



# Introduction to Artificial Intelligence

SUMMER 2023

Jiwoon Lee

# PROCESS

Session 0

## I ntroduce

- STUDY GOALS
- STUDY KEYPOINTS

## P lan & Details

- STUDY DETAILS
- WEEKLY TASKS
- STUDY PLANS

## Q &A

- Q&A & NETWORKING TIME



# Jiwoon Lee 이 지 운

Kwangwoon University  
Department of Computer Engineering

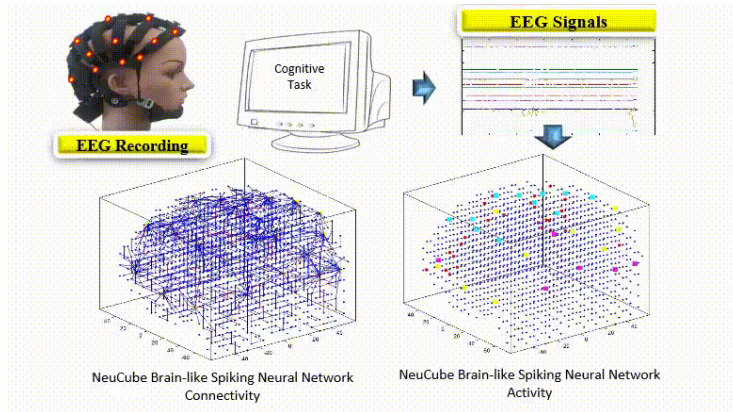
---

## Affiliation

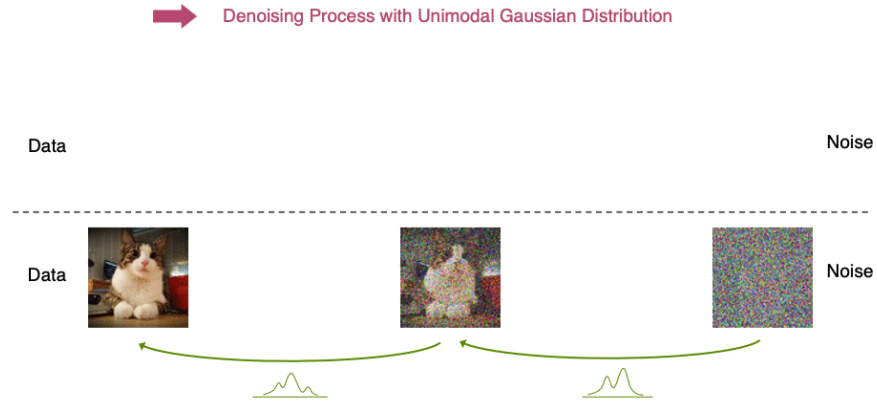
- Student researcher at Bio Computing and Machine Learning Lab
- (Former) Research intern at Qualcomm Institute – University of California, San Diego
- IEEE Seoul Section, CT Society Student Member
- Organizer of Machine Learning Community 'DeepUser'
- Organizer of Kwangwoon Univ. Machine Learning Community 'MI:RU'
- Organizer of Facebook Group '코딩이랑 무관합니다만,'



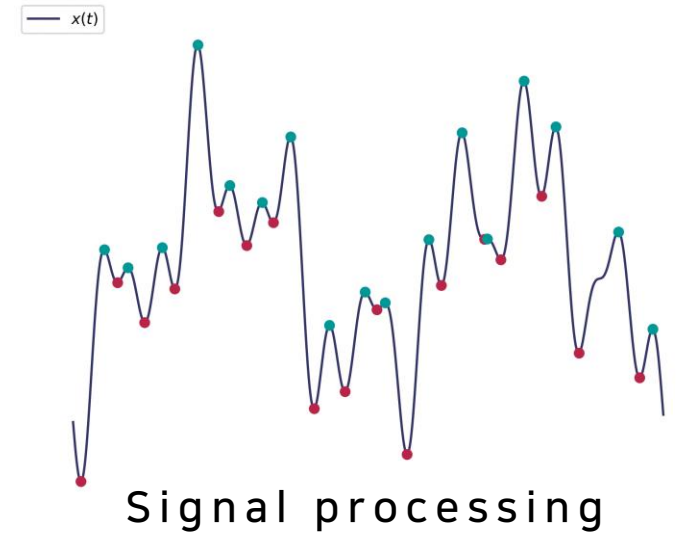
# Research area



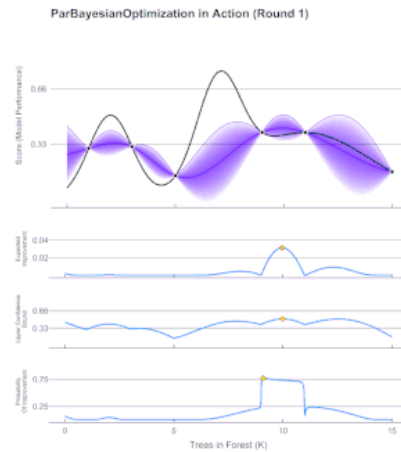
Computational neuroscience



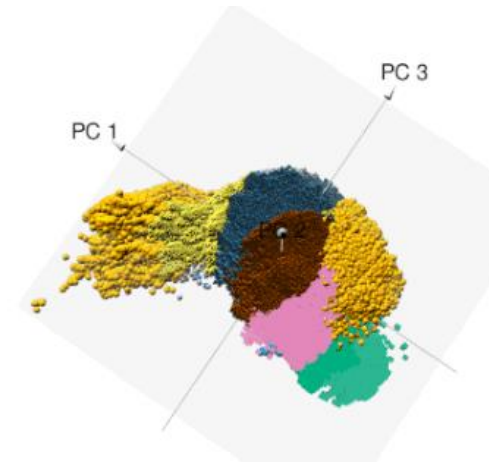
Generative AI



Signal processing



Information theory



Machine Learning

# INTRODUCE

## STUDY GOALS



INTERMEDIATE

Introduction to Artificial Intelligence (Summer 2023)

## Journey through the fascinating world of AI

Learn about the basic algorithms of Machine Learning and Artificial Intelligence

Explore the core concepts, techniques, and applications of AI and ML  
Preparing you for a deep dive into advanced topics

Overall Level

Basic

Intermediate

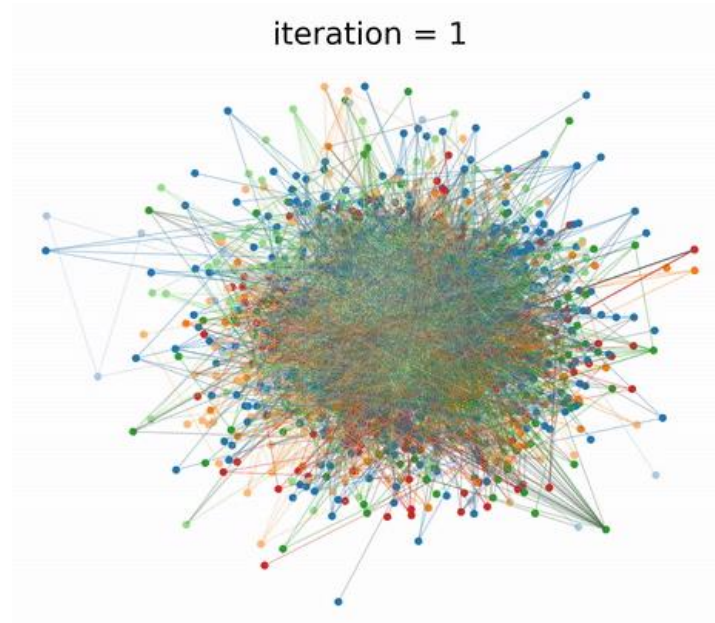
Hard

Average



# INTRODUCE

LECTURE KEYPOINTS



Intuitive understanding

With visualization

# PLAN

WEEKLY TASKS



PERSONAL TASK

Submit Preliminary  
research about topic of the  
next presentation

**'Warning'**

If task is not prepared

**3 warnings == expulsion**

## Section PPT

Depth  
**A**

---

Definition of SVM

Support Vector Machine

Input - set of (input, output) training pair samples

Output - set of weights  $w$

In SVM use the optimization of maximizing  
the margin ("street width") to reduce the number of weights

Neural Nets

< Intro >

Depth: 1

### Schematic approach for SVM

: Linear SVM classifier [Term]

**Margin :**  
Gap of two plus-plane and minus-plane

**Support Vector :**  
Vectors that on plus-plane or minus-plane  
Observations affecting the determination of margin

**SVM Classifier :**  
SVM based classifier if its decision boundary is only depend on support vector (ignore other vectors)

---

Depth: 2

### Schematic approach for SVM

: Linear SVM classifier [Hard Margin-Soft Margin]

**Hard margin :**  
All samples are correctly determined by the decision boundary.  
All data class have to be classified linearly if not it does not work.  
Have a high level of anomaly input.

**Soft Margin :**  
Margin model for greater flexibility by ignoring C anomalies.  
Model flexibility is improved but margin error is also increase

< Graphical Inst >

## Background knowledge for mathematical access of svm

: Linear SVM classifier [Lagrange multiplier method]

$f(x)$ : linear  
 $g(x)$ : constraint  $\rightarrow$  at least one coordinate of  $g(x)$  meets  $f(x)$

Then— which value is minimum(maximum) ?

"Finding maximum or minimum value  
in constraint condition"

$f(x)$ 's y-intercept takes it

## Schematic approach for SVM

: Linear SVM classifier [Karush-Kuhn-Tucker & Duality]

If  $x^*$  and  $u^*, v^*$  are primal and dual solutions, with zero duality gap, then  $x^*, u^*, v^*$  satisfy the KKT conditions

$x^*, u^*, v^*$ 가 KKT 조건을 만족

←strong duality

$x^*$ 는 primal 문제의 solution  
 $u^*, v^*$ 는 dual 문제의 solution

For a problem with strong duality (e.g., assume Slater's condition: convex problem and there exist a strictly satisfying non-affine inequality constraints).

$x^*$  and  $u^*, v^*$  are primal and dual solutions  
 $\Leftrightarrow x^*$  and  $u^*, v^*$  satisfy the KKT conditions

In the theory of optimization, **duality** is the principle that certain optimization problems can be viewed from two perspectives: the fundamental problem and the dual problem.

원래형(quality)이란 어떤 최적화 문제가 원초문제(the primal problem)의 쌍대문제(the dual problem)의 두 가지 관점에서 볼 수 있다는 원칙

< Mathematical Inst >

## RNN Implement

그림 5-20 RNN 계층의 계산 그래프(컴퓨팅 포워드)

The diagram illustrates the forward pass of an RNN layer. It shows the flow of data from inputs to the final hidden state and its derivative. Inputs include the previous hidden state  $h_{prev}$  and its derivative  $dh_{prev}$ , weight matrices  $W_h$  and  $dW_h$ , input vector  $x$  and its derivative  $dx$ , and bias vector  $b$  and its derivative  $db$ . The graph shows the calculation of the next hidden state  $h_{next}$  and its derivative  $dh_{next}$  through a  $\tanh$  activation function and a linear layer.

```
def backward(self, dh_next):
    W_h, b = self.params
    h_prev, h_next = self.cache
    # dh_next = 1을 기본값으로 초기화
    dh = dh_next + (1 - h_next ** 2)
    dx = no_grad().zero_() # dx를 0으로 초기화
    dh = no_grad().zero_()
    dh_prev = no_grad().zero_()
    dh = no_grad().zero_()
    dx = no_grad().zero_()
    dx = no_grad().zero_()

    self.grad[0][...] = dh
    self.grad[1][...] = dh
    self.grad[2][...] = dh

    return dh, dh_prev
```

```
# 초기화
batch_size = 10
wordvec_size = 100
hidden_size = 100 # RNN의 은닉층의 크기
time_step = 5 # Unrolled RNN의 총 단계
lr = 0.1
num_epoch = 100

# 학습
corpus = word_embeddings.load_data('train')
corpus_size = 1000
corpus = corpus.corpus()
vocab_size = int(corpus_size * 1)

xs = corpus[0:-1] # 입력
ys = corpus[1:] # 출력
data_size = len(xs)
print('데이터 크기: %d, 입력 크기: %d' % (corpus_size, vocab_size))

# 학습 시 사용될 변수
max_iters = data_size // (batch_size * time_step)
time_step = 0
total_loss = 0
loss_list = []
pel_list = []

# 모델 생성
model = SimpleRNN(vocab_size, wordvec_size, hidden_size)
optimizer = SGD(lr)
```

Part of Pen TreeBank Dataset

|     |                            |
|-----|----------------------------|
| VRZ | Verb, 3rd ps, sg, present  |
| WDT | Wh-determiner              |
| WP  | Wh-pronoun                 |
| WPS | Possessive wh-pronoun      |
| WRB | Wh-adverb                  |
| #   | Pound sign                 |
| \$  | Dollar sign                |
| .   | Sentence-final punctuation |
| ,   | Comma                      |
| :   | Colon, semi-colon          |

< Implement >



# EXAMPLE

## Weekly Assignment

Jiwoon Lee

Stanford CS230: Deep Learning

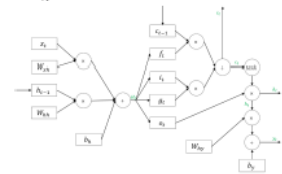
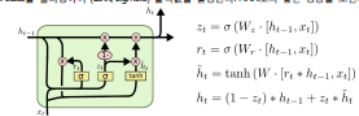
Lecture 5 - AI + Healthcare

Lecture 6 - Deep Learning Project Strategy

KW ML STUDY #3, #4 | 재욱자 주해경 | 2019.08.01

### Topic #4 Long Short-Term Memory (LSTM)

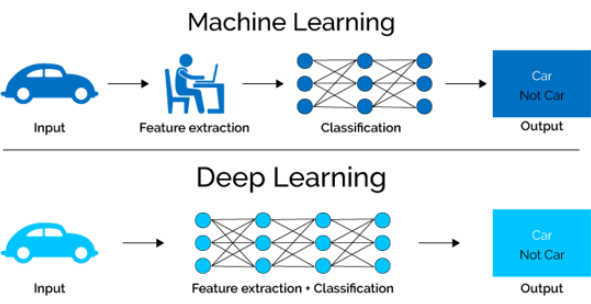
LSTM은 RNN의 긴 의존 기간의 vanishing gradient 문제를 극복하기 위해 설계되었다. 기본적인 RNN은 NN 모듈을 반복하는 재인 형태를 갖는다. LSTM은 각 반복 모듈에 cell-state를 추가한 재인 형태이다. 직전 시점의 그래디언트 값에 영향을 받아 cell-state와 하온 state가 재귀적으로 구성된다. Cell-state는 gate (sigmoid, dot) 를 통해서 정보를 더하거나(가) 제거하는(잊음) 기능을 수행한다. 새 게이트는 새로운 정보를 잊을지 결정하고, 새로운 정보를 기억할 지 결정하여 다음 cell-state를 갱신하고, 이 새 cell-state를 밀려들려 (tanh, sigmoid) 출력값을 결정한다. RNN보다 좋은 성능을 보인다.



Healthcare questions -- levels

1. What happened? (Descriptive)
2. Why did it happen? (Diagnostic)
3. What will happen? (Predictive)
4. What should we do? (Prescriptive)

Paradigm shift of deep learning



BCL Machine Learning Study 2019 Winter 소프트웨어학부 2018203067 이원빈

CS230 Lecture #1 > Class Introduction and Logistics

CS230 Lecture #2 > Deep Learning Intuition

[Lecture Summary]

CS230 Lecture #1 > Class Introduction and Logistics

Machine Learning과 Deep Learning의 발전이 두드러진 이유는 모든 산업에서의 데이터가 아날로그에서 디지털 데이터의 기록으로 바뀌고 현재 그러한 상황이 확립되고 있기 때문이다.

데이터의 양이 많아지면 많아질수록 Logistic Regression, Decision Tree, SVM과 같은 전통적인 기계학습보다는 Neural Network에서 효율을 더 높이 보이는데 그중에서도 Small보단 Medium이, Medium보단 Large Neural Network에서 더 큰 효율을 보인다. 특히 방대한 규모의 데이터를 다루는 현재 상황에서 PGM(Probabilistic Graphic Model), Planning Algorithms, Searching Algorithms, Knowledge Representation, Game Theory 등도 시간에 대해 성능이 점차 상승하고 있지만, Deep Learning과 Machine Learning은 급진적인 성능 향상을 보인다. 규모 면에서도, 계산적 측면에서도 뛰어난 측면을 나타낸다.

CS230 Lecture #2 > Deep Learning Intuition







특정 이미지 데이터는 색깔 정보를 담고 있는 숫자로 구성된 3D 행렬로 구성되어 있는데 이러한 데이터를 학습하기 위해선 학습 모델(Architecture)과 Parameter들을 찾아야 한다. 여기서, Parameter란 입력값에서 CNN을 거쳐 추출되는 출력값으로 변환해주는 함수에 사용되는 인자인데 이는 수많은 수로 이루어졌기 때문에 이를 찾는 것이 학습에 있어 가장 궁극적인 목표라 할 수 있다. 또한, 실제 데이터와 출력값 간의 오차를 줄이는 것도 중요한 쟁점이다.

해당 데이터가 동물이라 할 때, 데이터마다 다른 동물을 대표할 것이고 이를 위해 각 동물에 해당하는 Neuron을 따로 구성하여 따로 학습 후 분류해야 한다. (Multi-Logistic-Regression) 어떤 Neuron이 어떤 동물을 대표하는지 알기 위해서는 One-hot-Encoding을 통해 추출된 벡터를 통한 Labelling 작업이 매우 중요하다. 이미지 데이터 행렬값을 할

<Preliminary research>

# LECTURES

## List of Topic

-  Lecture #1 - Probability Theorems and Metrics for Data Analysis
-  Lecture #2 - Introduction to Neural Network
-  Lecture #3 - Curse of Dimensionality and Dimensionality Reduction
-  Lecture #4 - Introduction to Time Series Analysis
-  Lecture #5 - Markov Chains and Value-Based Reinforcement Learning
-  Lecture #6 - Policy-Based Reinforcement Learning

Easy



Hard

# LECTURES

## List of Topic

 Lecture #1 - Probability Theorems and Metrics for Data Analysis

**Part 1: Introduction to Metrics for Data Analysis**

**Part 2: Data Scaling**

**Part 3: Probability Theorems**

**Part 4: Bayes Theorem**

**Part 5: Practice: Calculating Loss**

**Part 6: Normalization and the Normal Distribution**

**Part 7: Probability Theorems and Metrics Practices**

# LECTURES

## List of Topic

 Lecture #2 - Introduction to Neural Network

**Part 1: Perceptron and Artificial Neurons**

**Part 2: Quick Review of Partial Derivatives**

**Part 3: Activation Functions**

**Part 4: Backpropagation Algorithm**

**Part 5: Convolutional Neural Networks (CNN)**

**Part 6: Recurrent Neural Networks (RNN)**

# LECTURES

## List of Topic

 Lecture #3 - Curse of Dimensionality and Dimensionality Reduction

**Part 1: Introduction to Curse of Dimensionality**

**Part 2: Norm in Linear Algebra**

**Part 3: Projection and Dimensionality Reduction**

**Part 4: Visualization Techniques**

**Part 5: Curse of Dimensionality in Machine Learning**

**Part 6: Dimensionality Reduction Practices**

**Part 7: Visualization Practices**

# LECTURES

## List of Topic

 Lecture #4 - Introduction to Time Series Analysis

**Part 1: Introduction to Time Series Analysis**

**Part 2: Time Series Components and Patterns**

**Part 3: Time Series Modeling: ARIMA**

**Part 4: Time Series Modeling: Seasonal ARIMA**

**Part 5: Time Series Modeling: Exponential Smoothing**

**Part 6: Time Series Modeling: ARIMA vs. Exponential Smoothing**

**Part 7: Time Series Analysis Practices**

# LECTURES

## List of Topic

 Lecture #5 - Markov Chains and Value-Based Reinforcement Learning

**Part 1: Markov Chains**

**Part 2: Introduction to Reinforcement Learning**

**Part 3: Value-Based Reinforcement Learning**

**Part 4: Deep Q-Networks (DQN)**

# LECTURES

## List of Topic

 Lecture #6 - Policy-Based Reinforcement Learning

**Part 1: Introduction to Policy-Based Reinforcement Learning**

**Part 2: Policy Parameterization and Policy Gradients**

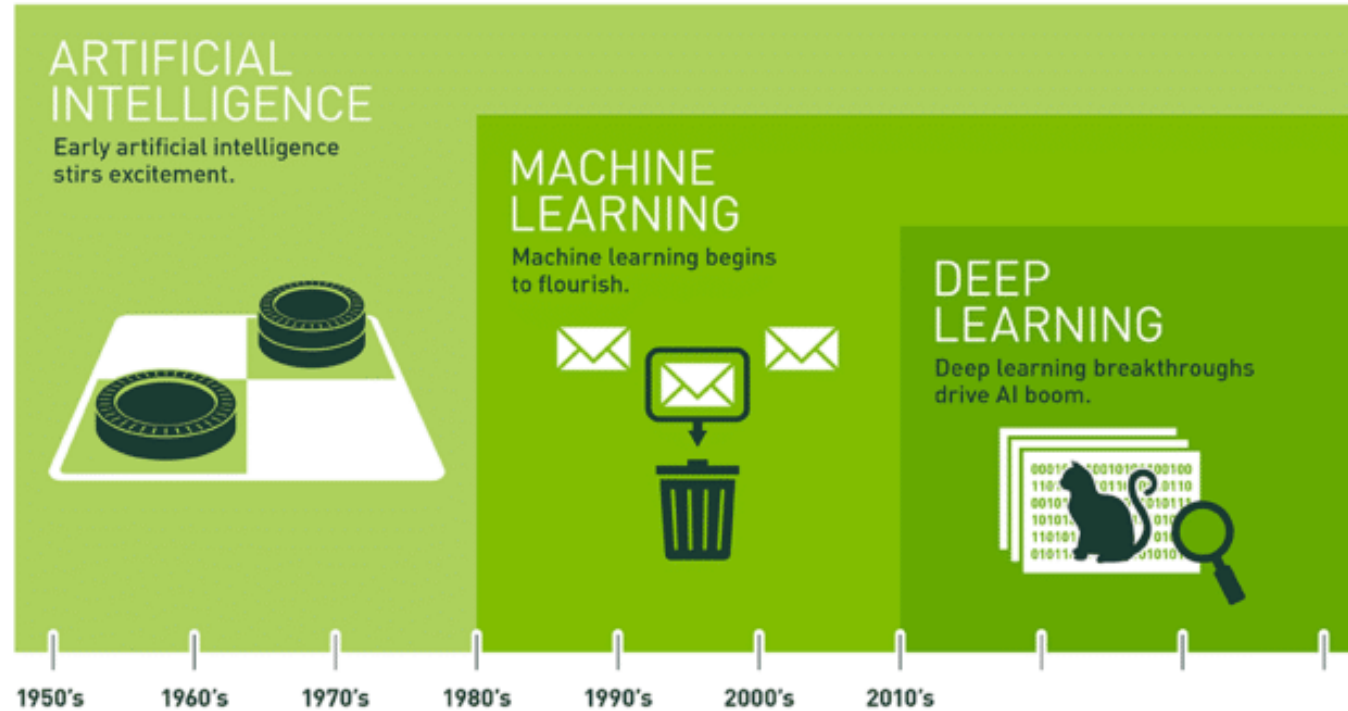
**Part 3: Actor-Critic Methods**

**Part 4: Proximal Policy Optimization (PPO)**



# Part 1: What is AI?

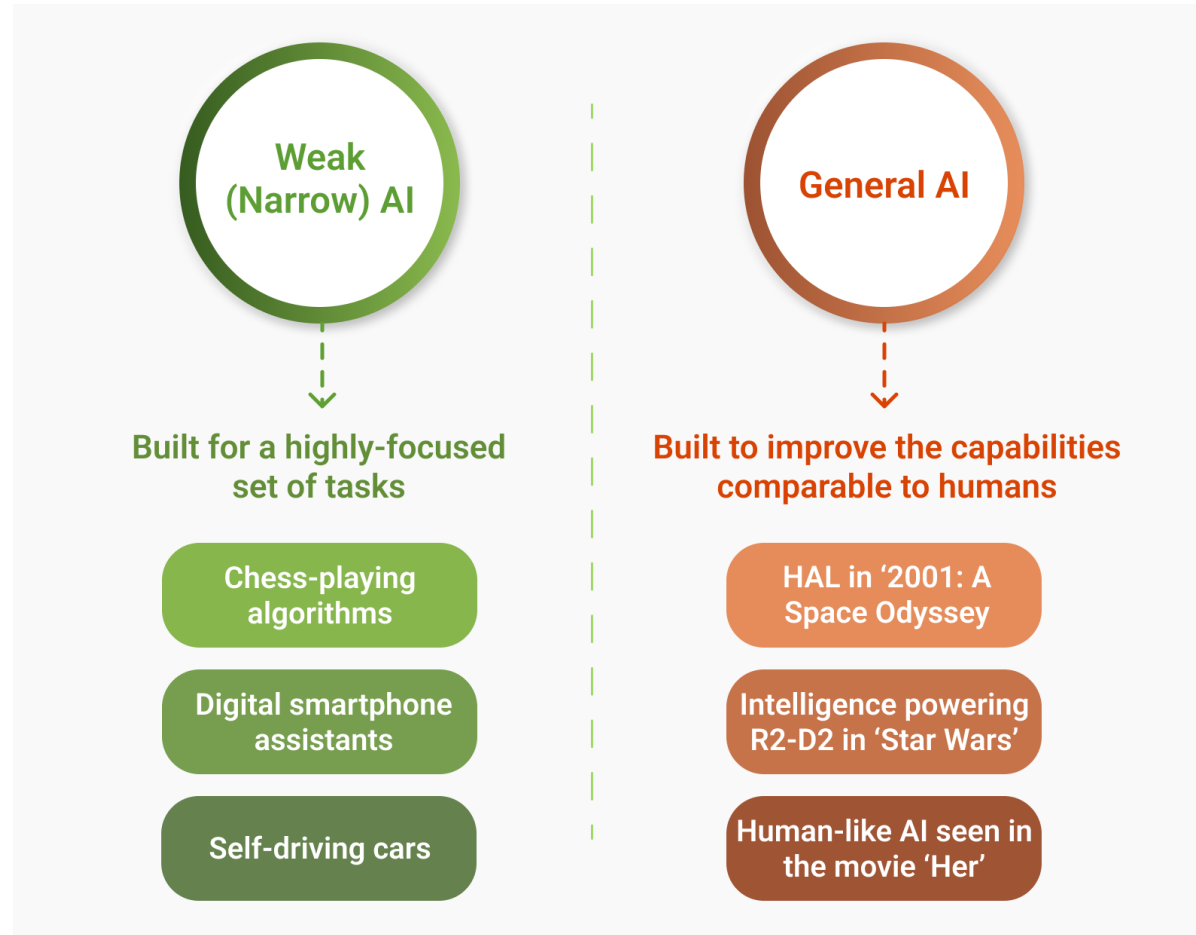
## Narrow AI vs General AI



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

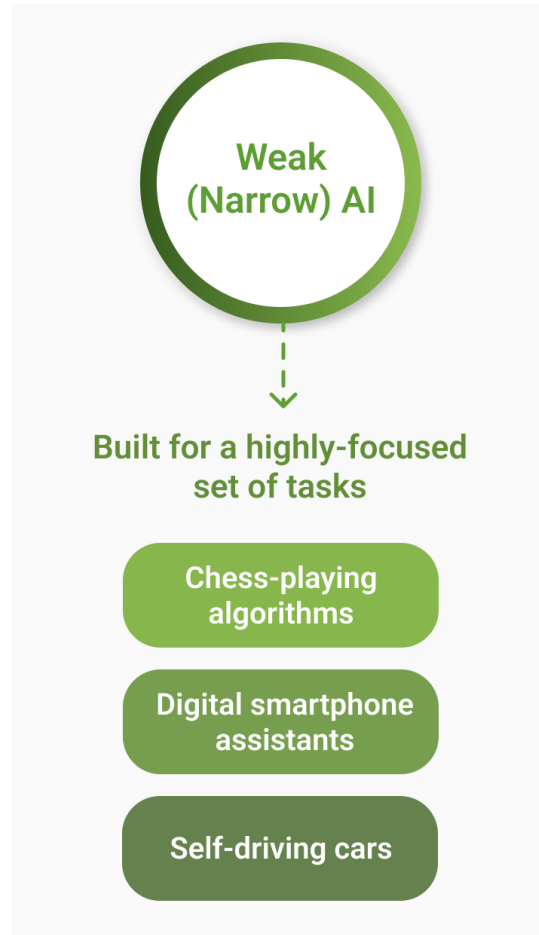
# Part 1: What is AI?

## Narrow AI vs General AI



# Part 1: What is AI?

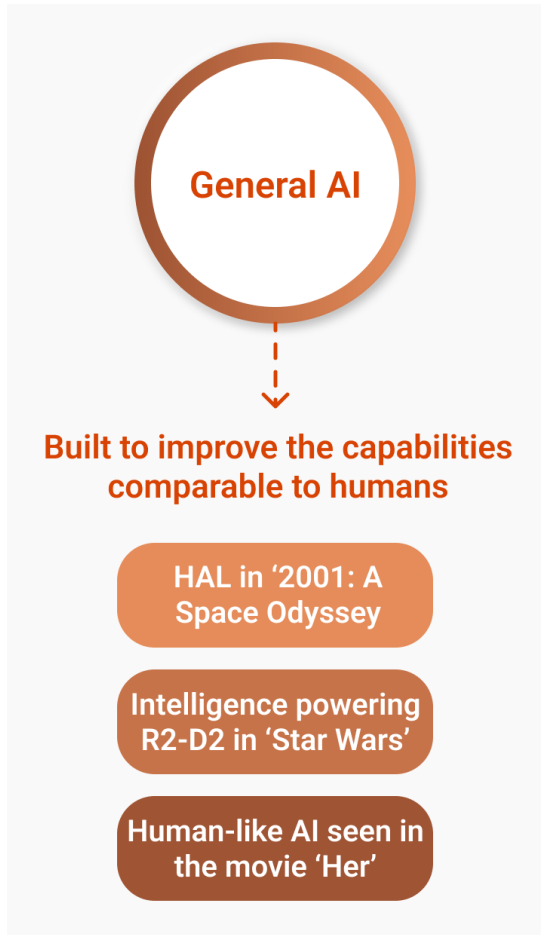
## Narrow AI vs General AI



- Chess playing AI
- Image classification
- Image generation
- ...

# Part 1: What is AI?

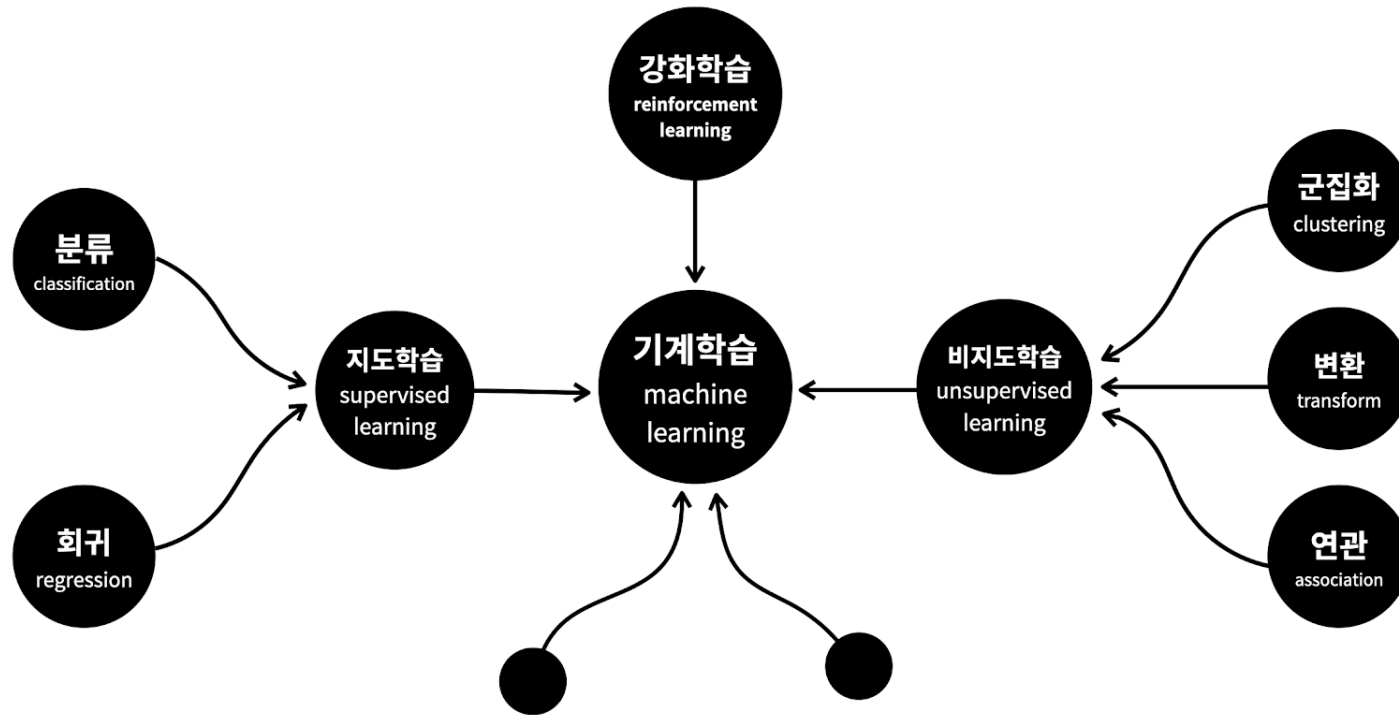
## Narrow AI vs General AI



- Chatbots (ChatGPT, Bard)
- AI robots like in the movie
- ...

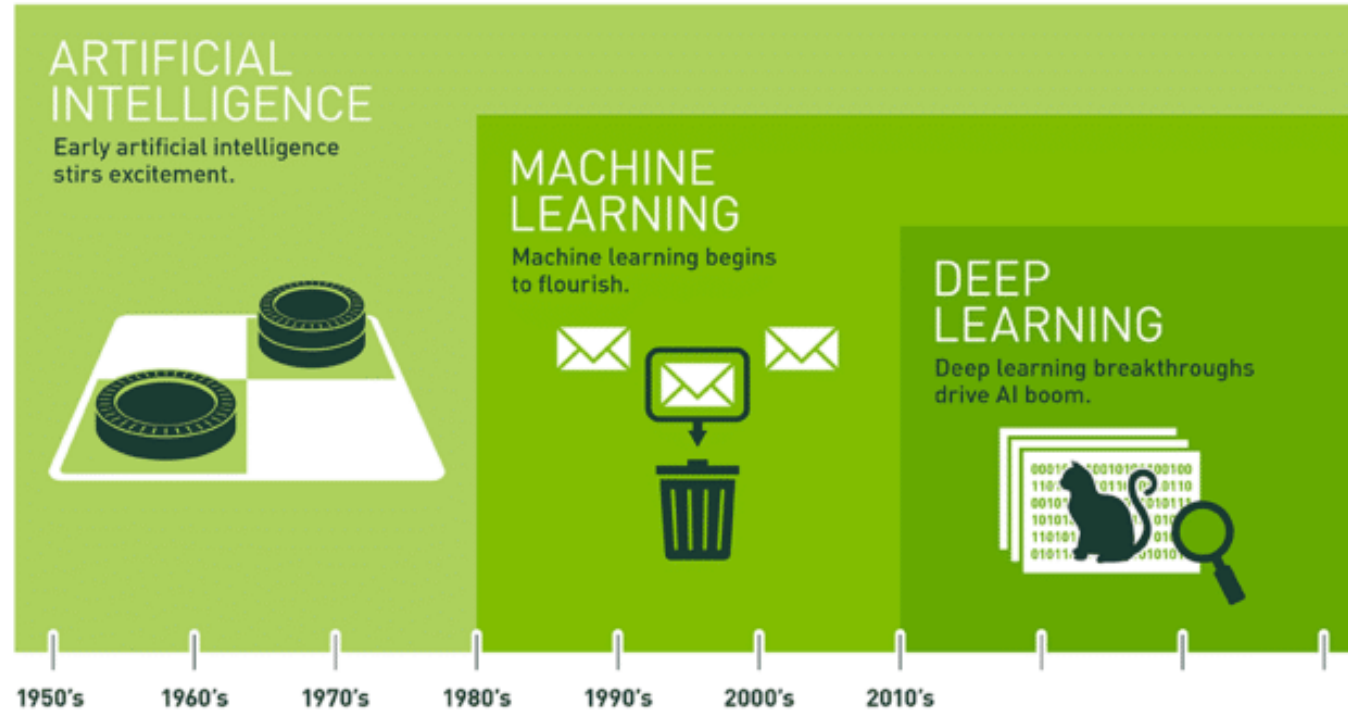
## Part 2: Machine Learning Overview

### Machine Learning



# Part 2: What is AI?

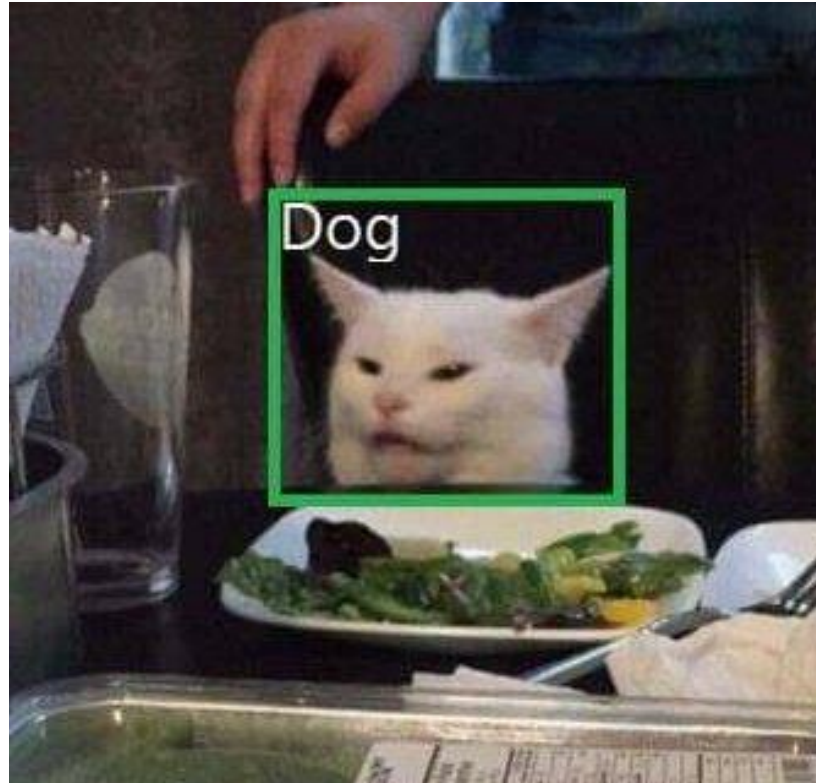
## Machine Learning



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

## Part 3: Classification

Meme



## Part 3: Classification

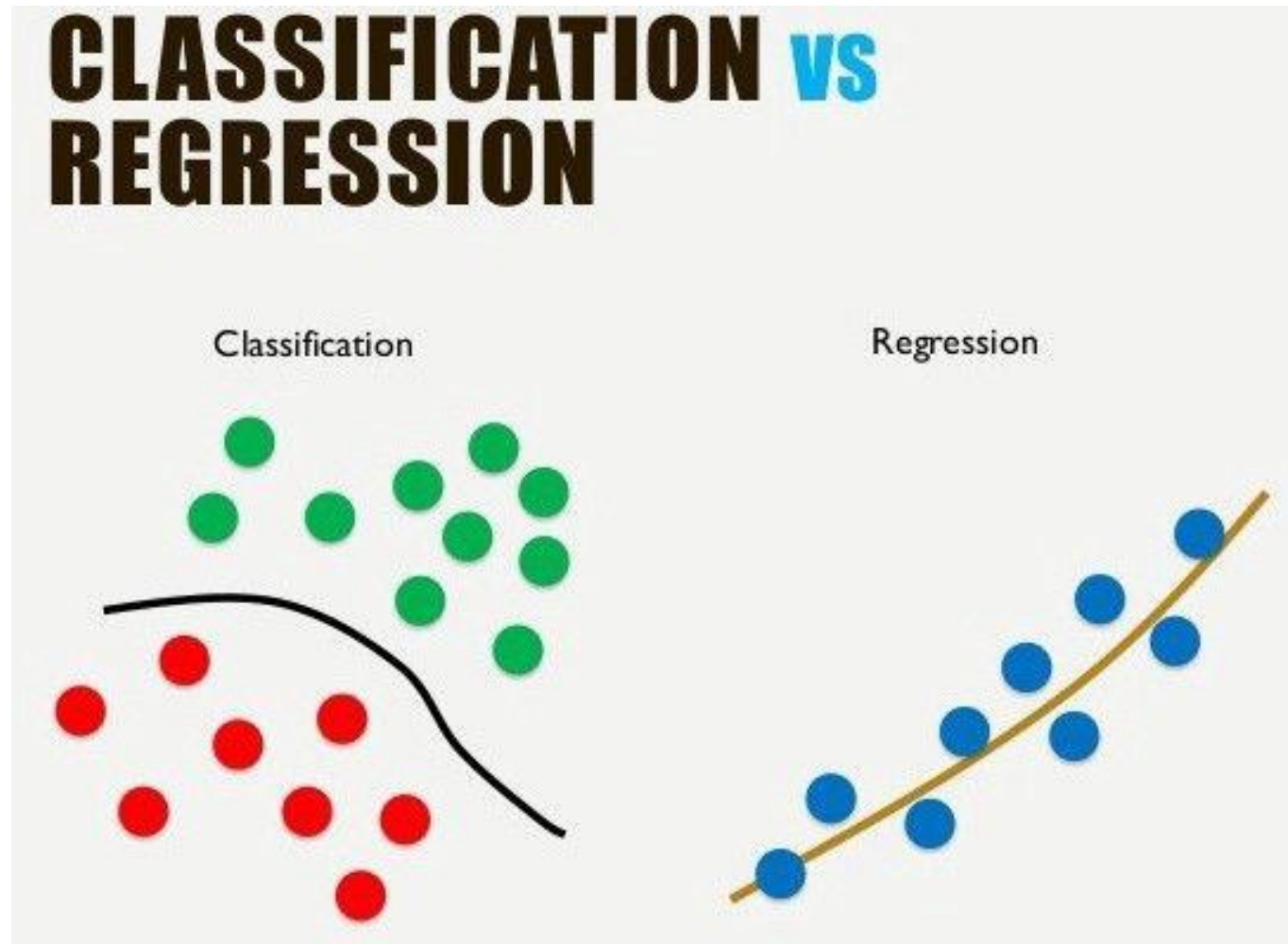
### Examples

| 독립변수<br>(Independent variable) | 종속변수<br>(Dependent variable) | 데이터를 만드는 방식                               |
|--------------------------------|------------------------------|---|
| 공부시간                           | 합격 여부(합격, 불합격)               | 사람들의 공부 시간을 입력받은 후 합격 여부를 확인              |
| 꽃의 잎 길이, 줄기 길이, 넓이             | 꽃의 종                         | 꽃마다 잎의 길이와 넓이, 줄기의 길이 등의 데이터와 꽃의 종을 같이 기록 |
| 키, 몸무게, 시력, 지병                 | 현역, 공익, 면제                   | 키, 몸무게, 시력, 지병 등을 토대로 현역, 공익, 면제인지 확인     |



## Part 3: Classification

Machine Learning



## Part 3: Classification

### Notation in the industries

숫자

양적(Quantitative)

얼마나 큰지, 얼마나 많은지, 어느 정도인지를 의미

아름

범주(Categorical)

같은 특성을 지닌 부류나 범위를 의미

## Part 4: Regression

### Machine Learning



| 날짜       | 요일 | 온도 | 판매량 |
|----------|----|----|-----|
| 2020.1.3 | 금  | 20 | 40  |
| 2020.1.4 | 토  | 21 | 42  |
| 2020.1.5 | 일  | 22 | 44  |
| 2020.1.6 | 월  | 23 | 46  |
| 2020.1.7 | 화  | 24 | 48  |
| 2020.1.8 | 수  | 25 | ?   |

원인 → 결과

과거의 데이터

미지의 데이터

# Part 4: Regression

## Machine Learning

| 독립변수 | 종속변수 |
|------|------|
| 20   | 40   |
| 21   | 42   |
| 22   | 44   |
| 23   | 46   |



## Part 4: Regression

Machine Learning



# Part 4: Regression

## Machine Learning



| 날짜       | 요일 | 온도 | 판매량 |
|----------|----|----|-----|
| 2020.1.3 | 금  | 20 | 40  |
| 2020.1.4 | 토  | 21 | 42  |
| 2020.1.5 | 일  | 22 | 44  |
| 2020.1.6 | 월  | 23 | 46  |
| 2020.1.7 | 화  | 24 | 48  |
| 2020.1.8 | 수  | 25 | ?   |

원인 → 결과

과거의 데이터

미지의 데이터

## Part 4: Regression

### Examples

| 독립변수<br>(Independent variable) | 종속변수<br>(Dependent variable) | How to make data?                    |
|--------------------------------|------------------------------|--------------------------------------|
| 공부시간                           | 시험점수                         | 사람들의 공부 시간을 입력받은 후 점수를 확인            |
| 온도                             | 레모네이드 판매량                    | 온도, 그날의 판매량 기록                       |
| 집의 위치, 역세권 여부 등                | 집의 가격                        | 집부터 역까지의 거리, 주변 편의시설 개수 등<br>의 값을 기록 |

# Quiz

| 면적(평) | 온도(섭씨) | 판매량(개) |
|-------|--------|--------|
| 1000  | 10     | 100    |
| 200   | 26     | 200    |
| 300   | 31     | 300    |

| 계절 | 날씨 | 휴가지 |
|----|----|-----|
| 봄  | 비  | 바다  |
| 여름 | 흐림 | 산   |
| 가을 | 맑음 | 강   |



## Part 5: Data preprocessing

### Data Preparation



## Part 5: Data preprocessing

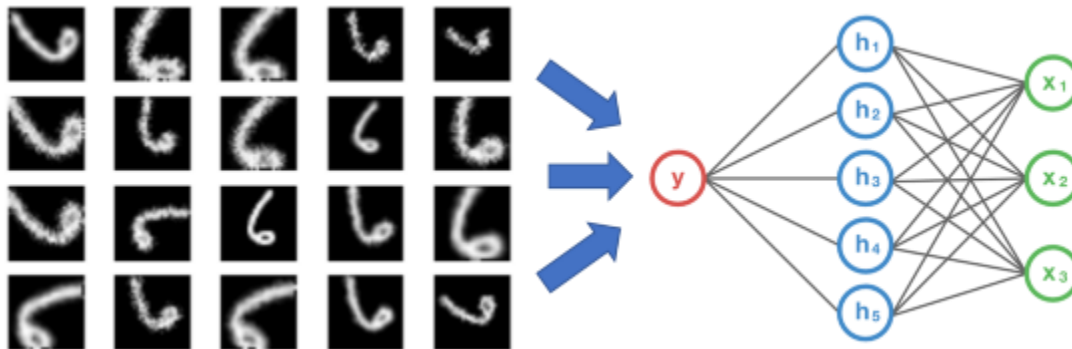
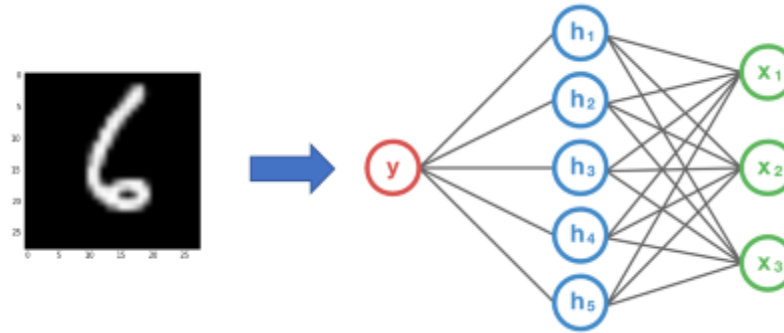
Why preprocessing is important?



## Part 5: Data preprocessing

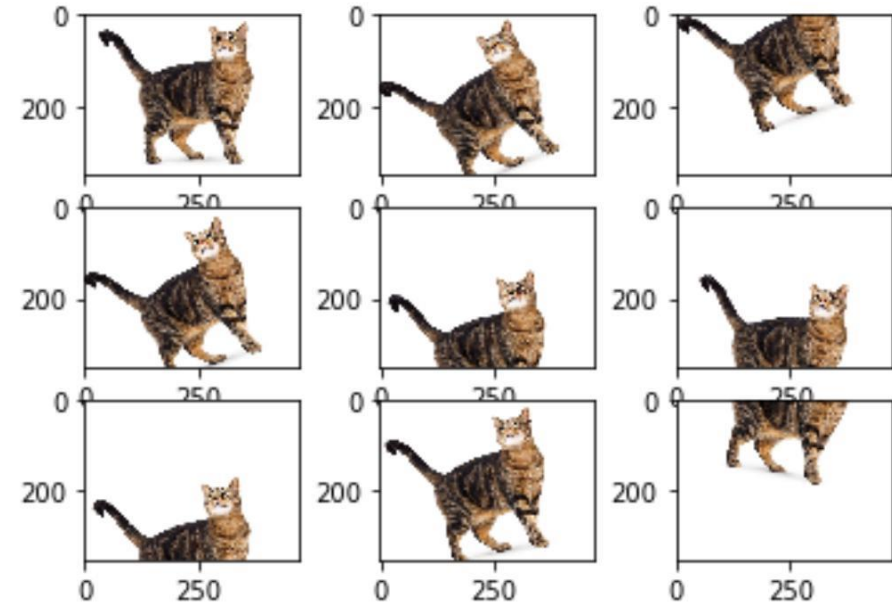
Why preprocessing is important?

Training



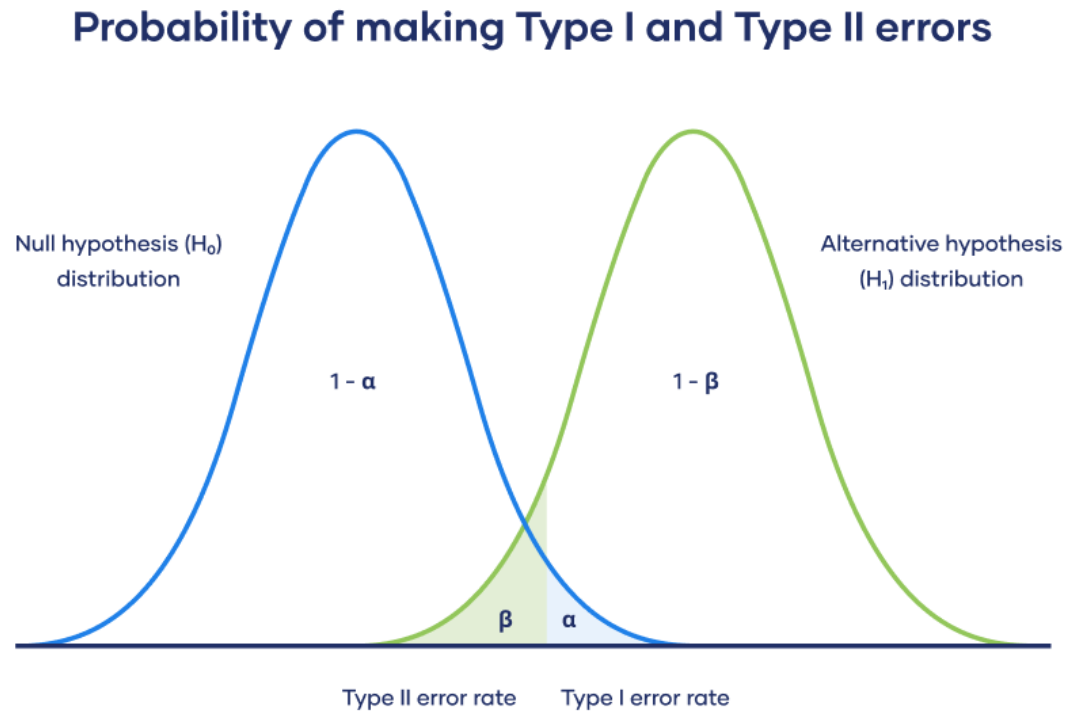
## Part 5: Data preprocessing

Why preprocessing is important?



# Part 5: Data preprocessing

Why preprocessing is important?



Better data  
Better classification model



Better Result!  
(Lower error)

# Trends



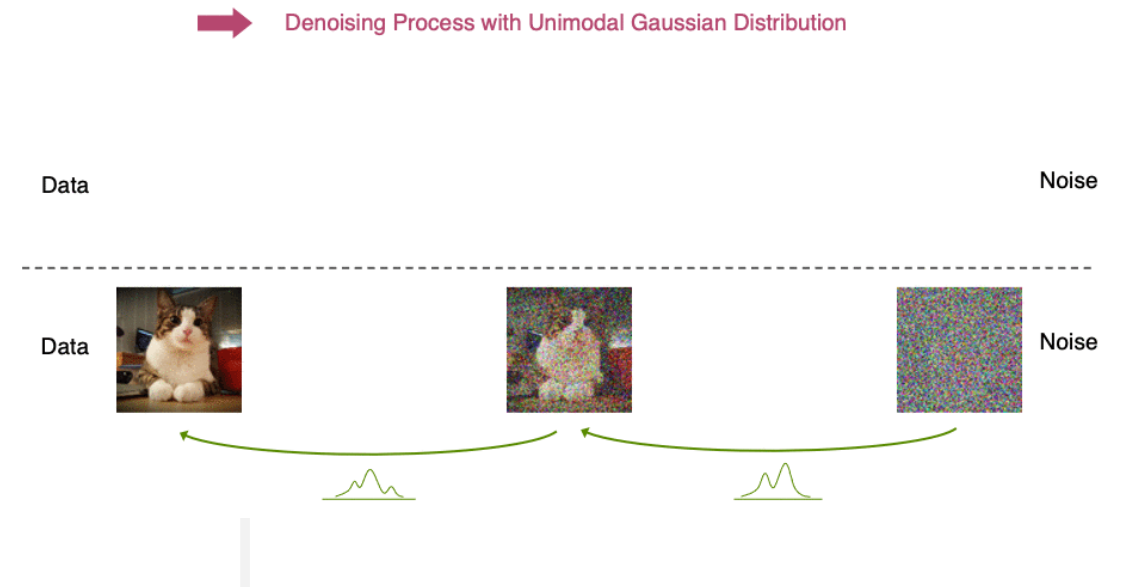
PyTorch

TensorFlow

# Trends



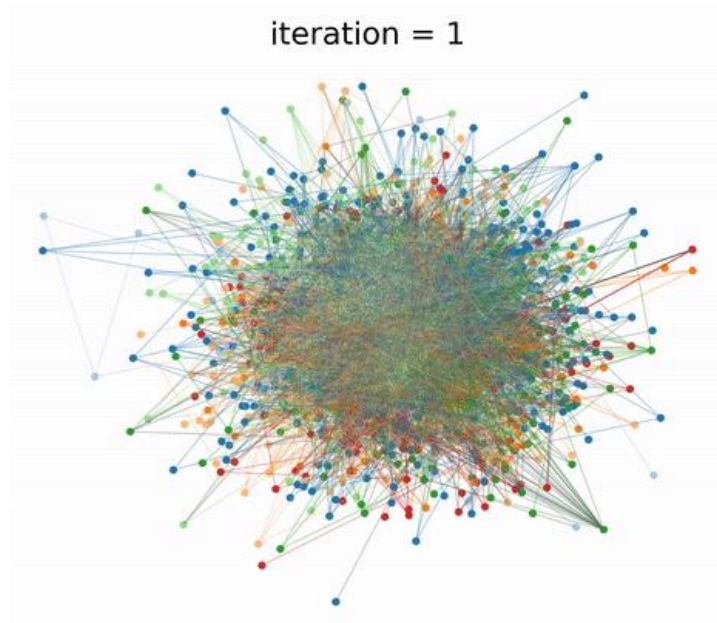
Large Language Model



Generative AI

# Reminder

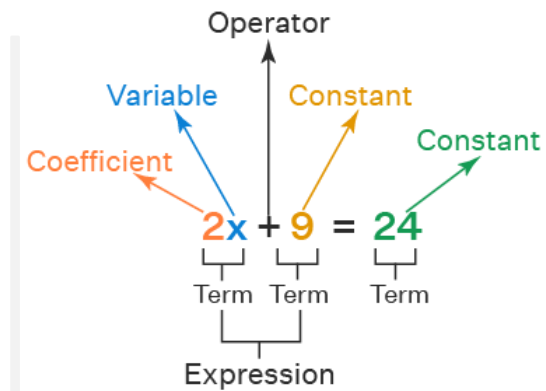
LECTURE KEYPOINTS



Intuitive understanding

With visualization

Parts of an Equation

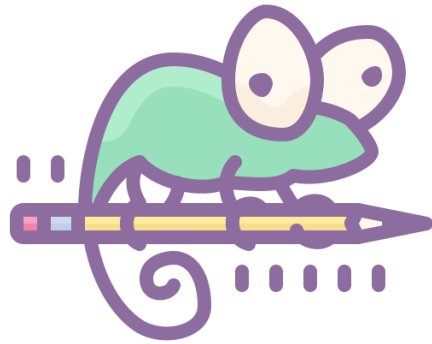


Left hand side = Right hand side

Understanding of equation



# Q&A



---

**Jiwoon Lee**  
**이지운**

Email. [metr0jw@outlook.com](mailto:metr0jw@outlook.com)

Phone. +82-10-3326-2914

Kwangwoon University

School of Computer Information and Engineering

Member of BCML Lab.