Machine Learning

Understanding AI & Machine Learning

Dept. SW and Communication Engineering

Prof. Giseop Noh (kafa46@hongik.ac.kr)

Professor Info

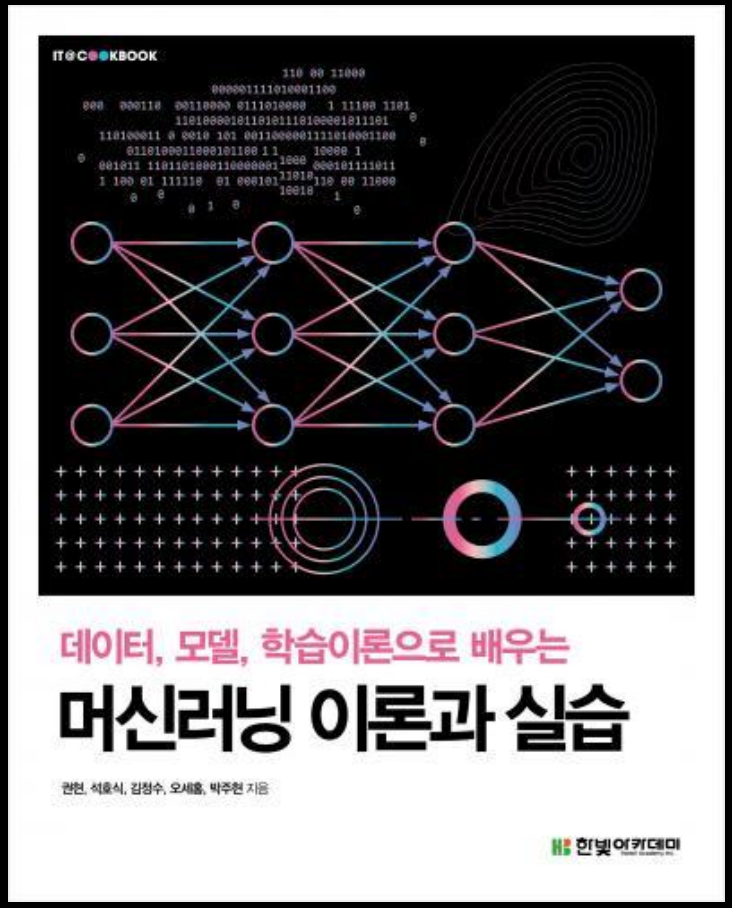
- **Name: Giseop Noh**
  - Email: kafa46@hoingik.ac.kr (backup: kafa46@gmail.com)
  - Cell: 010-5746-0257, Office: D402

- **Job Career**
  - ~ 2006. 12: R.O.K Airforce, Officer in Logistics
  - ~ 2016. 12: Defense Acquisition Program Administration
  - ~ 2018. 02: R.O.K Airforce Academy, Assistant Professor
  - ~ 2025. 02: Cheongju University, Assistance Professor
  - ~ Now: Hongik University, Assistance Professor

- **Academic Background**
  - R.O.K. Airforce Academy, Bachelor in Industrial Engineering
  - University of Colorado Denver, Master in Computer Science
  - Seoul National University, Ph.D in Computer Engineering

Textbook

Main Textbook



- Author: 권현 et al.
- Pub Date: 2024-07-15
- Total pages: 404
- ISBN :9791156640349
- eISBN :97911566405561

About class

- **The first 10 minutes of class will be used for**
  - Check attendance,
  - PPT setup,
  - and other necessary preparations.
- **If there is a consecutive class after this one,**
  - I will try to finish about 5 minutes early whenever possible.
  - However, the class may still end exactly on time.
  - If necessary, please adjust your individual course schedule accordingly,
    - considering the time required to move between buildings or classrooms.

Attendance Check

- **Attendance Check:**
  - Online attendance verification
  - If necessary, offline attendance may be used.

Grading

- **Grading Distribution**
  - A+ / A0: 20–30%B+ / B0: 30–40%C+ / C0: 20–30%D+ / D0: 10–20%
  - Subject to change according to school policy (advance notice will be given if changes occur).

- **Grade Processing for Insufficient Attendance**
  - Students who do not meet the required attendance days will receive an F grade according to university regulations.

- **How to Receive an F Grade**
  - Academic misconduct (cheating, plagiarism, etc.)
  - Final score below 30 points on average
  - Missing more than 1/3 of total class hours
  - Failure to take the midterm or final exam without a valid reason

Excused Absence

Excused Absence Policy

- Attendance will be recognized only if a request for approval of excused absence is submitted for reasons specified in Academic Regulations Article 75
- and approved by the dean of the affiliated college.
- Academic Regulations:
  - https://www.hongik.ac.kr/kr/introduction/regulations-book.do?mode=list&srCategoryId=110

Excused Absence Application Procedure

- Check Eligibility
- Verify that the reason for absence qualifies under Academic Regulations Article 75.
- Prepare Documents: Complete the "Request for Approval of Excused Absence" form and gather supporting documents (e.g., medical certificate, official notice).
- Submit Application: Submit the completed form and supporting documents to the affiliated college office.
- Dean’s Approval: The dean reviews and approves the request.
- Notification: Approved absences will be officially recognized, and the student will be informed of the outcome.

Changes in Lecture Progress

- **The lecture plan may be adjusted if deemed reasonable based on student surveys or suggestions.**
  - After the midterm exam,
  - adjustments may be made based on lecture evaluation feedback, including:
    - The need for additional explanations or deeper learning
    - Adjustments to progress and assessment scope to align with other course sections

- **If any modifications to the lecture plan are necessary,**

  students will be notified in advance.

Course overview

1	Introduction to Artificial Intelligence and Machine Learning	Chapter 1
2	Basic Mathematics for Machine Learning	Chapter 2
3	Numpy Basics and Practice	Chapter 2
4	Data Visualization Basics	
5	Pandas Basics	
6	Pandas Applications and Regression Preview	
7	Model Performance Evaluation	
8	Mid-term exam	
9	Linear Regression Theory	Chapter 3
10	Linear Regression Practice	Chapter 3
11	Logistic Regression Theory	Chapter 4
12	Logistic Regression Practice	Chapter 4
13	K-Nearest Neighbors (KNN)	Chapter 7
14	Decision Trees	Chapter 8
15	Final exam	

Contents

- **Concept of Artificial Intelligence**
- **Development of Artificial Intelligence**
- **Basics on Machine Learning**

Goals

- **Understand the relationship among artificial intelligence, machine learning, and deep learning.**
- **Understand the explosive growth of artificial intelligence in terms of data, models (algorithms), and computing power.**
- **Understand the basic principles of machine learning, including supervised learning, unsupervised learning, and reinforcement learning.**

Concept of Artificial Intelligence

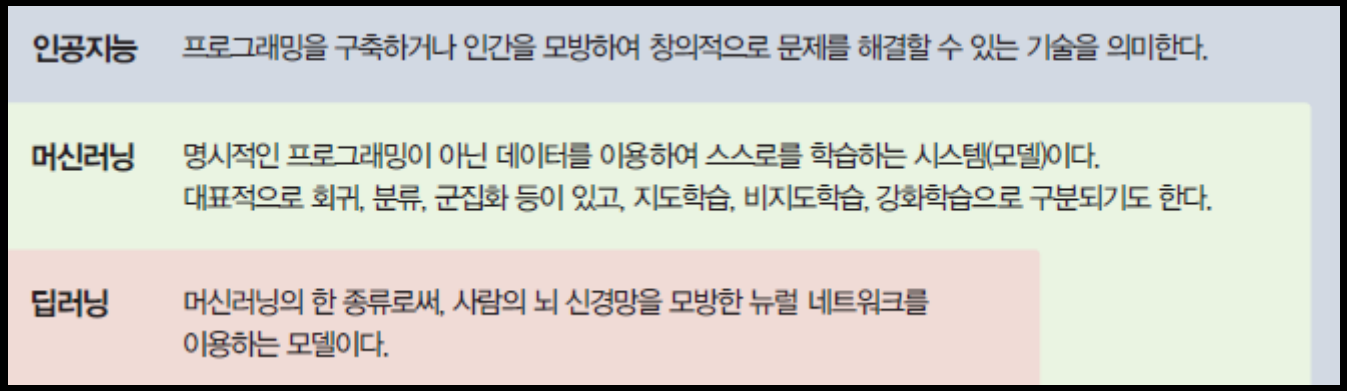
인공지능 개념

■ Artificial Intelligence, AI

- It is a technology inspired by human intelligence and serves as a key technology in the Fourth Industrial Revolution.
- Its goal is to implement systems that perceive, decide, and act like humans.

■ Classification of Artificial Intelligence

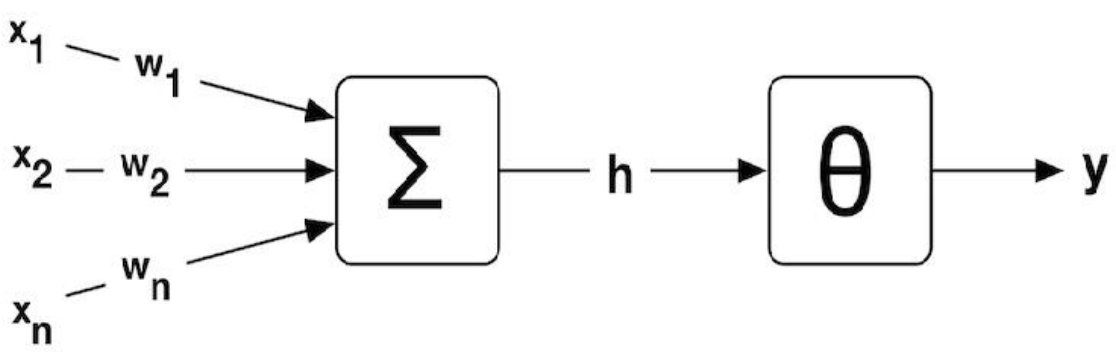
- Weak Artificial Intelligence: AI specialized in a specific domain (e.g., image classification in computer vision)
- Strong Artificial Intelligence: A general-purpose AI that processes diverse data and mimics human thought processes (e.g., analyzing images and providing natural language descriptions)

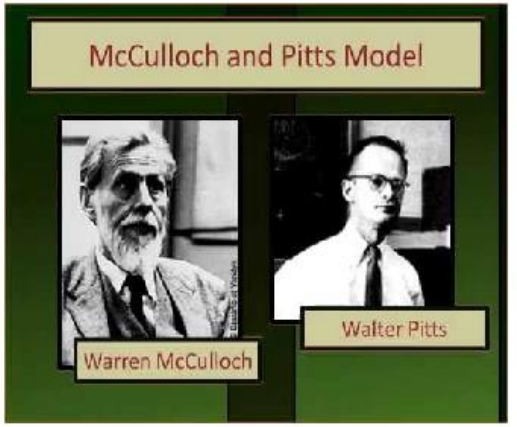


Development of Artificial Intelligence

History of Artificial Intelligence (1/2)

- **1943**: Warren McCulloch and Walter Pitts proposed that the functioning of neurons could be analyzed using propositional logic and that neurons capable of expressing logical operations could be constructed.

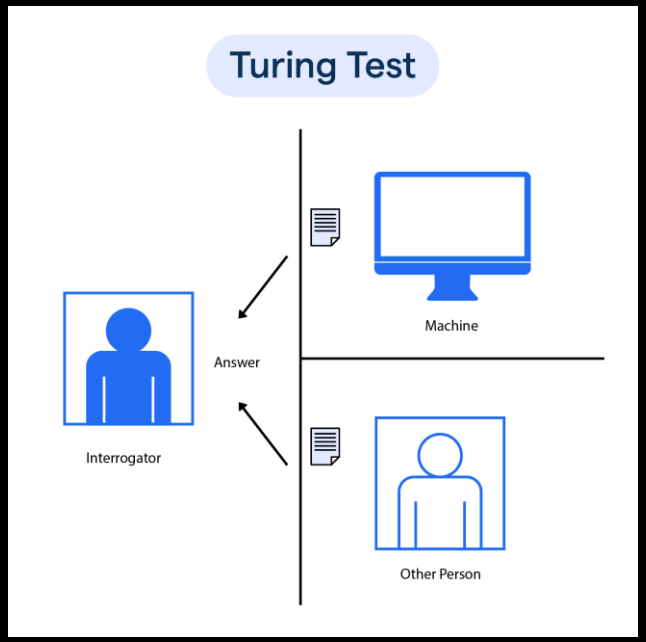


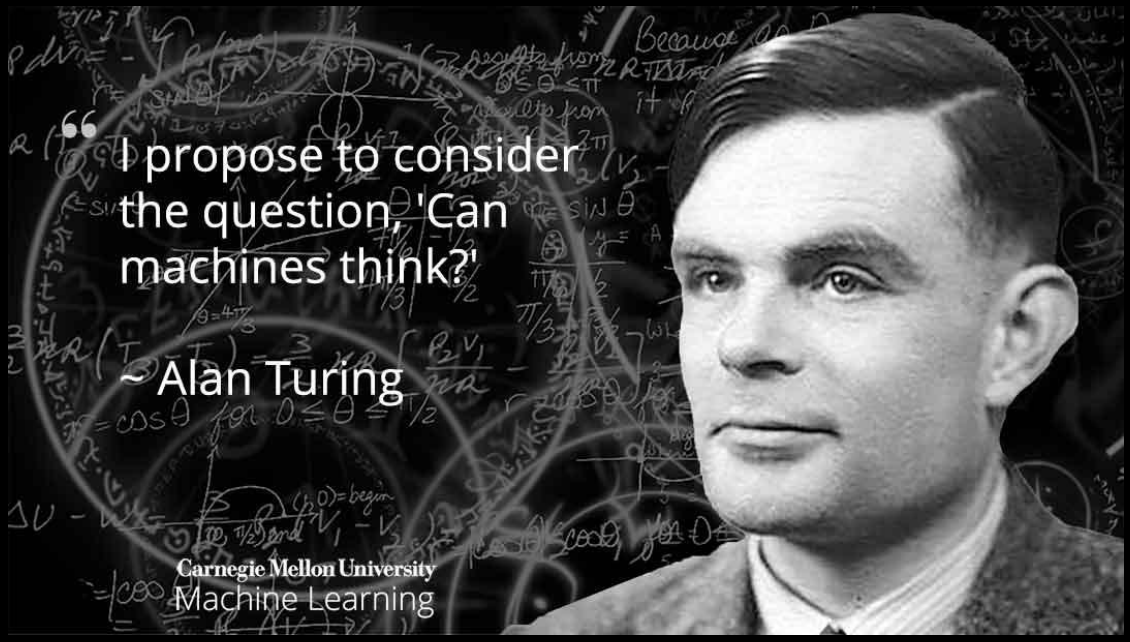


https://www.linkedin.com/pulse/mcculloch-pitts-neuron-bidyut-bikash-

History of Artificial Intelligence (1/2)

- **1950**: Alan Turing proposed the **Turing Test**, suggesting that if a conversational partner cannot be distinguished as either a human or artificial intelligence, the AI is considered to have passed the test.

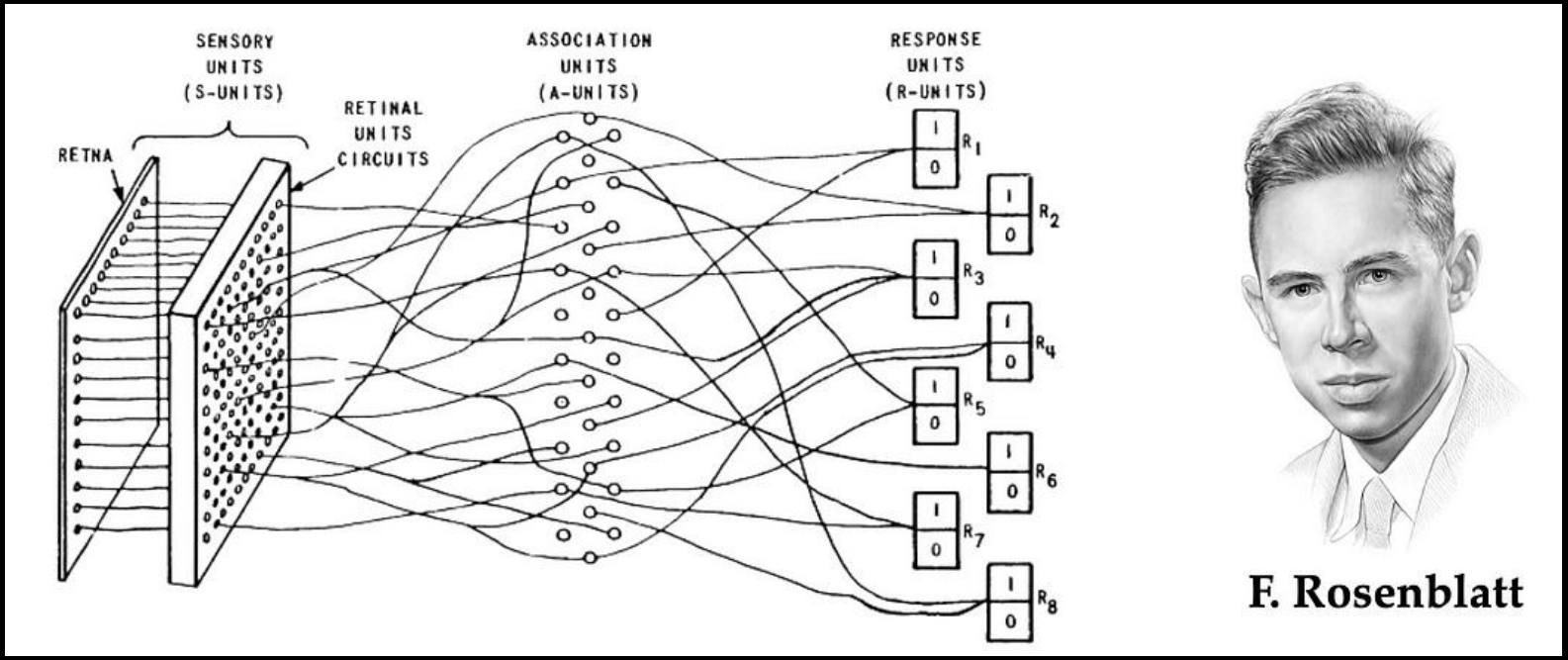




History of Artificial Intelligence (1/2)

- **1957**: Frank Rosenblatt introduced the **Perceptron**

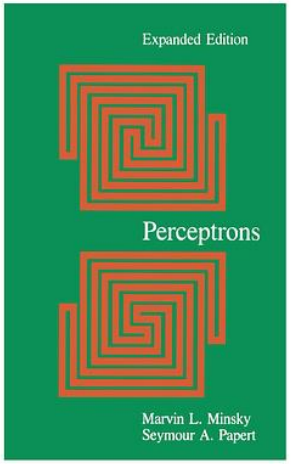
- The first program that mimics the function of neural cells.

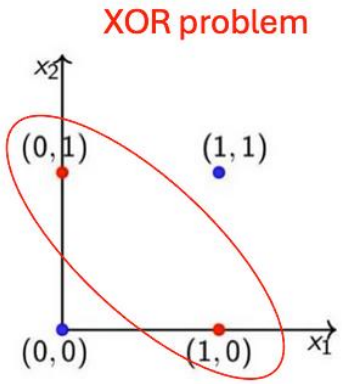


History of Artificial Intelligence (2/2)

- **1969**: Marvin Minsky and Seymour Papert pointed out the limitations of the Perceptron, stating that it could not solve nonlinear problems such as XOR.

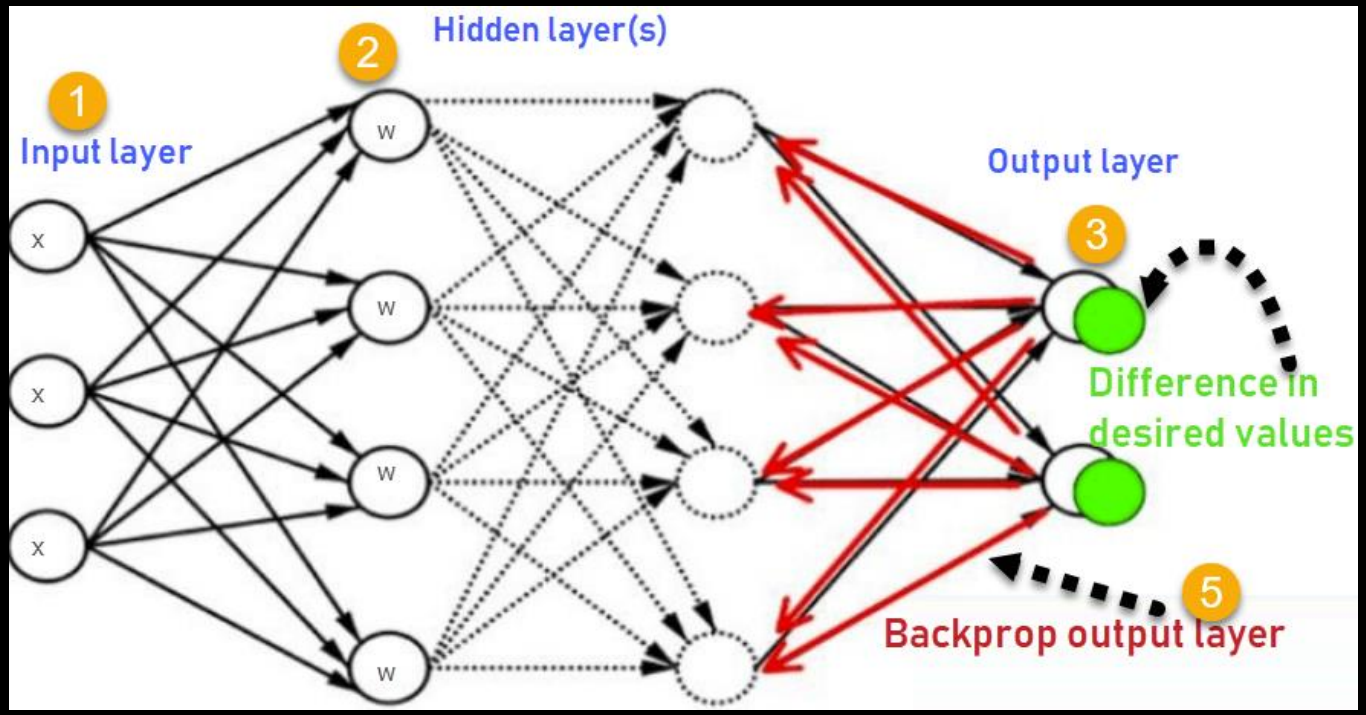
Minsky and **Papert**: Perceptrons: An introduction to computational geometry. MIT Press, 1969.





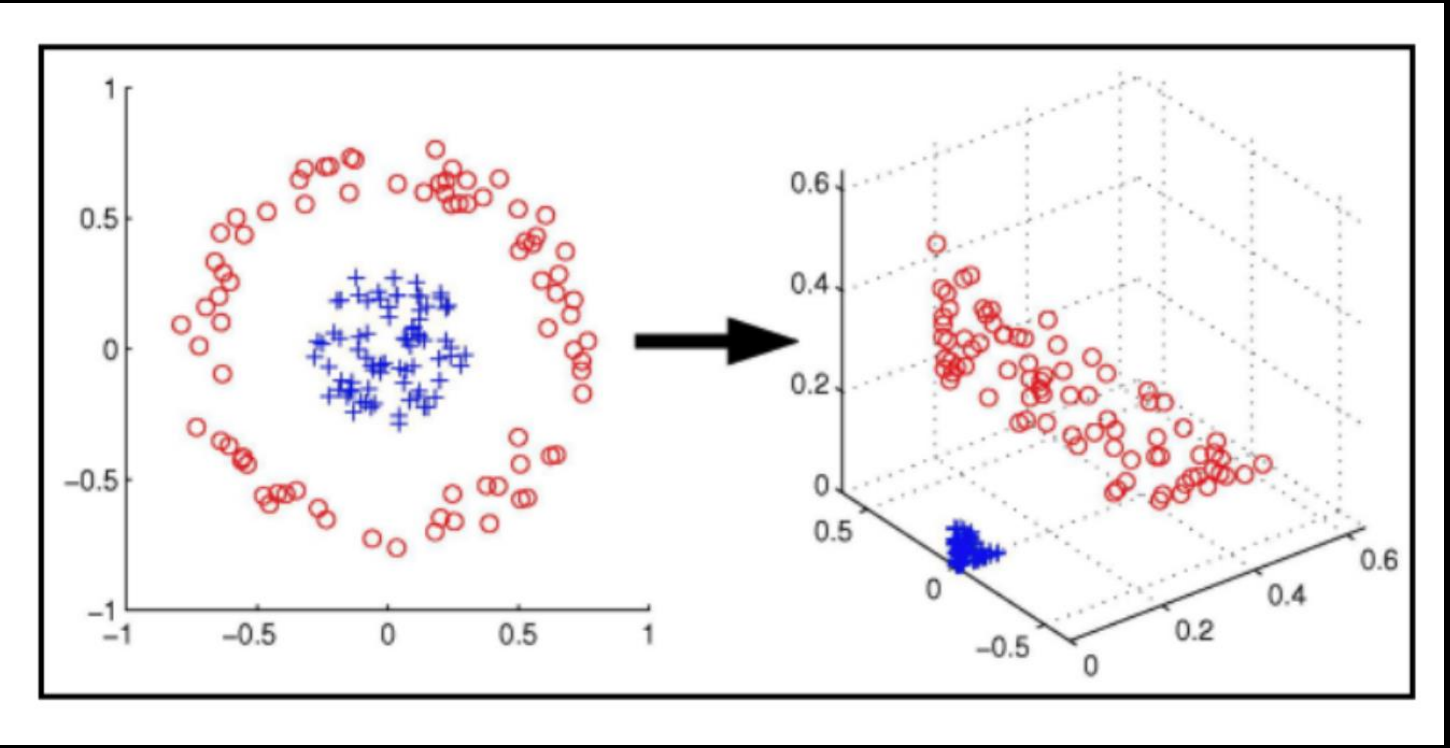
History of Artificial Intelligence (2/2)

- **1986**: David Rumelhart and colleagues proposed a learning method using **error backpropagation** in multi-layer perceptrons.



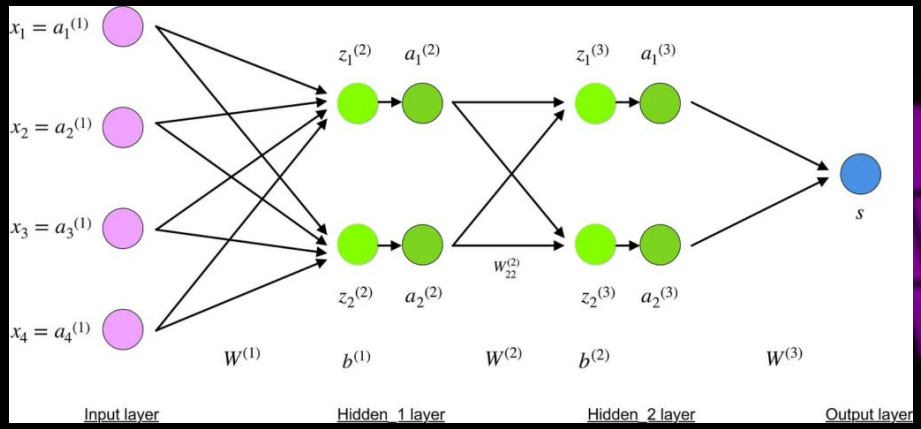
History of Artificial Intelligence (2/2)

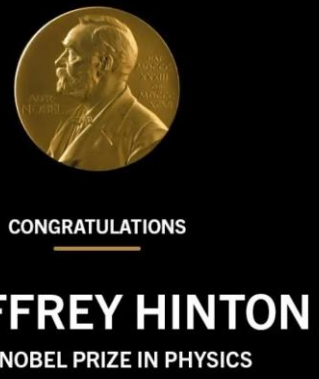
- **1995**: Vladimir Vapnik and Corinna Cortes introduced the **Support Vector Machine (SVM)** model.



History of Artificial Intelligence (2/2)

- **2006**: Geoffrey Hinton proposed **Deep Neural Networks**, marking the beginning of deep learning.





- 2024: Geoffrey Hinton won Novel Prize in Physics

Time line of AI

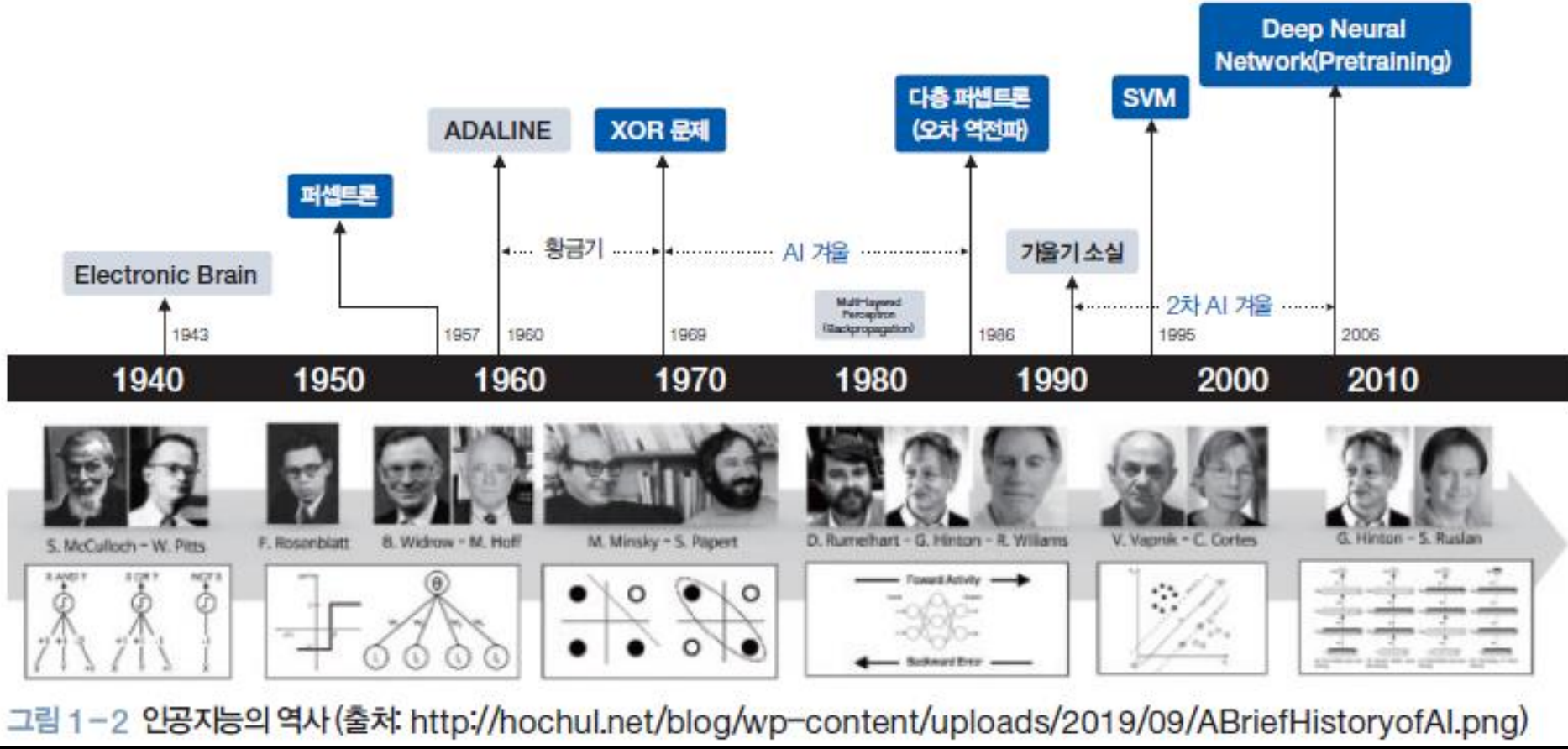


그림 1-2 인공지능의 역사 (출처: http://hochul.net/blog/wp-content/uploads/2019/09/ABriefHistoryofAI.png)

Explosive Growth of AI - Data

- **Data**
  - **Structured Data**
    - Organized in rows and columns following a specific format.
    - **Example:** Excel-style data with rows and columns.
  - **Unstructured Data**
    - Not arranged in a predefined manner.
    - **Example:** Videos, audio files, text (documents).
  - **Semi-structured Data**
    - Not stored in a structured format like relational databases but still retains some level of organization.
    - **Example:** Documents stored in **JSON (JavaScript Object Notation), XML, or HTML** format.

Explosive Growth of AI - Algorithm (1/2)

- **Algorithm (Model)**
  - Traditional Methods (Before 2010)
    - Extracted key features from images and used traditional machine learning algorithms such as Support Vector Machine (SVM) for classification.
  - Introduction of Deep Learning (2012)
    - Krizhevsky et al. proposed a deep learning-based model.
    - Used the raw image as input and produced classification results directly.
  - GoogLeNet (2014)
    - Utilized multiple small convolutional filters (leveraging receptive fields).
    - Introduced a deep stacking architecture with multiple layers.
  - ResNet (2015)
    - Introduced the Residual Learning approach, allowing previous information to be passed to the next layer.
    - Used multiple small convolutional filters (leveraging receptive fields).
    - Reduced the number of model parameters while maintaining high performance.

Explosive Growth of AI - Algorithm (2/2)

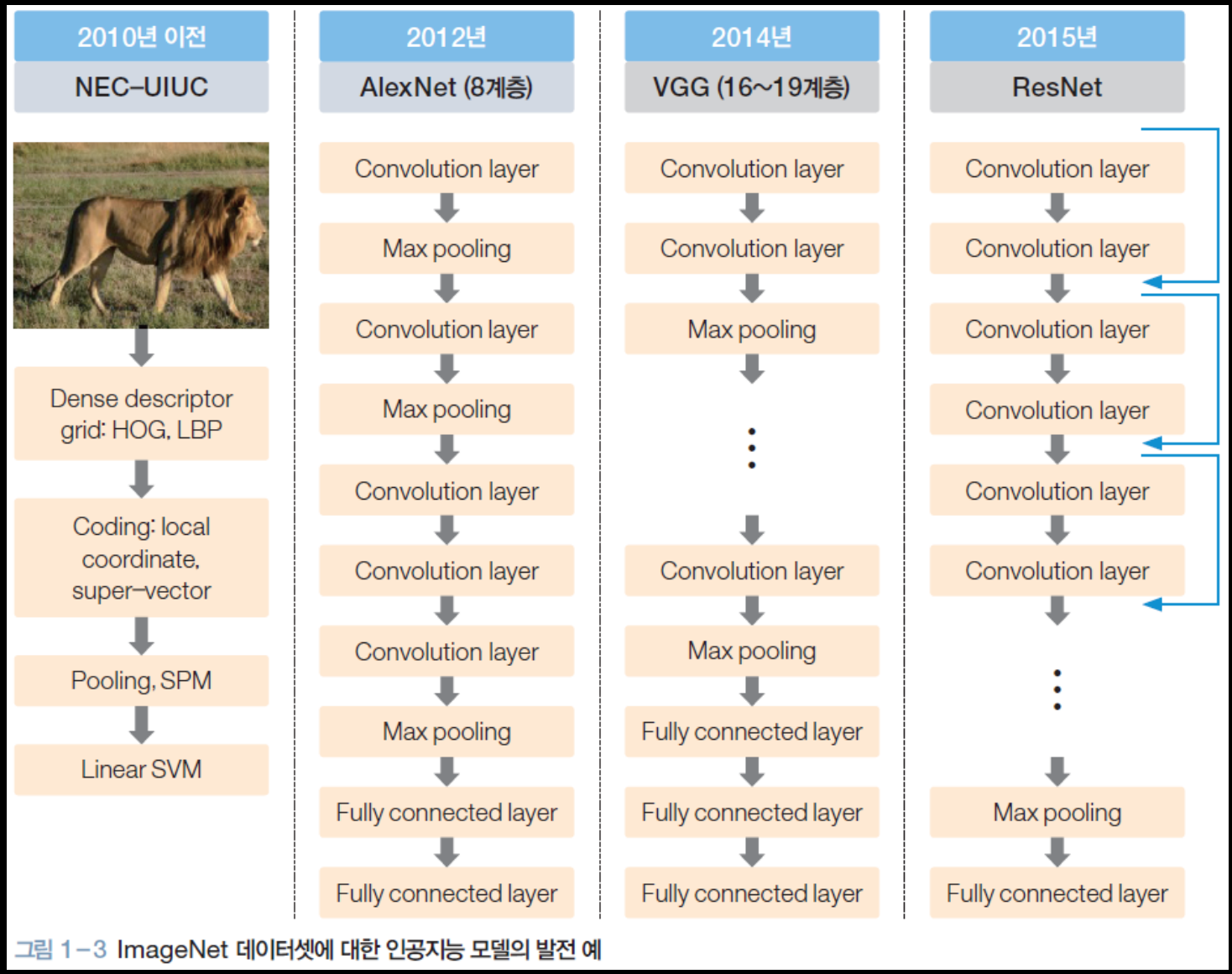


그림 1–3 ImageNet 데이터셋에 대한 인공지능 모델의 발전 예

Explosive Growth of AI - Computing

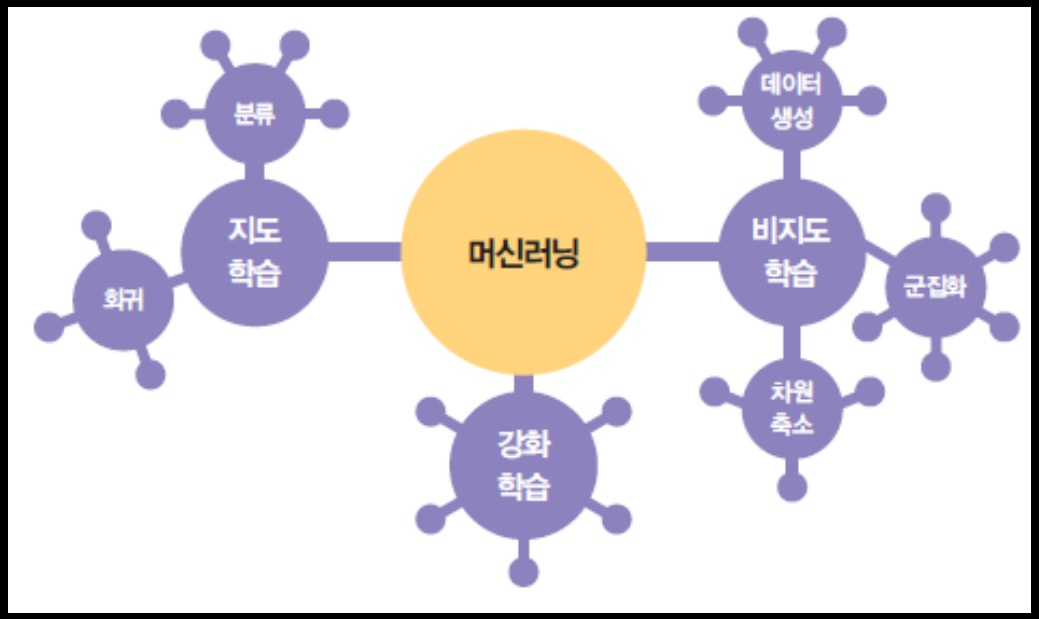
- **Computation (Computing Power)**
  - **Parallel Processing Capability of GPUs**
    - Performs multiple computations simultaneously, significantly improving learning efficiency.
  - **Performance Enhancement**
    - Greatly increases the learning capacity of AI models, enabling the training of **large-scale models**.
  - **Development of AI-Specific Processing Units**
    - Continuous improvement in computational performance through AI-enabled devices such as **IoT and AI semiconductors**.

Basics of Machine Learning

Fundamentals of Machine Learning

- **Performance improvement due to the increase in data, enhanced computing power, and improved algorithm models.**

- **The emergence of applications utilizing machine learning technology across various fields.**



Categories in ML

- **Supervised Learning**
  - Learns using labeled training data (with actual/true values) and then performs prediction and classification on test data.
  - Common applications include **regression** and **classification**.

- **Unsupervised Learning**
  - Trains the model using **unlabeled data**, performing tasks such as **clustering**, generating new data, and dimensionality reduction.

- **Reinforcement Learning**
  - Without predefined data, an **agent** interacts with the environment, learning through trial and error to choose actions that maximize rewards.

Supervised Learning - Procedures

- Training Phase: Generates the optimal function using observed data (Best Function)
- Testing Phase: Predicts the result using the learned function (Predict by Using Function).

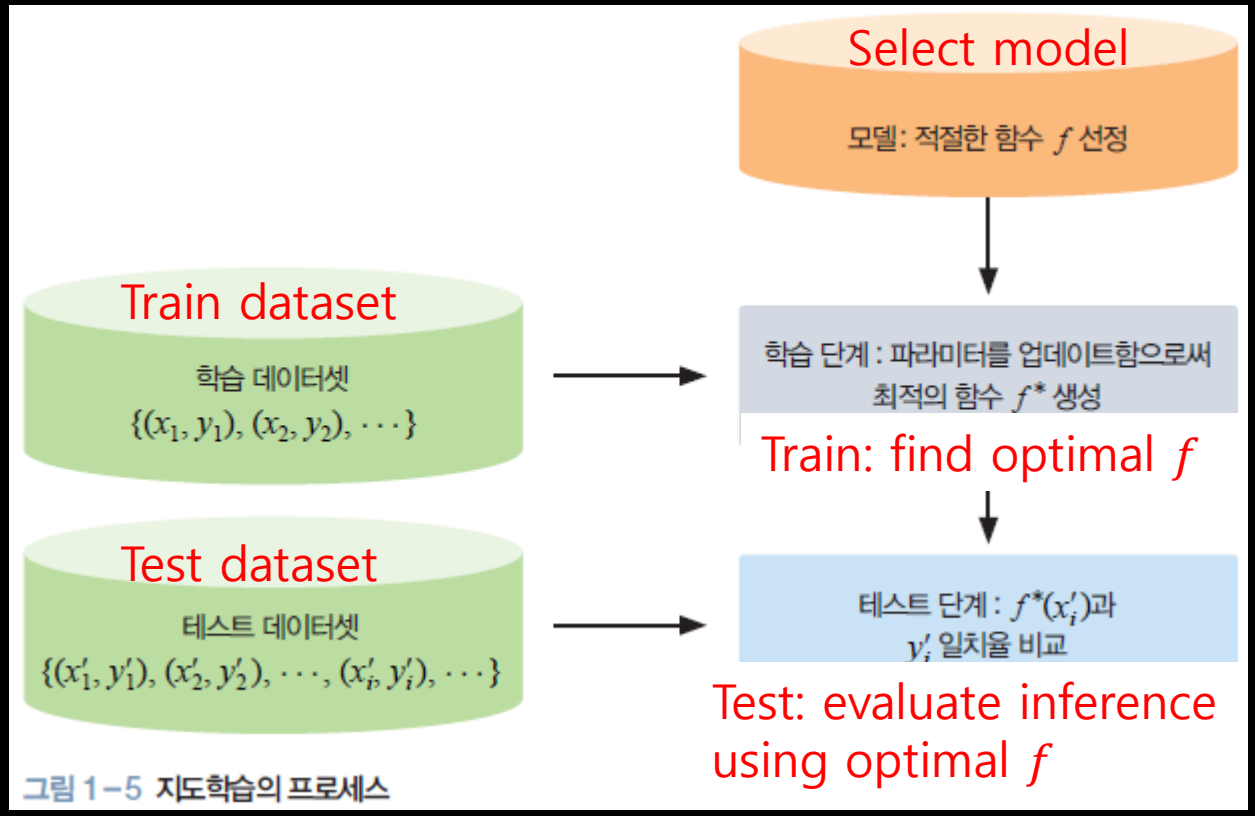
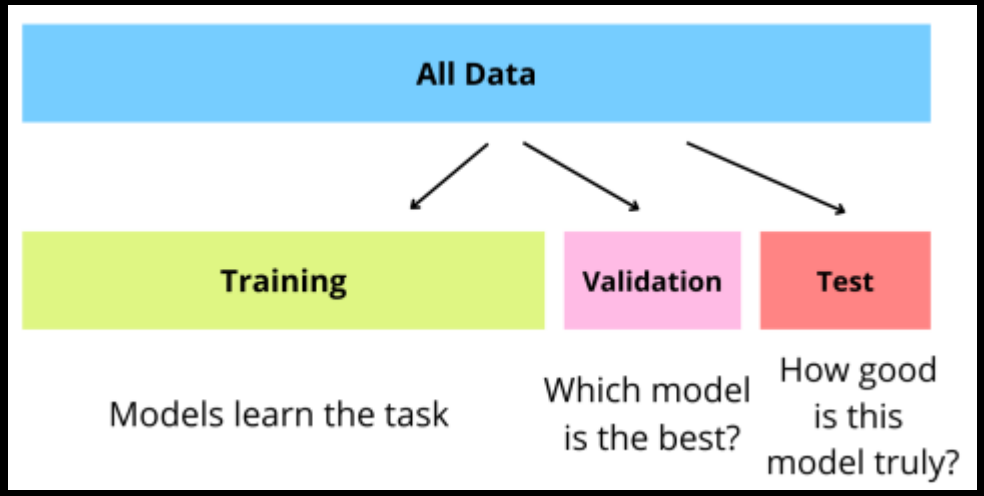


그림 1-5 지도학습의 프로세스

Supervised Learning - Dataset

Dataset



- Composed of a Training Dataset and a Test Dataset
- Typical split ratios:
  - 7:3 or 8:2 (Training Dataset : Test Dataset).
- Validation Dataset
  - Used during training to prevent the model from overfitting to the training dataset.

What's in next lecture

Main Topics

- Mathematics and NumPy Library for Machine Learning
  - Basic Mathematics for Machine Learning
  - NumPy Library for Machine Learning

Preparation

- Laptop or Desktop in Lab
- Download Practice Materials

  (GitHub Link, https://github.com/KMA-AIData/ML)

수고하셨습니다 ..^^..
Thank you!