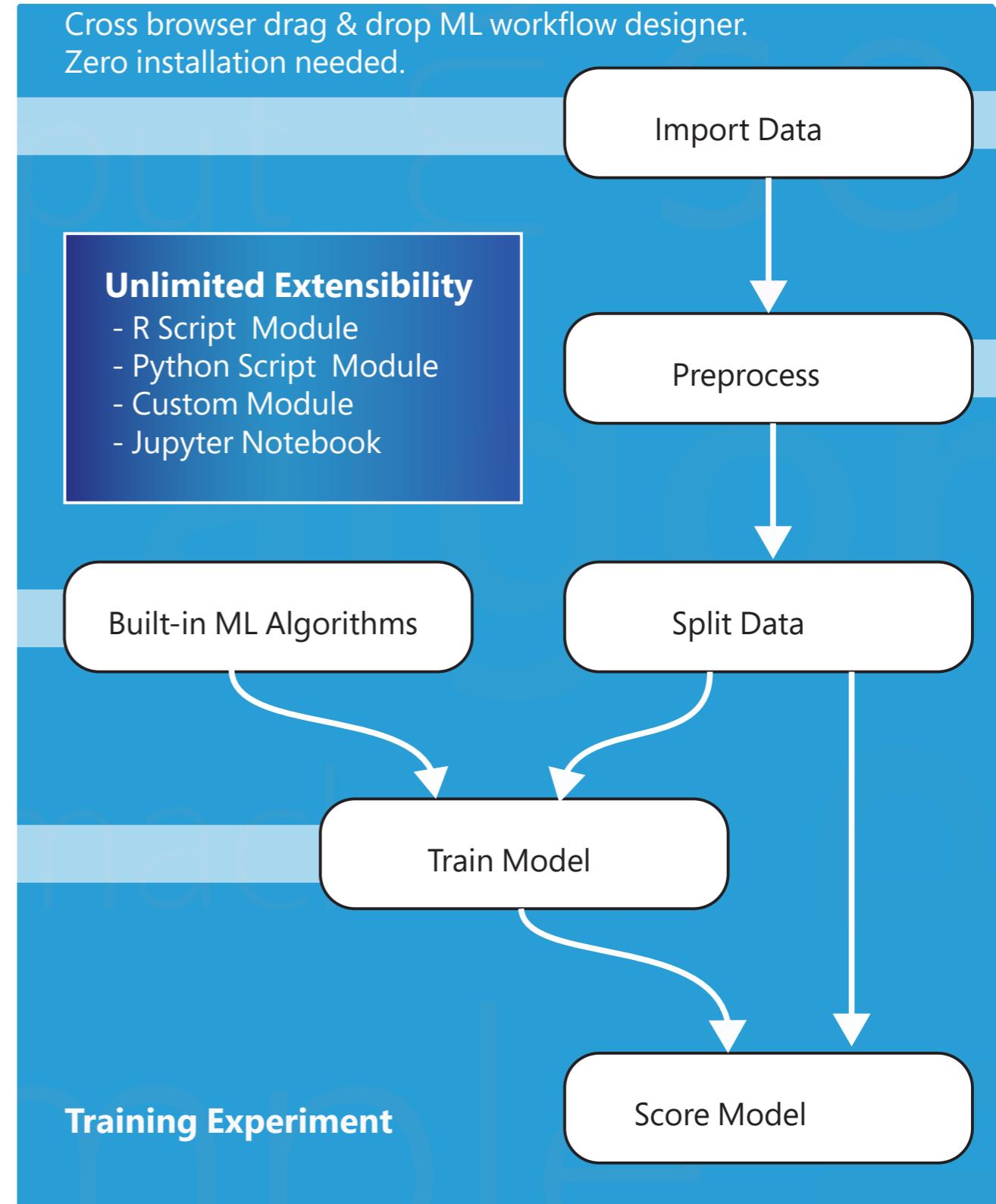
문제 정의

데이터셋 준비

모델 설정

모델 훈련 / 평가

모델 활용



Webs for more ML

- 캐글 데이터셋
- UCI 데이터셋
- Azure 갤러리
- Azure ML Studio에서 R 사용하기 실습
- Azure Machine Learning Service
- Azure Notebook
- Azure Cognitive Service

Books for more ML

입문자가 읽기 좋은 머신러닝 책

- 핸즈온 머신러닝 오렐리앙 제롬(한빛미디어)
- 처음 배우는 머신러닝 김승연, 정용주(한빛미디어)
- 모두의 딥러닝 조태호(길벗)

그리고 딥러닝에 대해 더 공부하고 싶다면

- 블록과 함께하는 파이썬 딥러닝 케라스 이야기 김태영(디지털북스)
- 케라스 창시자에게 배우는 딥러닝 프랑소와 쏠레(길벗)
- 밑바닥부터 시작하는 딥러닝 사이토 고키(한빛미디어)

수고하셨습니다



ninevincentg@gmail.com