Introduction to Artificial Intelligence [AICS223]

AI, Data & Preprocessing (W02)

고려대학교

인공지능사이버보안학과

KOREA UNIVERSITY

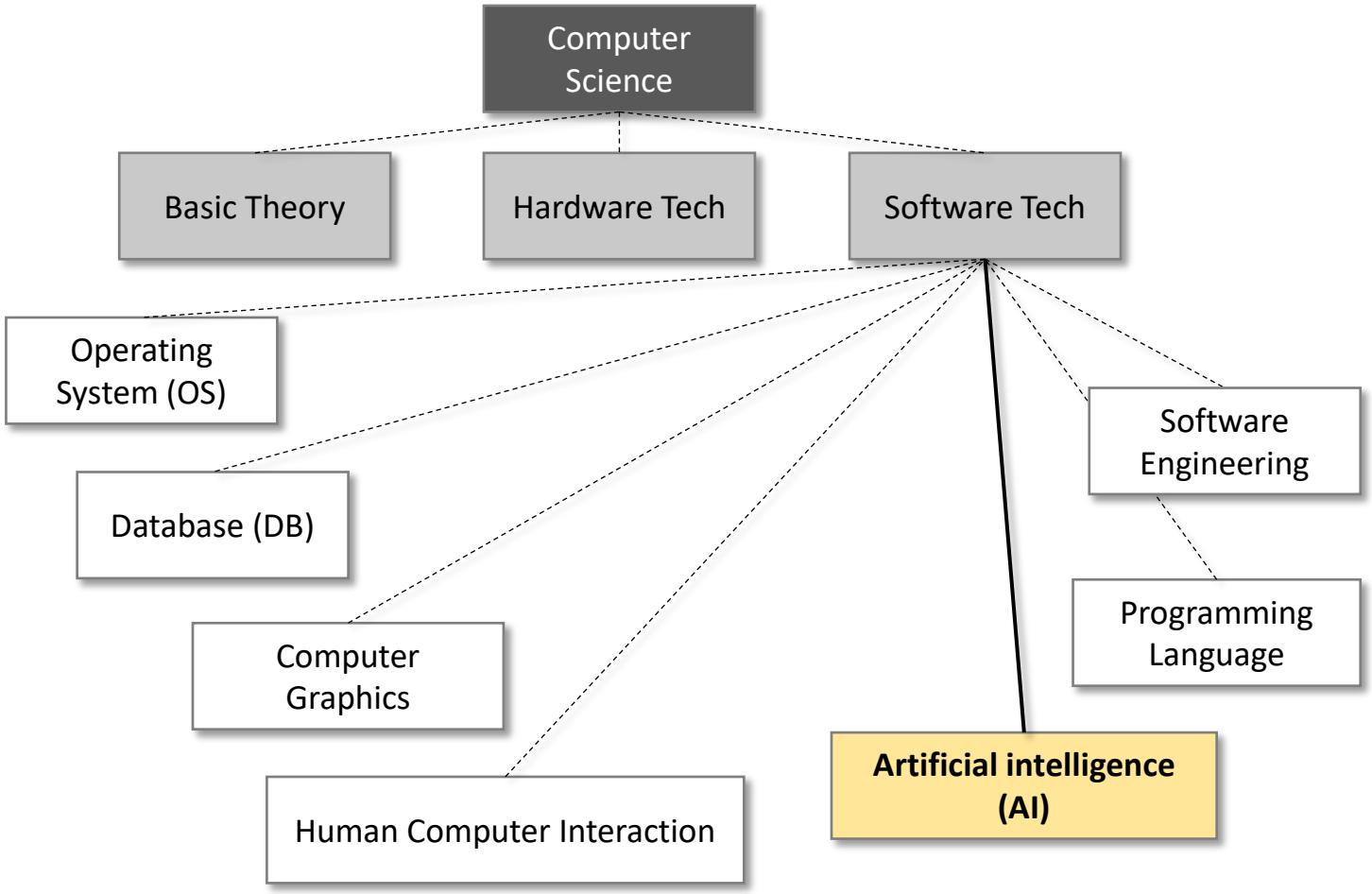
Prof. Mee Lan Han (aeternus1203@gmail.com)

CONTENTS

Artificial intelligence (AI)

■ What is Artificial Intelligence?

- □ 생물의 구조나 지적 활동에서 힌트를 얻은 소프트웨어 기술
- □ 인간의 **학습능력**과 **추론능력**, **언어이해능력**을 컴퓨터 프로그램으로 구현하는 기술



Artificial intelligence (AI)

■ What is Artificial Intelligence?

□ 컴퓨터가 학습하고 생각하여 스스로 판단할 수 있도록 만드는 기술로 발전

관점	개념 설명
사전적 개념	철학적인 개념으로, 지성을 갖춘 존재 또는 시스템에 의해 만들어진 인공적인 지능을 의미
전통적 개념	컴퓨터가 인간의 지능적인 행동을 모방할 수 있도록 하는 소프트웨어로, 인간이 가진 지적 능력의 일부 또는 전체를 구현한 것
기술적 개념	인간의 지능으로 할 수 있는 사고, 학습, 자기계발 등을 컴퓨터가 할 수 있도록 하는 방법을 연구하는 컴퓨터공학 및 정보기술의 한 분야

Artificial intelligence (AI)

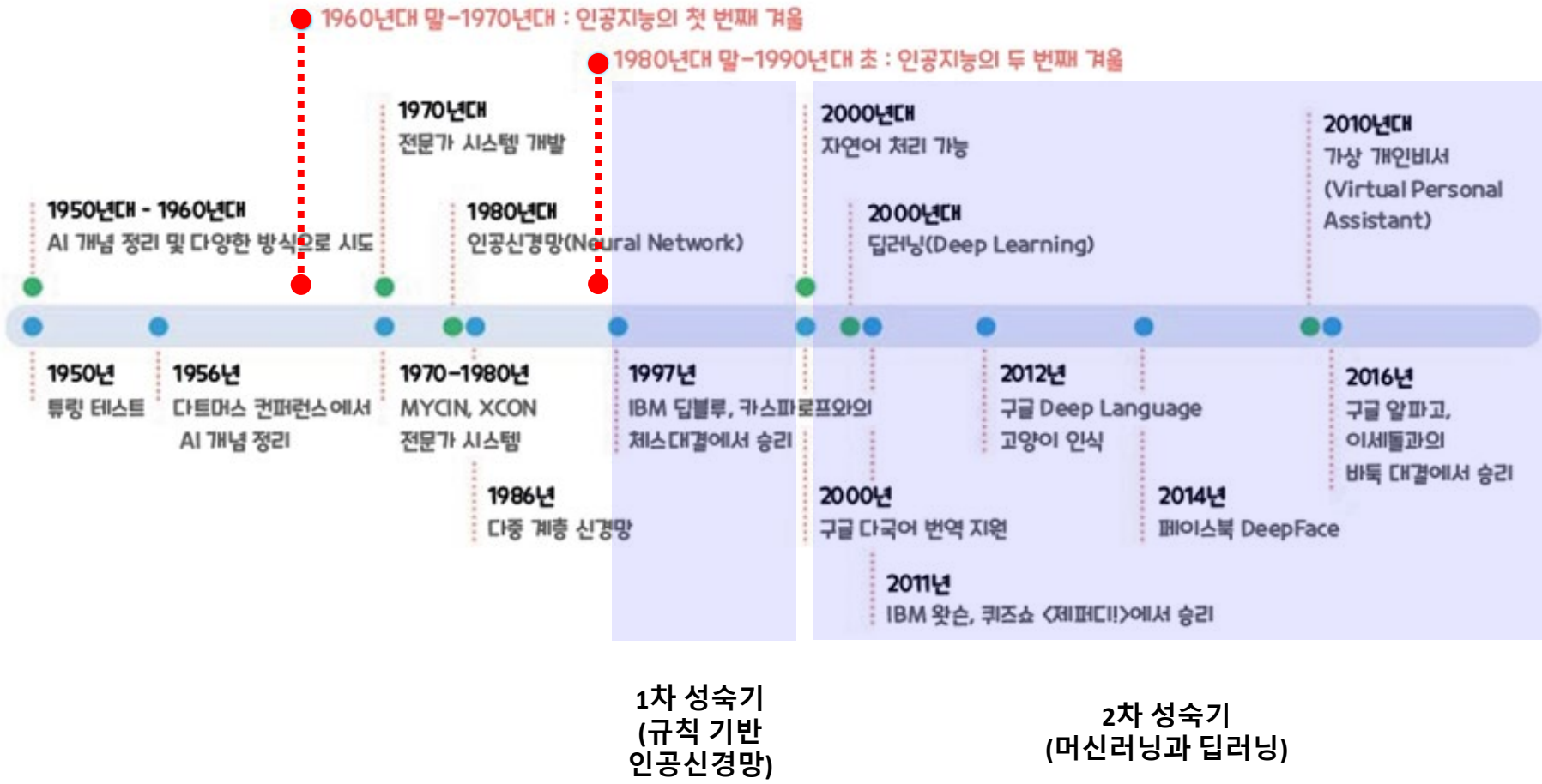
■ What is Artificial Intelligence?

□ 인공지능의 시작부터 최근까지의 연대를 시간 흐름순으로 정리

Artificial intelligence (AI)

■ What is Artificial Intelligence?

□ 인공지능의 시작부터 최근까지의 연대를 시간 흐름순으로 정리



Artificial intelligence (AI)

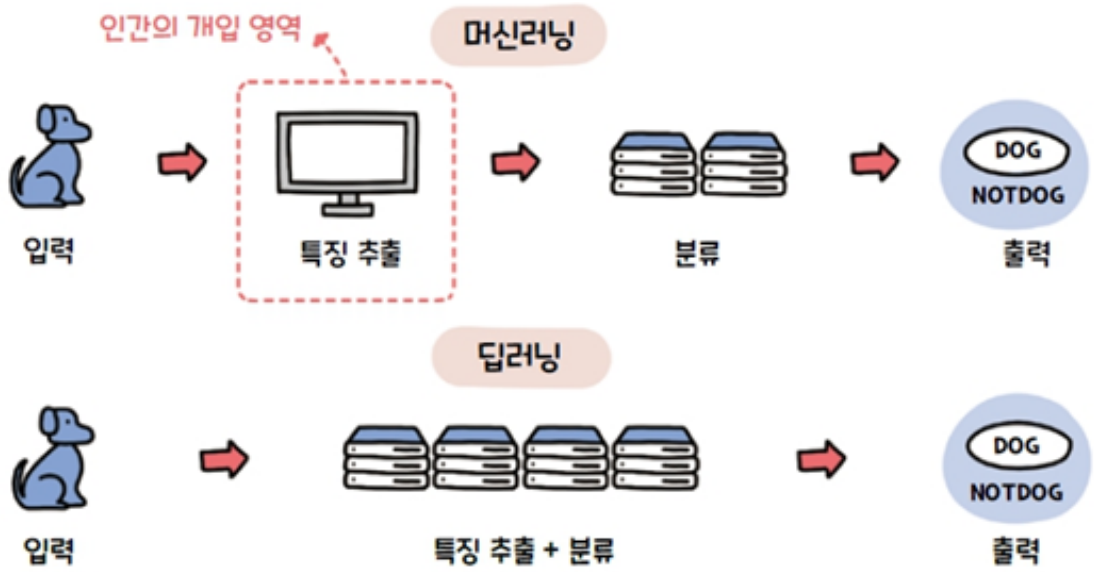
■ What is Artificial Intelligence?

□ 머신러닝

- **보유한 지식 (학습 데이터)를 기반으로 목적에 맞는 모델 생성 후 추론 및 탐색하는 인공지능**

□ 딥러닝

- 머신러닝보다 발전하여 **특정 데이터는 전처리 과정 없이 학습가능한 수준의 알고리즘 보유**
- 특정 영역의 데이터는 (이미지, 문장, 신호 등)는 모델 생성 시 인간의 개입 (전처리 수행) 없이 특징을 추출하고 모델링 할 수 있는 알고리즘 보유

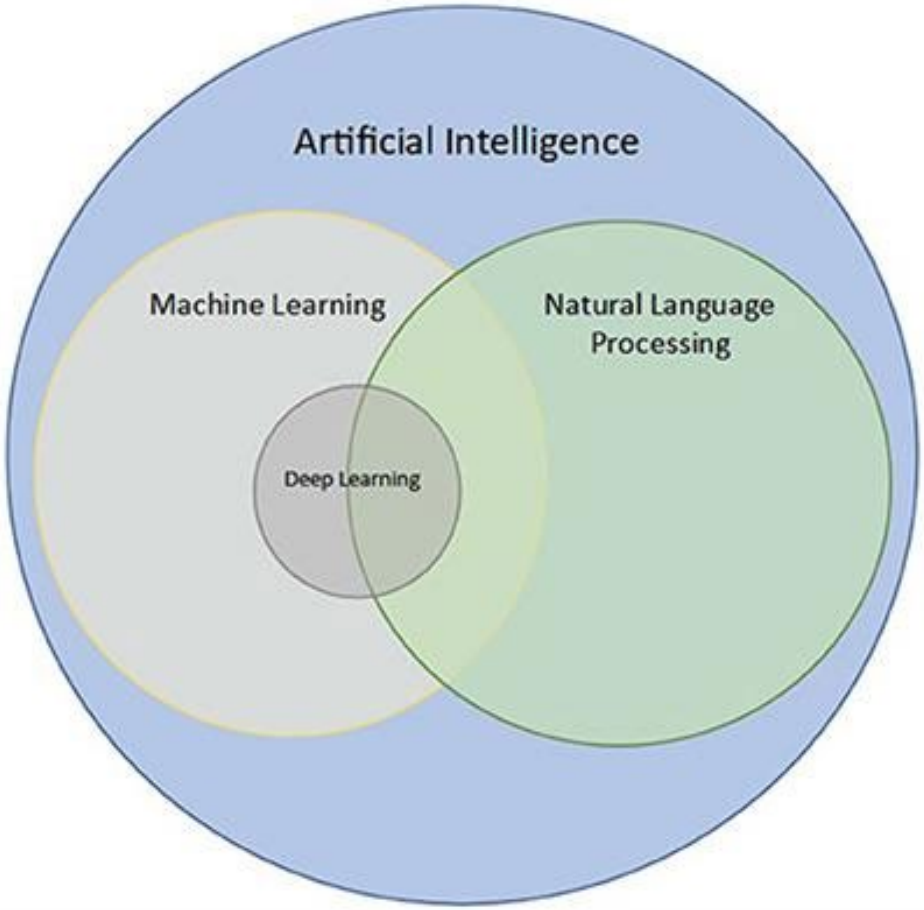


필요에 의해 수행해야하는 특징추출과 전처리 과정의 어려움을 덜어줌

Artificial intelligence (AI)

■ Fields of Artificial Intelligence

□ 서로 밀접한 관계를 맺고 있는 인공지능 분야의 기술



Artificial intelligence (AI)

■ Fields of Artificial Intelligence

□ 서로 밀접한 관계를 맺고 있는 인공지능 분야

Artificial Intelligence

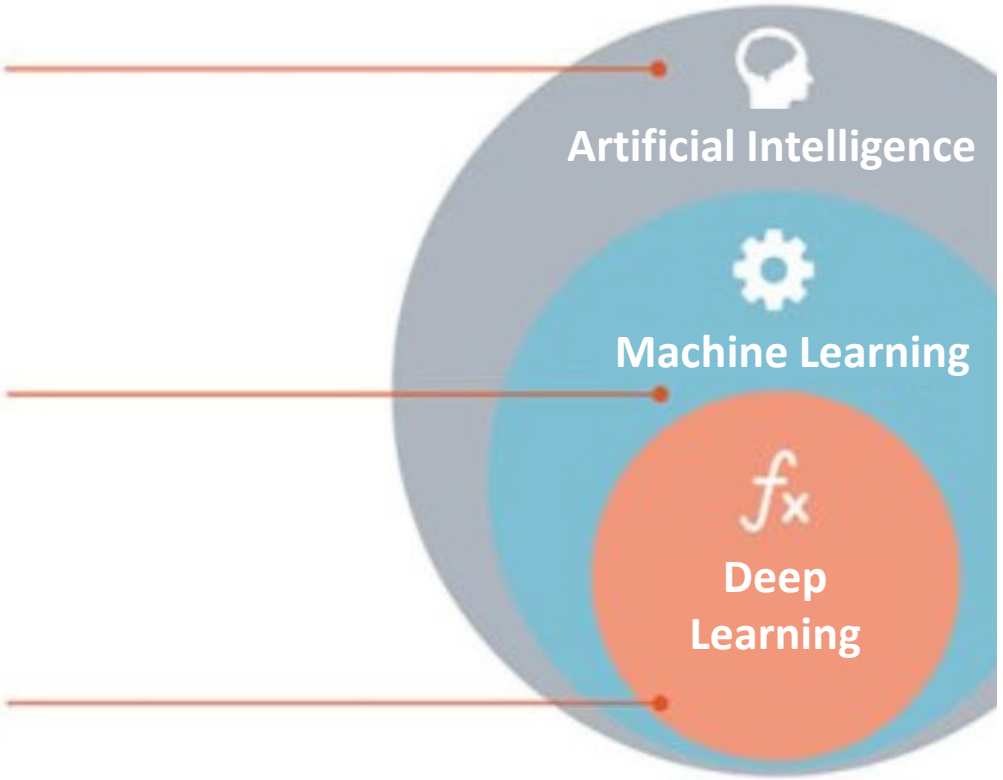
- Natural Language Processing
- Image Recognition
- Swarm Intelligence

Machine Learning

- Decision Tree
- Random Forest
- Support Vector Machine

Deep Learning

- Convolution Neural Network, CNN
- Generative Adversarial Network, GAN
- Recurrent Neural Network, RNN
- Long Short Term memory, LSTM



A step towards artificial intelligence, machine learning and deep learning, Soledad Musella Rubio, May 28, 2020

Artificial intelligence (AI)

■ Fields of Artificial Intelligence

□ 서로 밀접한 관계를 맺고 있는 인공지능 분야

Artificial Intelligence

- Natural Language Processing
- Image Recognition
- Swarm Intelligence

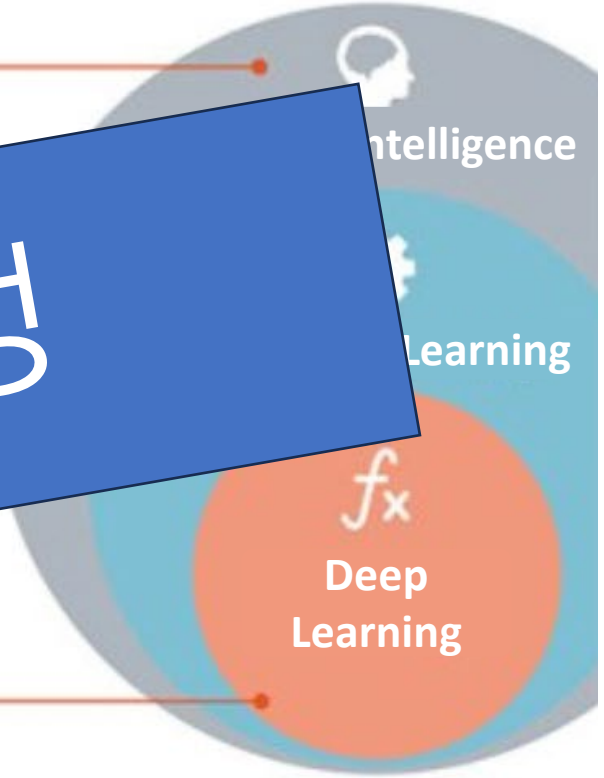
Machine Learning

-
-
-

Deep L

- Convolution Neural Network, CNN
- Generative Adversarial Network, GAN
- Recurrent Neural Network, RNN
- Long Short Term memory, LSTM

Model 생성



A step towards artificial intelligence, machine learning and deep learning, Soledad Musella Rubio, May 28, 2020

Artificial intelligence (AI)

■ What is modeling in artificial intelligence (AI)?

□ **Discriminator(Discriminative) 판별모델** vs. Generator(Generative) 생성모델

- 일반적으로 규칙과 논리에 기반
- 특정 패턴이나 특징을 학습하여 분류, 예측, 판별하는 모델을 도출함
- Class 차이에 주목하여 어떤 Class에 들어가야 할지 결정해 주는 모델

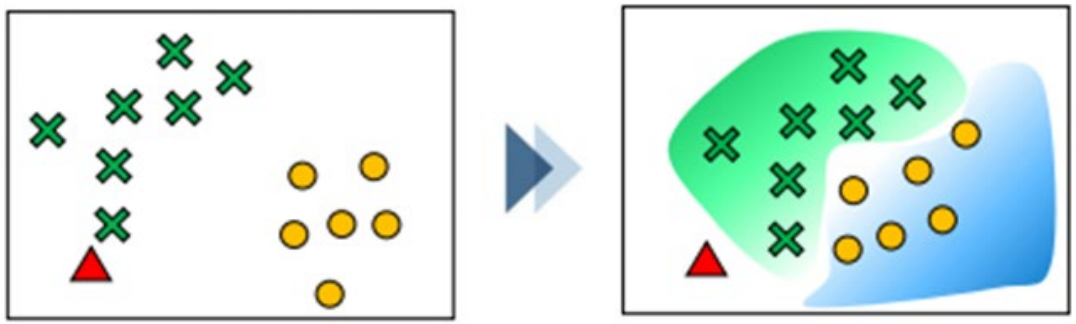
✓ 판별 모델은 정답 (Ground Truth, GT)이 존재하므로 모델의 출력을 정답과 비교하기 용이
✓ 범주형 데이터를 사용하는 경우 (분류 모델)
✓ 연속형 데이터를 사용하는 경우 (회귀 분석 모델)

Artificial intelligence (AI)

■ What is modeling in artificial intelligence (AI)?

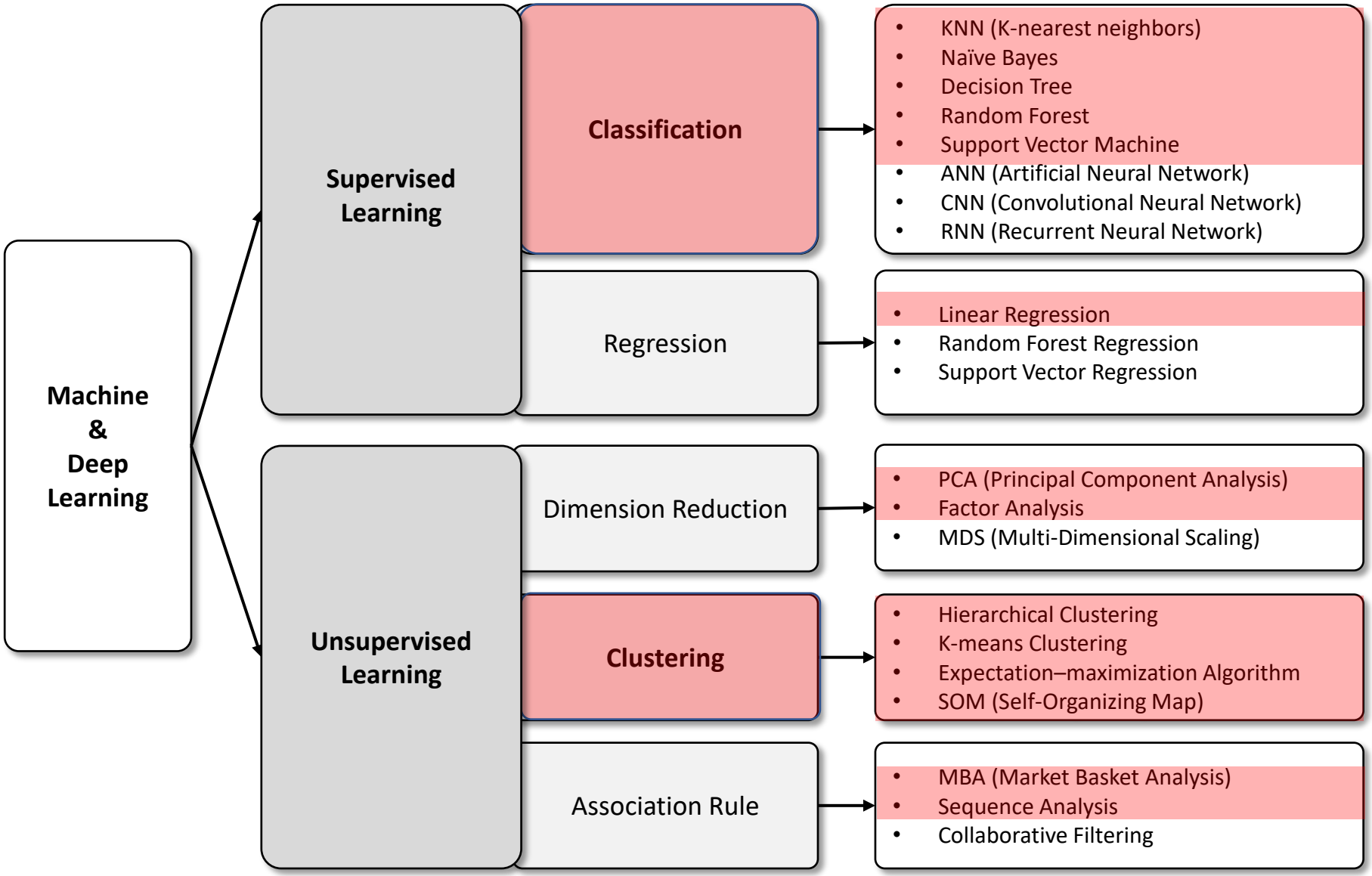
□ Discriminator(Discriminative) 판별모델 vs. **Generator(Generative) 생성모델**

- 학습 데이터를 기반으로 **새로운 데이터 샘플을 생성**하는데 사용되는 기계 학습 모델의 한 유형
- 학습 데이터의 분포를 따르는 유사한 데이터를 생성하는 모델

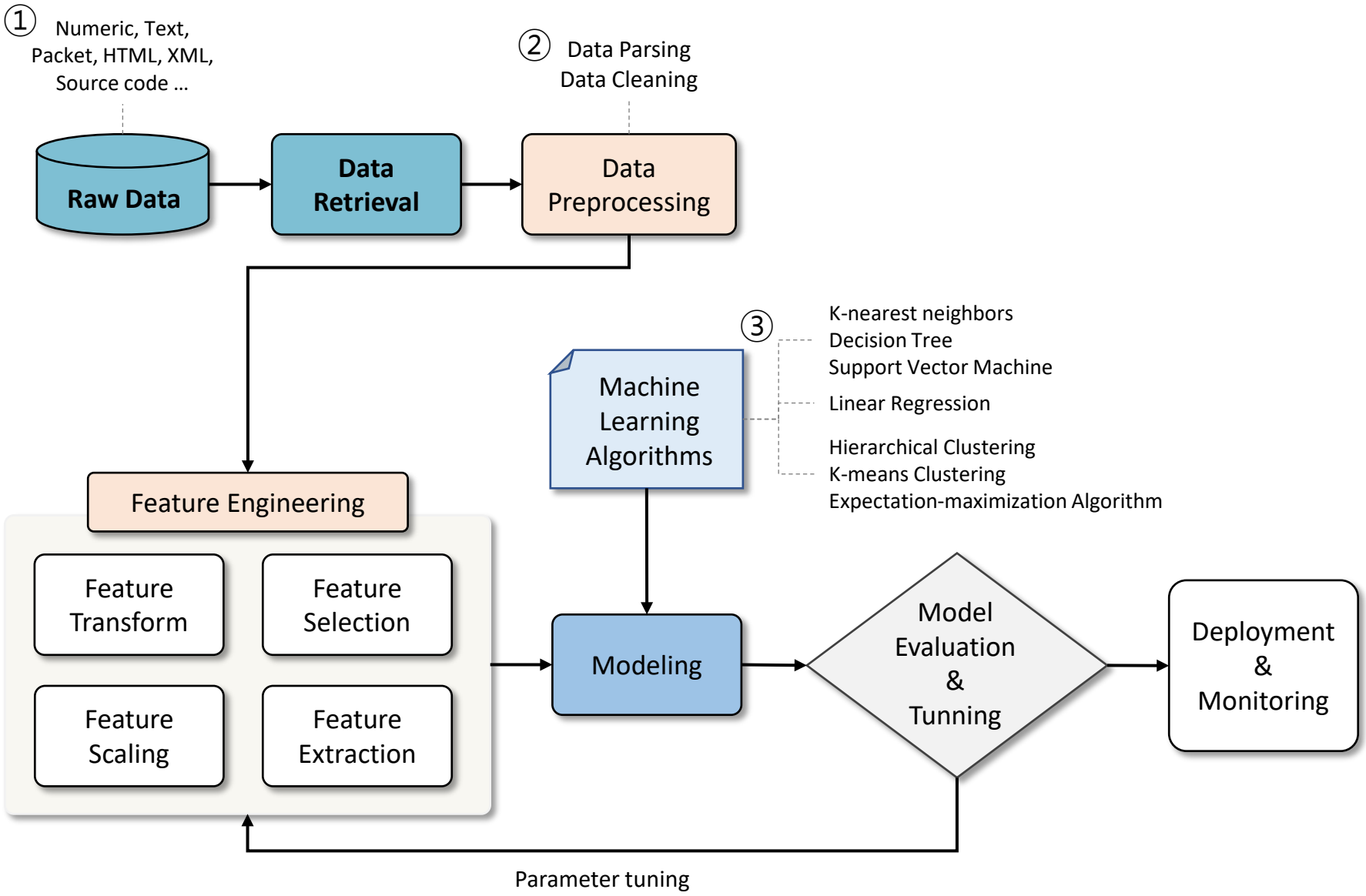


✓ 판별 모델과 달리 비교할 정답이 존재하지 않아 결과를 직접적으로 비교할 대상이 없음
✓ 훈련 데이터를 정답으로 사용할 경우, 훈련 데이터를 그대로 복제하는 현상 발생할 수 있음
✓ 개인의 주관이 개입되지 않아야 하고, 연구자들이 공감할 수 있는 객관적인 지표가 필요함

Machine Learning



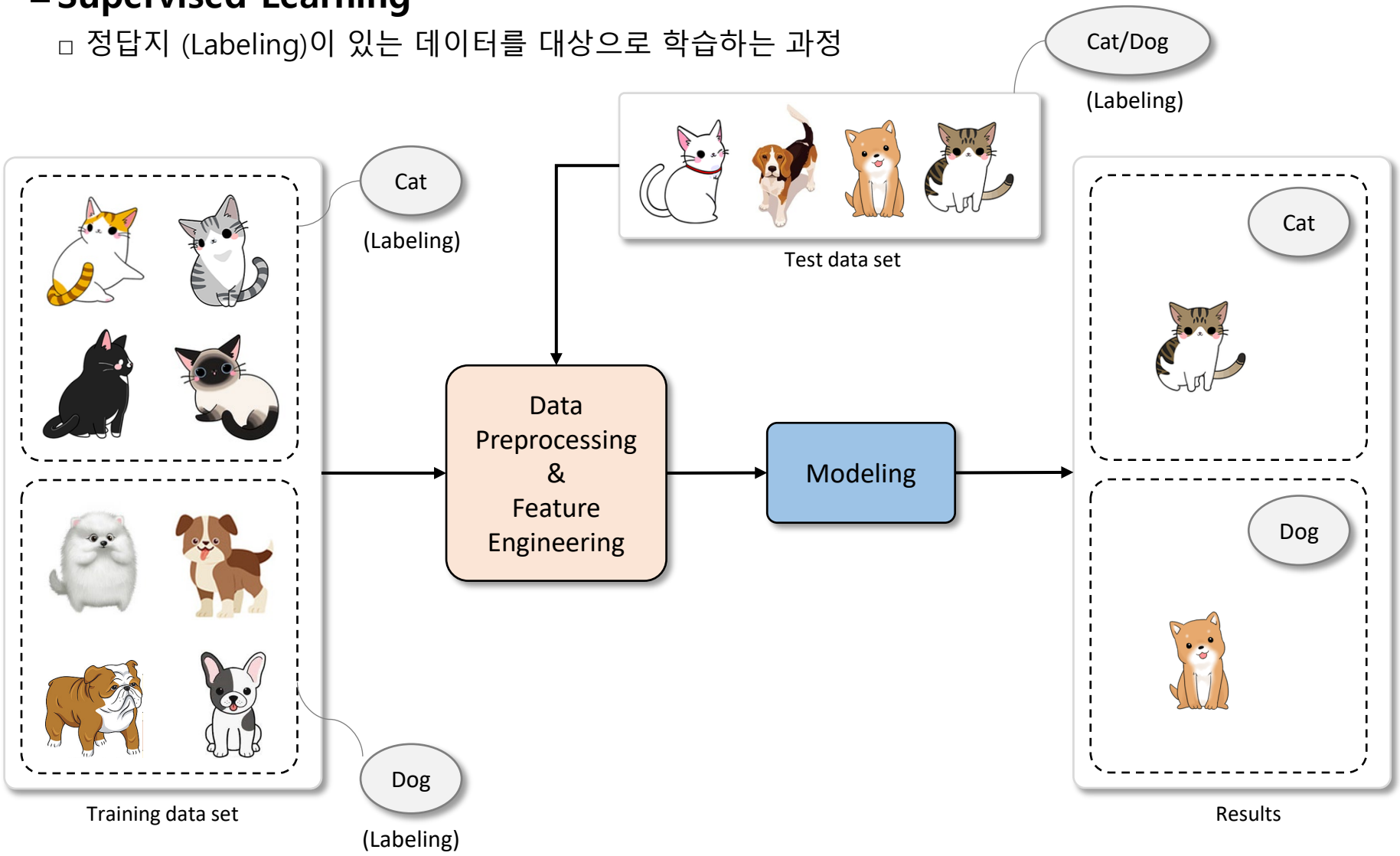
Machine Learning Pipeline



Machine Learning

■ Supervised Learning

□ 정답지 (Labeling)이 있는 데이터를 대상으로 학습하는 과정



Machine Learning

■ Supervised Learning

□ 정답지 (Labeling)이 있는 데이터를 대상으로 학습하는 과정

□ 실제 데이터에서의 정답지 표현 (Class라고 기재된 레이블)

Features

ID	Mean	Var	Q1	Q3	IQR	Skew	Kurt	Class
0x545	19.5	3	18	21	3	0	-2	Fuzzy
0x131	19	3	18	19.5	1.5	0.7071068	-1.5	Fuzzy
0x350	19.75	4.25	18	21.25	3.25	0.1153172	-1.847751	Fuzzy
0x370	19.5	3.6666667	18	20.5	2.5	0.4933822	-1.371901	Fuzzy
0x2B0	19.75	4.25	18	21.25	3.25	0.1153172	-1.847751	Fuzzy
0x329	19.5	3	18	21	3	0	-2	Fuzzy
0x1F1	29	271	24	38.5	14.5	-0.704173	-1.5	Fuzzy
0x002	19.25	4.25	17.75	21	3.25	-0.115317	-1.847751	Fuzzy
0x4B1	38.5	0.5	38.25	38.75	0.5	0	-2	Fuzzy
0x153	19.25	0.25	19	19.25	0.25	1.1547005	-0.666667	Fuzzy
0x2A0	19.75	4.25	18	21.25	3.25	0.1153172	-1.847751	Fuzzy
0x260	19.75	4.25	18	21.25	3.25	0.1153172	-1.847751	Fuzzy
0x002	21.666667	108.33333	17.5	27.5	10	-0.528005	-1.5	Replay
0x153	20	169	16	27.5	11.5	-0.702403	-1.5	Replay
0x2A0	21	144.66667	16.75	29.25	12.5	-0.812266	-1	Replay
0x260	21	116.66667	15.25	29.25	14	-0.467325	-1.394971	Replay
0x440	18	52	14	21	7	0.4703305	-1.5	Replay
0x140	11.5	47.9	7.5	16	8.5	0.0416349	-1.152793	Replay
0x18F	18.666667	96.333333	13	21.5	8.5	0.7071068	-1.5	Replay
0x316	20.333333	161.33333	13	24	11	0.7071068	-1.5	Replay
0x2C0	17.8	151.7	5	25	20	-0.323556	-1.769181	Replay
0x43F	18	43	14.5	21	6.5	0.2736425	-1.5	Replay
0x430	40.5	420.5	33.25	47.75	14.5	0	-2	Replay
0x545	19.333333	225.33333	12	27	15	-0.081428	-1.5	Replay
0x131	16	73	16	17	1	-0.38719	-0.466457	Replay
0x350	21.333333	126.33333	16.5	27.5	11	-0.411326	-1.5	Replay

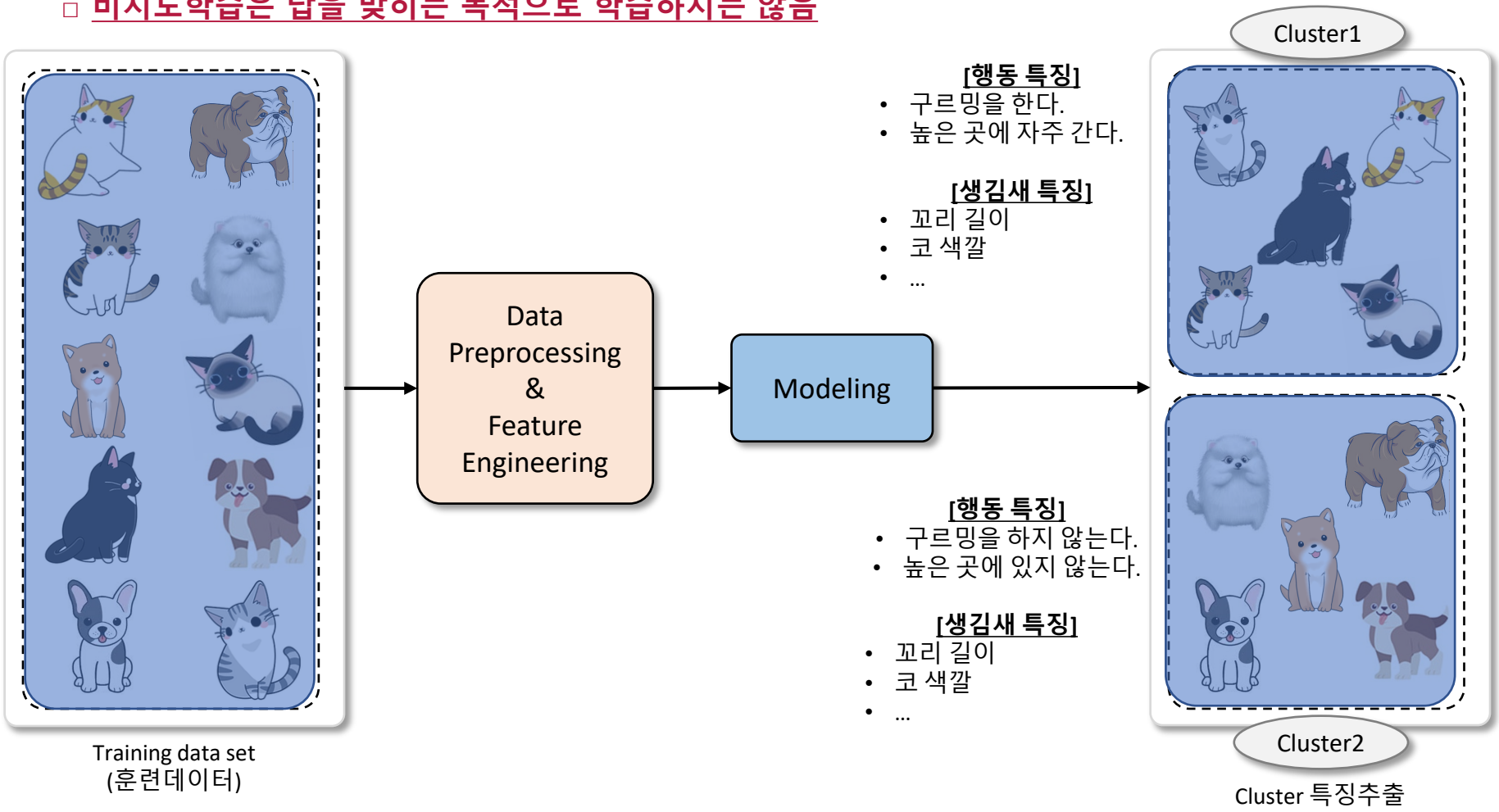
Labeling

Machine Learning

■ Unsupervised Learning

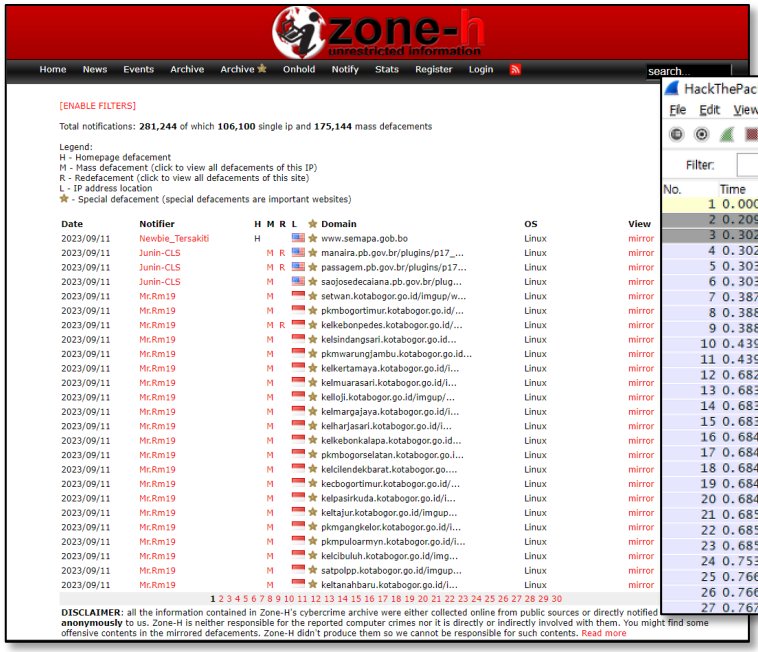
□ Labeling이 없는 데이터셋을 대상으로 **비슷한 특징이 있는지 알아보기 위해 군집화함으로써** 새로운 결과를 추론하는 학습과정

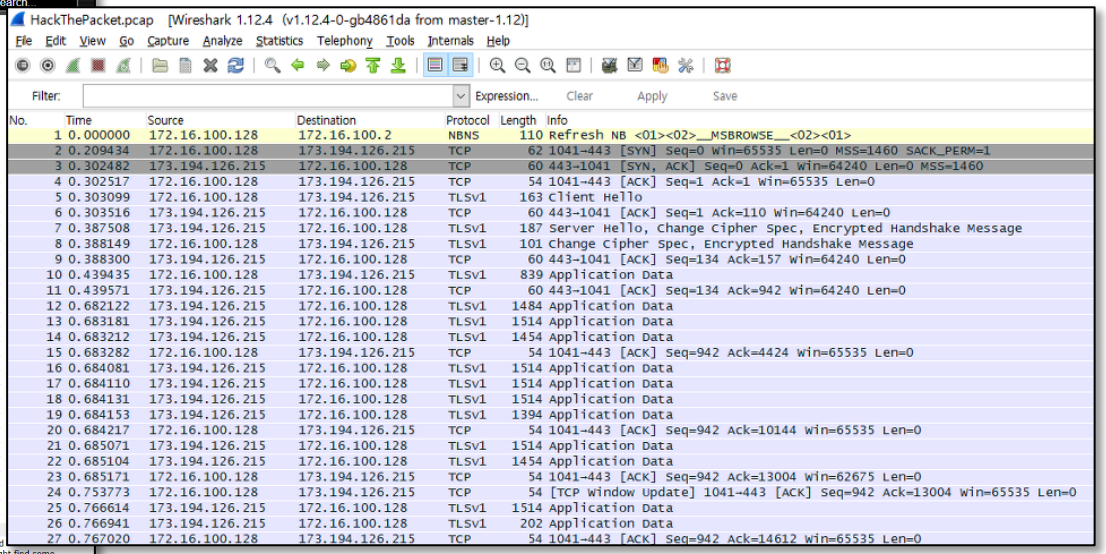
□ **비지도학습은 답을 맞히는 목적으로 학습하지는 않음**



Machine Learning

■ Unsupervised Learning





_ws.col.UT	_ws.col.Pr	ip.src	ip.dst	tcp.srcport	tcp.dstport	tcp.len	tcp.seq	tcp.ack
11:58:58	CDP							
11:59:00	LDAP	192.168.10	192.168.10	33898	389	403	1	1
11:59:00	TCP	192.168.10	192.168.10	33898	389	403	1	1
11:59:00	LDAP	192.168.10	192.168.10	389	33898	316	1	404
11:59:00	TCP	192.168.10	192.168.10	389	33898	316	1	404
11:59:00	TCP	192.168.10	192.168.10	33898	389	0	404	317
11:59:00	TCP	192.168.10	192.168.10	33898	389	0	404	317
11:59:00	LDAP	192.168.10	192.168.10	33904	389	403	1	1
11:59:00	TCP	192.168.10	192.168.10	33904	389	403	1	1
11:59:00	LDAP	192.168.10	192.168.10	389	33904	316	1	404
11:59:00	TCP	192.168.10	192.168.10	389	33904	316	1	404
11:59:00	TCP	192.168.10	192.168.10	33904	389	0	404	317
11:59:00	TCP	192.168.10	192.168.10	33904	389	0	404	317

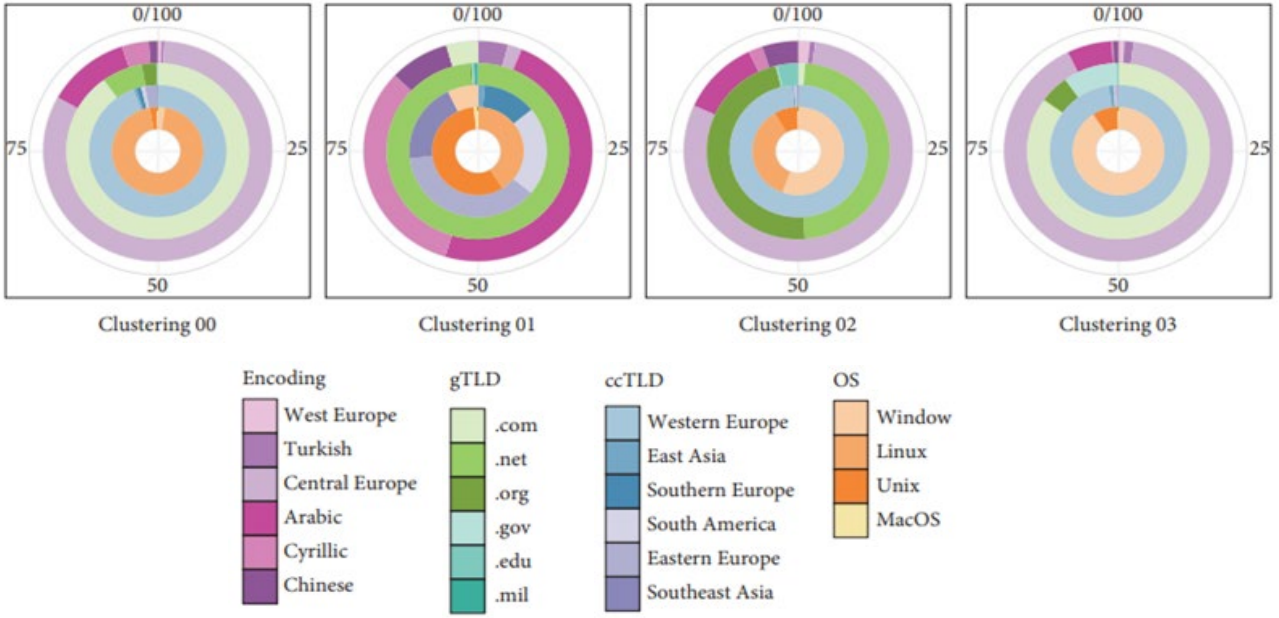
source ip	source port	destination	destination	attack type
172.16.0.1	32802	192.168.10	80	brute force
172.16.0.1	32822	192.168.10	80	brute force
172.16.0.1	32860	192.168.10	80	brute force
172.16.0.1	32880	192.168.10	80	brute force
172.16.0.1	32900	192.168.10	80	brute force
172.16.0.1	32938	192.168.10	80	brute force
172.16.0.1	32958	192.168.10	80	brute force
172.16.0.1	33016	192.168.10	80	brute force
172.16.0.1	33036	192.168.10	80	brute force

Machine Learning

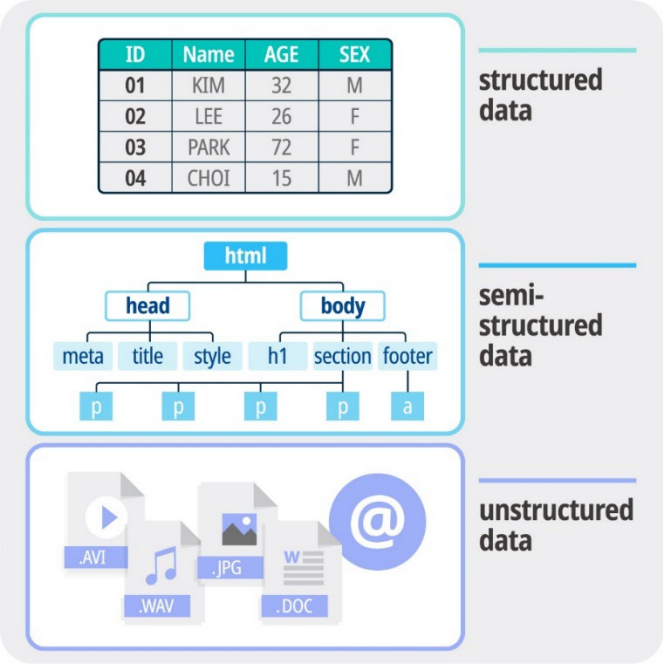
■ Unsupervised Learning

- □ 비지도학습은 데이터의 숨겨진 특징 (feature)이나 구조 or 패턴을 발견하는데 사용
- □ 범죄 데이터 분석
  - 범죄자 그룹핑 or 범죄자 수사범위 축소 (범죄자를 찾는 것이 목표가 아님)

No	Date	Notify	URL	gTLD	TLD_Groupir	ccTLD	OS_Grouping	OS	Encoding_Grouping	Encoding	IP	WebServer	Lang
1	2002-12-15	Affix	corp1ontheweb.com	com			Window	Win 2000	Taiwanese	big5	199.231.130.38	Unknown	Courier
2	2002-12-16	gB	www.albemarlebulkheads.com	com			Linux	Linux	Taiwanese	big5	208.155.64.39	Unknown	Courier
3	2002-12-14	Red Eye	www.kanggyeong.ms.kr		EastAsia	kr	Window	Win NT9x	Taiwanese	big5	211.251.36.98	Unknown	Courier
4	2009-12-12	spo0feR	apmab.ap.nic.in		SouthAsia	in	Linux	Linux	Chinese	gb2312	164.100.12.136	Apache	Courier
5	2009-06-16	AYYILDIZ	www.bikinialley.com	com			Window	Win 2003	Chinese	gb2312	216.139.216.143	IIS/6.0	Courier
6	2009-06-24	ir4dex	neyla.net	net				Unknown	WestEurope	iso-8859-1	119.235.22.31	Apache	Courier
7	2003-10-11	ghost_x	www.macvalleybaptist.org	org			Linux	Linux	WestEurope	iso-8859-1	12.129.206.109	Apache	Courier
8	2002-11-22	king9x	www.mmavideogames.com	com			Linux	Linux	WestEurope	iso-8859-1	12.129.206.109	Unknown	Courier

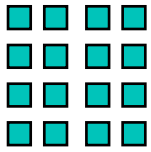


Data Characteristics

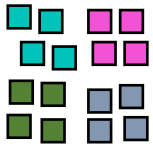


* 한국정보통신기술협회 정보통신용어사전 이미지 참조

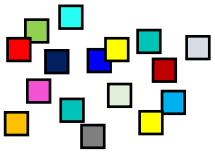
"고정된 틀& 패턴"



"일종의 패턴 "



"틀& 패턴 불규칙"



■ 정형 데이터

- □ 관계형 데이터베이스 시스템의 테이블과 같이 **고정된 컬럼에 저장되는 데이터와 파일**
- □ 지정된 **행과 열에 의해 데이터의 속성이 구별**되는 스프레드시트 형태의 데이터

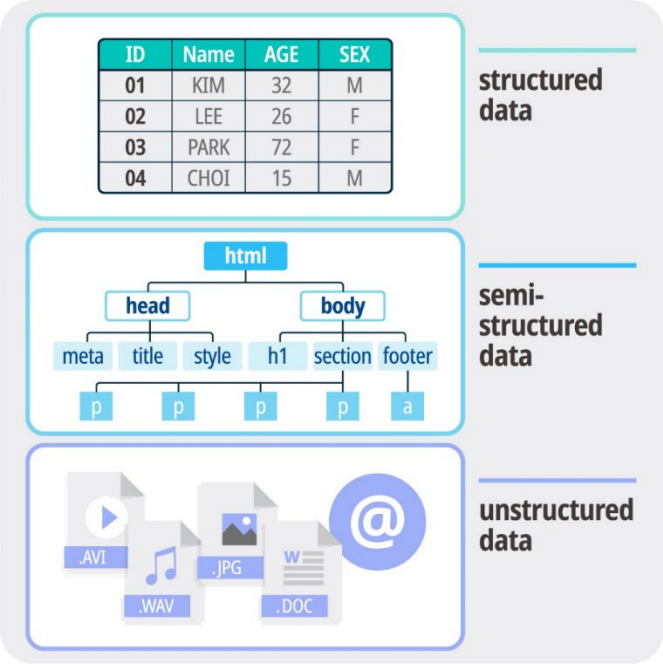
■ 반정형 데이터

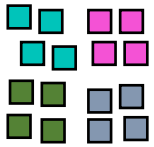
- □ 해당 파일 파싱 후 메타구조를 갖는 정형데이터 형태의 테이블 구조로 재생성
- □ 보통 API 형태로 제공되기 때문에 데이터 처리 기술이 요구
- □ HTML, XML, JSON 등

■ 비정형 데이터

- □ 고정된 필드에 저장되어 있지 않은 데이터
- □ HEX (16진수), 이미지, 비디오 스트림, 오디오 데이터, 자연어
- □ 수집 난이도가 높음

Data Characteristics





■ 데이터 종류

□ Numeric (숫자)
- 소수값, 2진수, 10진수, 16진수

□ Character (문자)
- string, symbolic constant

□ Date and Time (날짜/시간)
- date, time, timestamp 등

□ TRUE/FALSE (논리 연산값)

* 한국정보통신기술협회 정보통신용어사전 이미지 참조

Data Characteristics

■ Structured Data

□ 관계형 데이터베이스 시스템의 테이블과 같이 **고정된 컬럼에 저장되는 데이터와 파일**

□ 지정된 **행과 열에 의해 데이터의 속성이 구별**되는 스프레드시트 형태의 데이터

스키마에 의해 정의된 컬럼

Column 1	Column 2	…	Column N-1	Column N
data	data	data	data	data
data	data	data	data	data
data	data	data	data	data
data	data	data	data	data

data: 컬럼에 의해 정의된 데이터

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5.0	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa

Class
(라벨링)

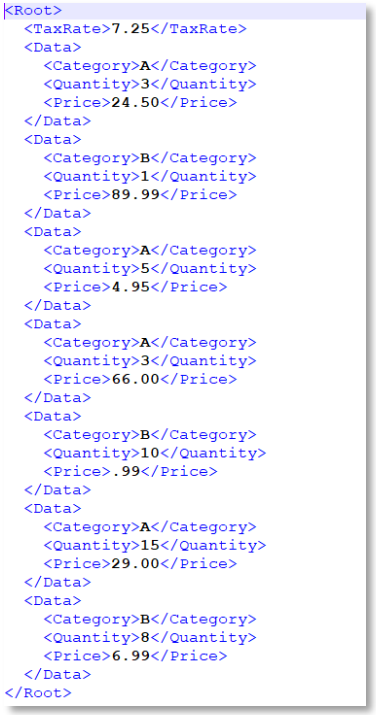
Data Characteristics

■ Semi Structured Data

- □ 스키마에 해당하는 메타 데이터가 데이터 내부에 존재
- □ 보통 API 형태로 제공되기 때문에 데이터 처리 기술이 요구
- □ 해당 파일을 파싱하여 메타구조를 갖는 정형데이터 형태의 테이블 구조로 재생성

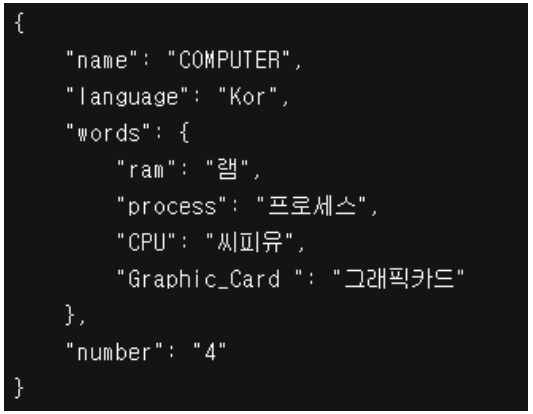
```
<div id="ntp-contents">
  <div id="logo">
    <!-- The logo that is displayed in the absence of a doodle. -->
    <div id="logo-default" title="Google"></div>
    <!-- Logo displayed when theme prevents doodles. Doesn't fade. -->
    <div id="logo-non-white" title="Google"></div>
    <!-- A doodle, if any: its link and image. -->
    <div id="logo-doodle">
      <div id="logo-doodle-container">
        <div id="logo-doodle-wrapper">
          <button id="logo-doodle-button">
            <img id="logo-doodle-image" tabindex="-1"></img>
          </button>
        </div>
      </div>
      <iframe id="logo-doodle-iframe" scrolling="no"></iframe>
    </div>
  </div>

  <div id="fakebox-container" hidden>
    <div id="fakebox">
      <div class="search-icon"></div>
      <div id="fakebox-text"></div>
      <input id="fakebox-input" autocomplete="off" tabindex="-1" type="url"
          aria-hidden="true">
      <div id="fakebox-cursor"></div>
      <button id="fakebox-microphone" class="microphone-icon" hidden></button>
    </div>
  </div>

  <div id="realbox-container" >
    <div id="realbox-input-wrapper">
      <div id="realbox-icon" data-default-icon="search.svg">
      </div>
      <input id="realbox" type="search" autocomplete="off" spellcheck="false"
          aria-live="polite" autofocus>
      <button id="realbox-microphone" class="microphone-icon" hidden></button>
      <div id="realbox-matches"></div>
    </div>
  </div>
```

HTML
URL 형태로 존재 or 접근



```
<Root>
  <TaxRate>7.25</TaxRate>
  <Data>
    <Category>A</Category>
    <Quantity>3</Quantity>
    <Price>24.50</Price>
  </Data>
  <Data>
    <Category>B</Category>
    <Quantity>1</Quantity>
    <Price>89.99</Price>
  </Data>
  <Data>
    <Category>A</Category>
    <Quantity>5</Quantity>
    <Price>4.95</Price>
  </Data>
  <Data>
    <Category>A</Category>
    <Quantity>3</Quantity>
    <Price>66.00</Price>
  </Data>
  <Data>
    <Category>B</Category>
    <Quantity>10</Quantity>
    <Price>.99</Price>
  </Data>
  <Data>
    <Category>A</Category>
    <Quantity>15</Quantity>
    <Price>29.00</Price>
  </Data>
  <Data>
    <Category>B</Category>
    <Quantity>8</Quantity>
    <Price>6.99</Price>
  </Data>
</Root>
```

XML
오픈 API 형태로 제공



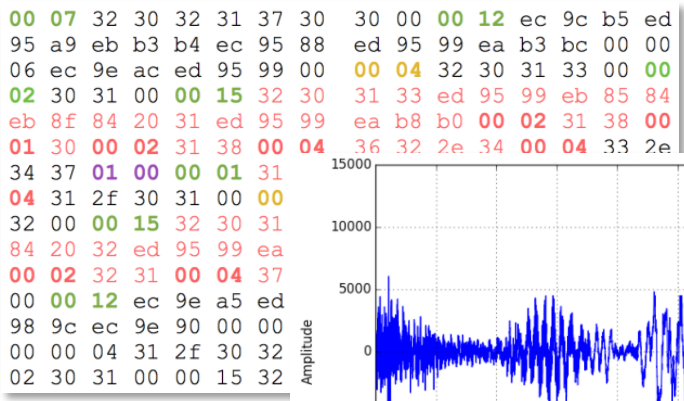
```
{
    "name": "COMPUTER",
    "language": "Kor",
    "words": {
        "ram": "램",
        "process": "프로세스",
        "CPU": "씨피유",
        "Graphic_Card ": "그래픽카드"
    },
    "number": "4"
}
```

JSON
오픈 API 형태로 제공

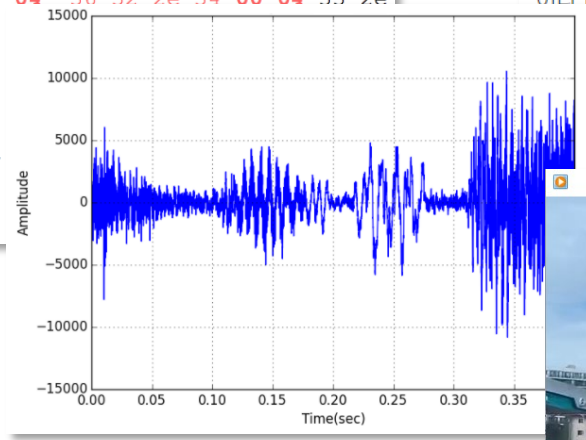
Data Characteristics

■ Unstructured Data

- □ 고정된 필드에 저장되어 있지 않은 데이터
- □ HEX (16진수), 이미지, 비디오 스트림, 오디오 데이터, 자연어
- □ 해당 파일을 파싱하여 메타구조를 갖는 정형데이터 형태의 테이블 구조로 재생성
- □ 수집 난이도가 높음



```
00 07 32 30 32 31 37 30  30 00 00 12 ec 9c b5 ed
95 a9 eb b3 b4 ec 95 88  ed 95 99 ea b3 bc 00 00
06 ec 9e ac ed 95 99 00  00 04 32 30 31 33 00 00
02 30 31 00 00 15 32 30  31 33 ed 95 99 eb 85 84
eb 8f 84 20 31 ed 95 99  ea b8 b0 00 02 31 38 00
01 30 00 02 31 38 00 04  36 32 2e 34 00 04 33 2e
34 37 01 00 00 01 31
04 31 2f 30 31 00 00
32 00 00 15 32 30 31
84 20 32 ed 95 99 ea
00 02 32 31 00 04 37
00 00 12 ec 9e a5 ed
98 9c ec 9e 90 00 00
00 00 04 31 2f 30 32
02 30 31 00 00 15 32
```

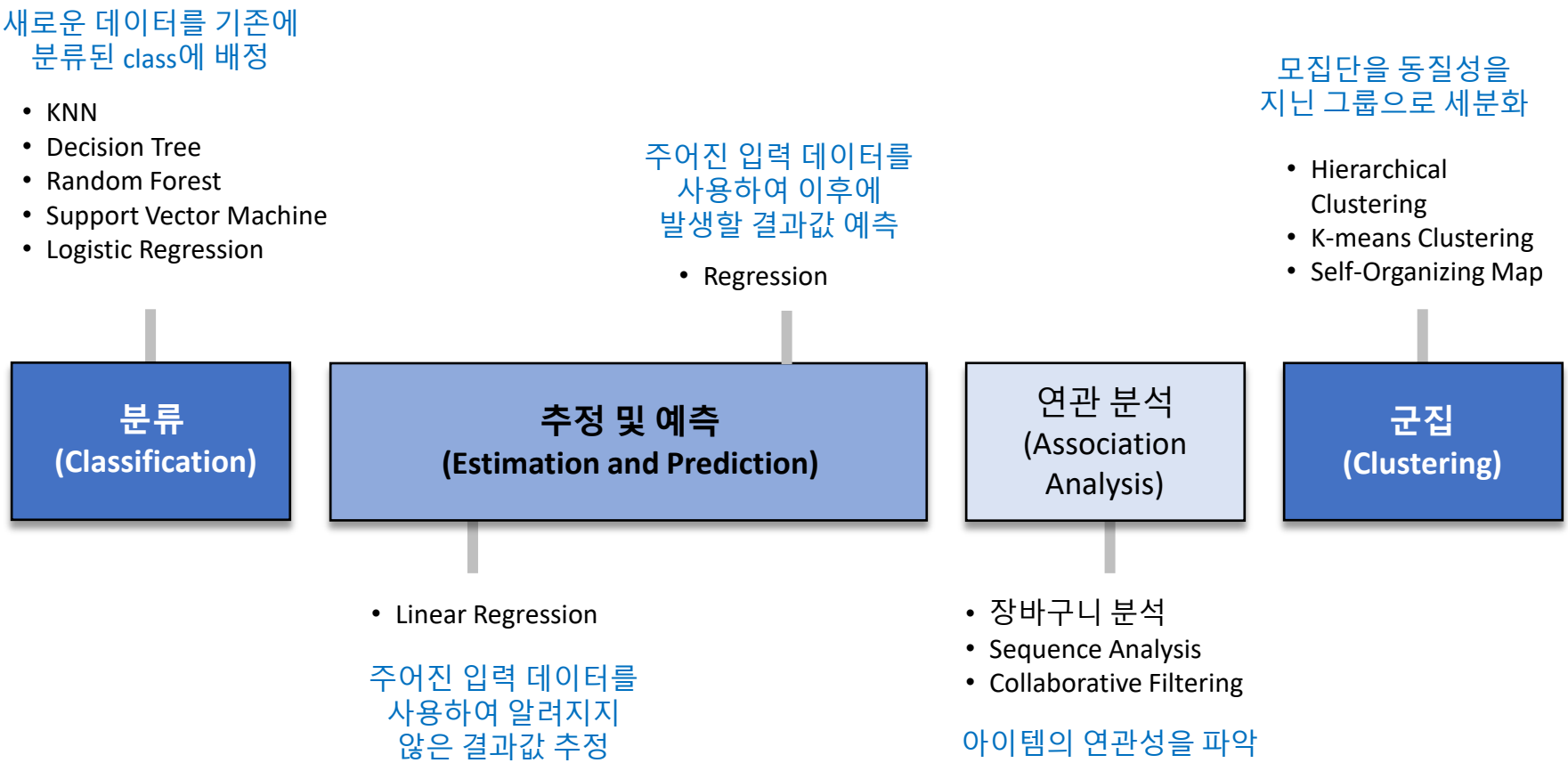


'본격적인 공격'이란 최근 랜섬웨어 공격자들 사이에서 유행하는 것 그대로 '파일 유출'과 '파일 암호화'를 말한다. 현재까지 밝혀진 바 원퍼센트는 AWS S3 스토리지 버킷, 파워셸, 코발트 스트라이크, 미미캐츠(Mimikatz), 샤프스플로잇(SharpSploit), 샤프캐츠(SharpKatz)와 같은 도구들을 활용하고 있다. 대부분 정상적 보안 연구를 위해 만들어진 도구들이지만 해커들이 더 잘 사용하게 된, 그래서 멀웨어라고 명확히 구분하기 힘든 것들이다.

트 피해자들에게 공격 사실을 알리는 메모를 시스템에 남기는데, 이 메모에는 "데이터가 암호화 했고 공격자의 손에 넘어가기도 했다"는 내용과, 어떻게 범인들에게 연락해야 하는지 알려주는 포함되어 있다. 공격자들은 토르에 채널을 운영하고 있으며, 피해자가 안내대로 공격자의 사이 런 연락을 하지 않으면 피해자에게 전 한다. 만약 전화를 걸었는데도 피해 "는 내용의 이메일을 보낸다.

Data Characteristics

■ Algorithms



Data Characteristics

■ Algorithms

Opinion mining

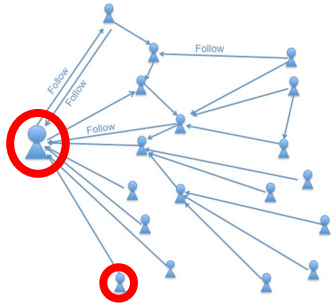
- **문맥과 연계된 감성 분석 (Sentiment Analysis) 활용하여 특정 텍스트의 어조와 감정을 파악**
- 사람들이 특정 제품/서비스를 좋아하거나 싫어하는 이유를 분석하여 대중의 여론을 확인
- 조건과 상황에 맞게 제품/서비스를 추천
- 자연어 처리 과정에서 감성 사전 구축

Text mining

- 텍스트 형태로 이루어진 비정형 데이터를 자연어 처리 방식을 통해 정보 추출하는 기법
- 특정 **키워드**나 문맥을 기반으로 **의미 추출**

Social network

- 소셜 네트워크 서비스에 내포된 **사용자간의 관계를 분석하는 기법**
- 소셜 네트워크는 **노드 (node), 개체와 엣지 (edge)**로 이루어진 자료 구조를 말함
- 노드 중요도 측정
  - Degree centrality, Betweenness centrality, Closeness centrality, Eigenvector centrality



Web mining

- 인터넷을 이용하는 과정에서 생성되는 웹로그 정보나 검색어로부터 추출되는 정보를 대상
- 대체적으로 반정형 or 비정형 데이터
- 페이지 링크 구조를 형성함 -> 하이퍼텍스트 형식의 비순차적

```
<div id="fakebox-container" hidden>
  <div id="fakebox">
    <div class="search-icon"></div>
    <div id="fakebox-text"></div>
    <input id="fakebox-input" autocomplete="off" tabindex="-1" type="url"
        aria-hidden="true">
    <div id="fakebox-cursor"></div>
    <button id="fakebox-microphone" class="microphone-icon" hidden></button>
  </div>
</div>
```

Thank you

KOREA
UNIVERSITY