Machine Learning Laboratory
POSTECH

[CSED/AIGS 515] - 2023 Spring

Machine Learning

Dongwoo Kim
POSTECH CSED/GSAI
dongwoo.kim@postech.ac.kr

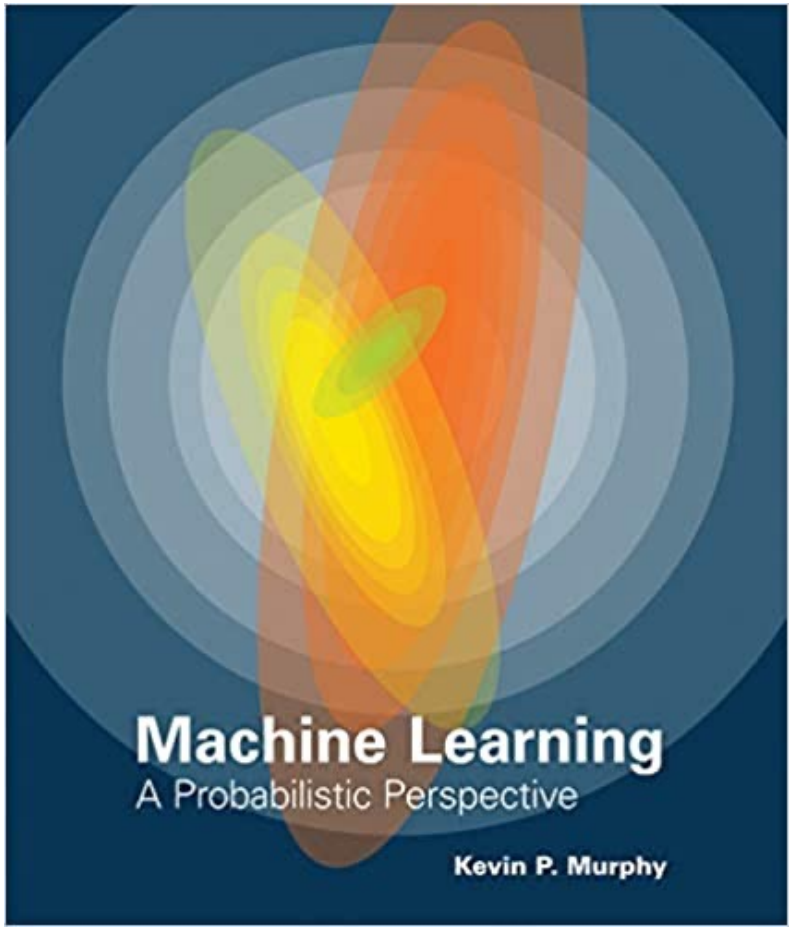
Basic Lecture Information

- Time: 11:00-12:15 (Tue/Thu)
- Classroom
  - 청암학술정보관 502호
- Office hour: contact by email
  - dongwoo.kim@postech.ac.kr

Main Textbook

- Machine Learning: A Probabilistic Perspective, by Kevin Murphy, MIT Press, 2012

- Draft of the book is available at the book webpage: https://probml.github.io/pml-book/book1.html



Attendance Check

- We will not check attendance for the first two weeks.
- We will check the attendance via mobile app.
  - The detailed instruction will be available soon.
  - But you need to satisfy minimum attendance requirement (3/4)

Exams & Assignments

- Midterm exam
- Final exam
- 5~6 assignments
  - Problem solve (handwriting)
  - Programming (numpy + pytorch)

- Midterm 30% / Final 30% / Assignment 40%
  - + need to satisfy minimum attendance requirement (3/4)

Important: **zero tolerance** for plagiarism/cheating

Questions

- Questions during lecture
  - Feel free to interupt
  - **Don’t be shy**

- Offline questions
  - Use Q&A board (Recommanded)
  - Contact by email for personal meeting (dongwoo.kim@postech.ac.kr)

- TA (assignment, grading, attendance,…)
  - Moonjeong Park + …
  - Use email address: csed515-ta@postech.ac.kr

Pre-requisite

- Mandatory
  - Basic programming (Python; numpy + pytorch)
  - Basic (vector) calculus

$$\frac{d \ln x}{dx} = \qquad\qquad \int \exp(-x) dx =$$

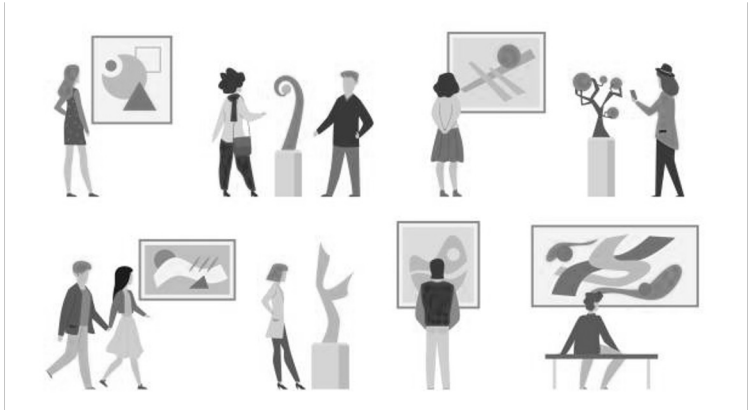
  - Probability (will review)
  - Linear algebra
- The first assignment will be about basic probability and algebra. If you feel not ready for these materials, you may consider taking this course later.
  - CSED490Y: Optimization for Machine Learning
  - CSED500: Advanced Linear Algebra for CSE
  - CSED530 Advanced Probability Theory for CSE

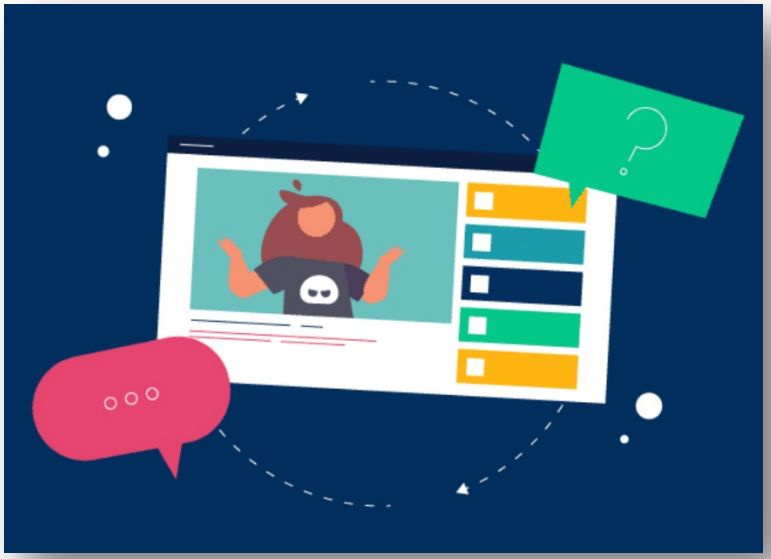
An Introduction to Machine Learning

- What is Machine Learning?
- Three Types of Machine Learning
- Course Overview

Era of Artificial Intelligence

- Recommandation



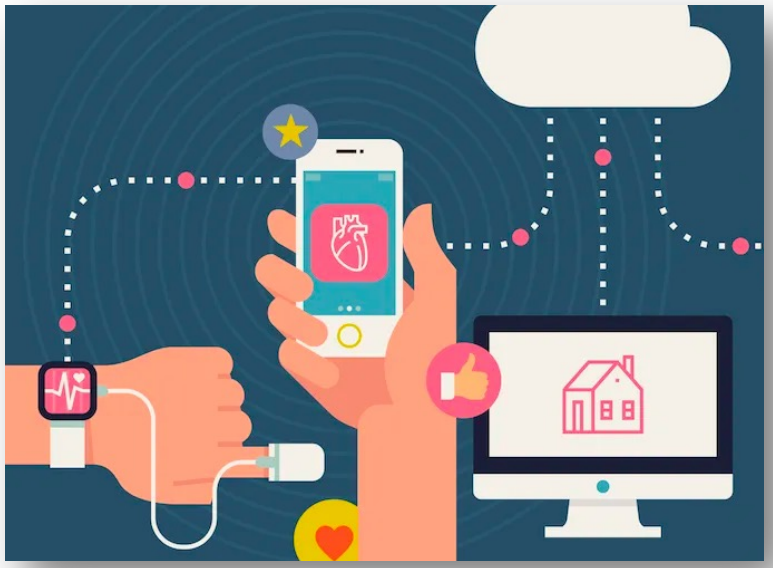


Era of Artificial Intelligence

- Self-Driving Car

Era of Artificial Intelligence

- Heath Care



Era of Artificial Intelligence

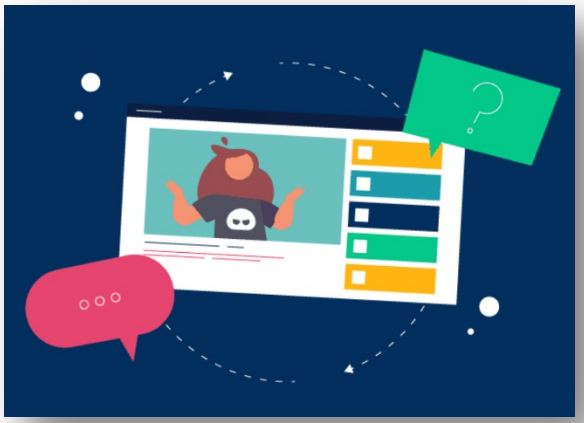
- Intelligent Chatbot

Do you know POSTECH?

Yes, I am familiar with POSTECH, which stands for Pohang University of Science and Technology. It is a private research university located in Pohang, South Korea, and it is considered one of the top universities in the country. POSTECH was founded in 1986 and has a strong focus on science and engineering research, with particular strengths in materials science, nanotechnology, energy, and biotechnology. The university has a highly competitive admissions process and is known for its rigorous academic programs and innovative research.

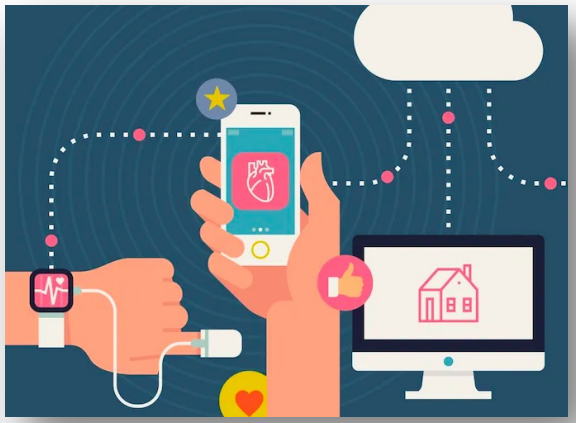
Era of Artificial Intelligence

Recommander System



Self-Drving Car

Healthcare



Household Robot

K. Wyrobek, E. Berger, H.F.M. Van der Loos, and K. Salisbury. ICRA, 2008.

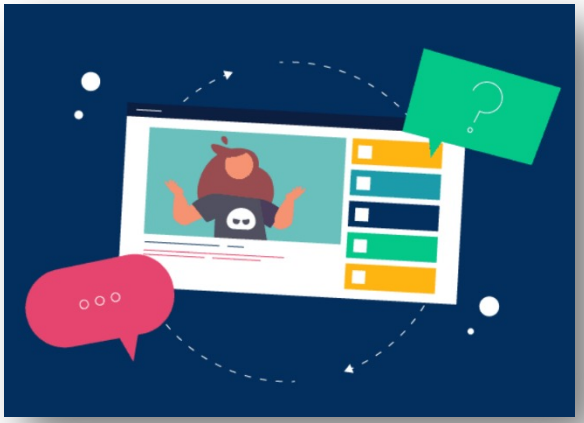
Key Challenge: Software!

- Tele-operated robots, i.e., hardware has been ready since 2008!

Images from CSAIL@MIT, 2017

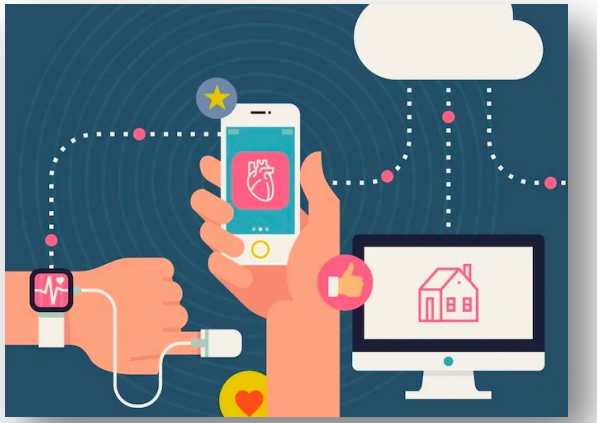
A Paradaim of Building Software

Recommander System



Self-Drving Car

Healthcare



Machine Learning

An Introduction to Machine Learning

- What is Machine Learning?
- Three Types of Machine Learning
- Course Overview

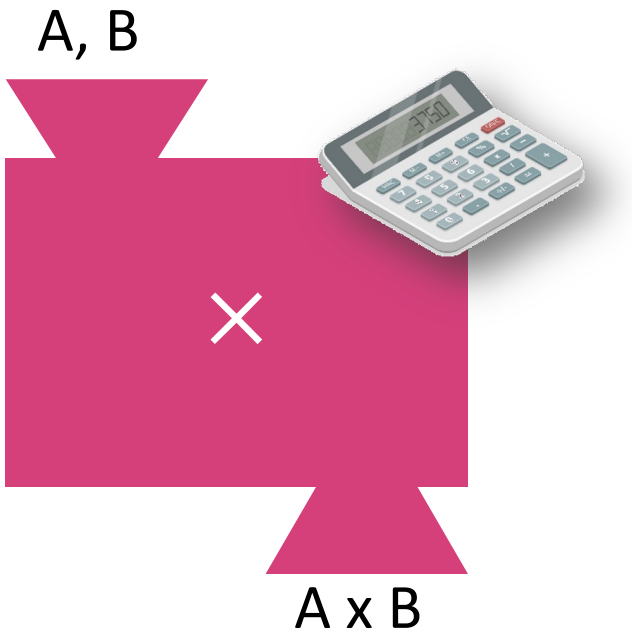
Machine

= Hardware/software to perform an intended function

Input

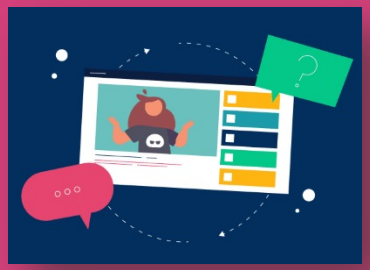
$$f(x) = y$$

Output



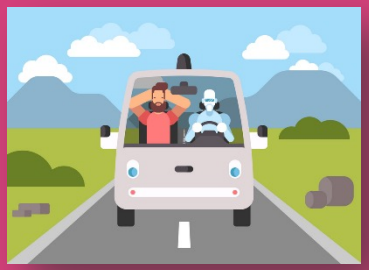
Examples

User history



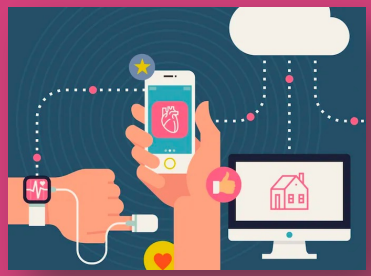
Item list
(preferred)

Destination



Car control
(safe, fast)

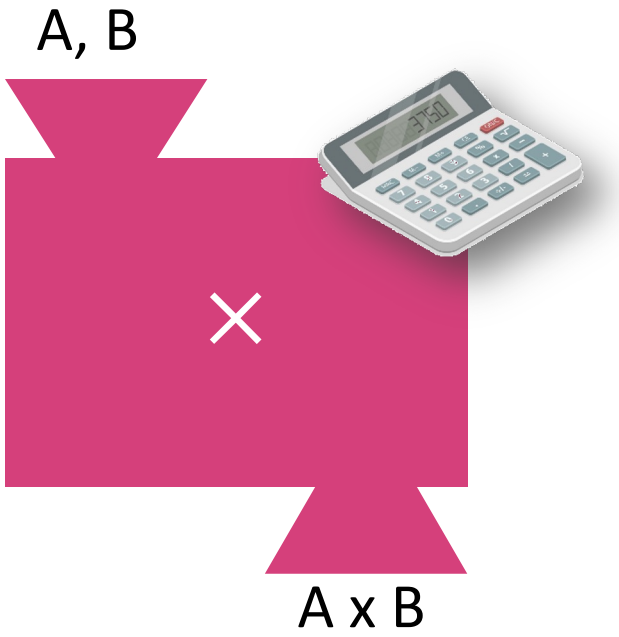
Health data



Diagnosis
(accurate)

Traditional Computer Science

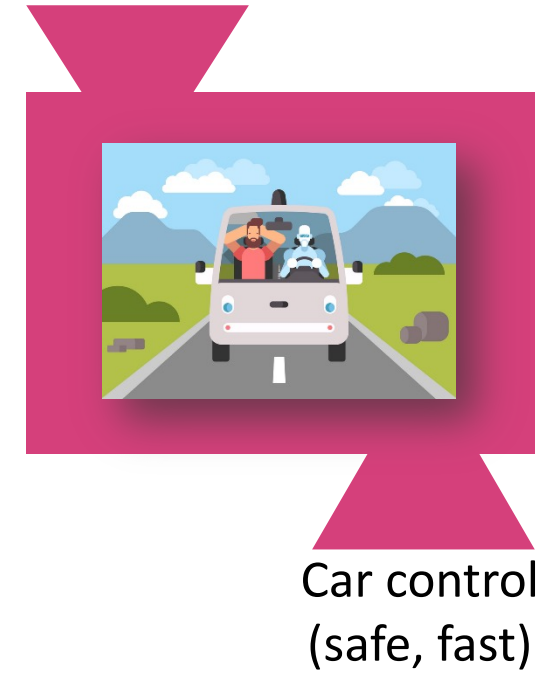
= writing a program code

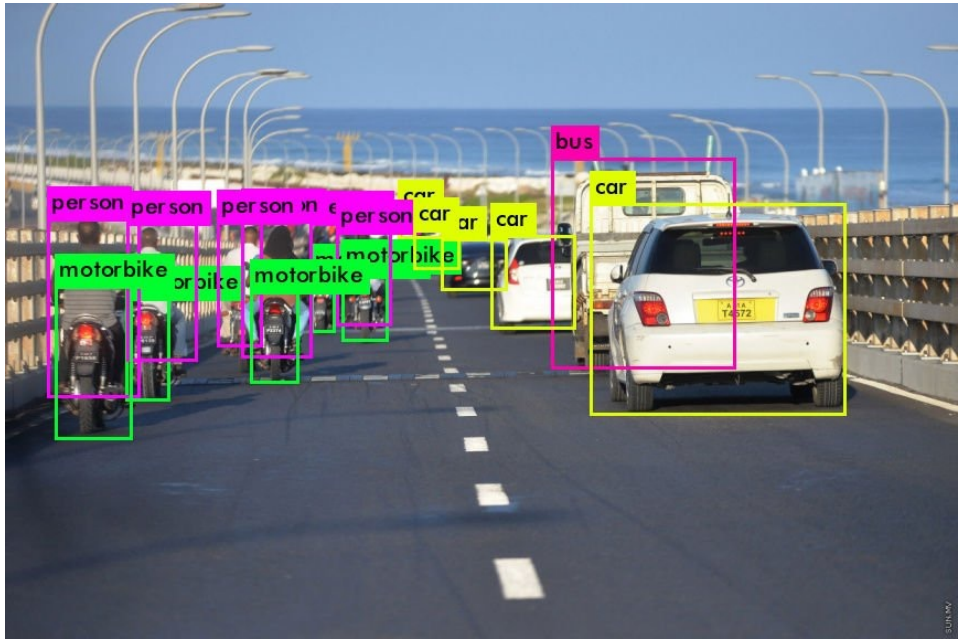


$$\underbrace{(A+A+\ldots+A)}_{\text{B times}}$$

Limitation of Traditional Paradigm (1)

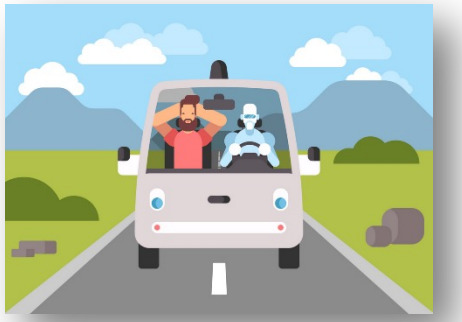
- Sometimes, we do not have a clear description about the function that we want

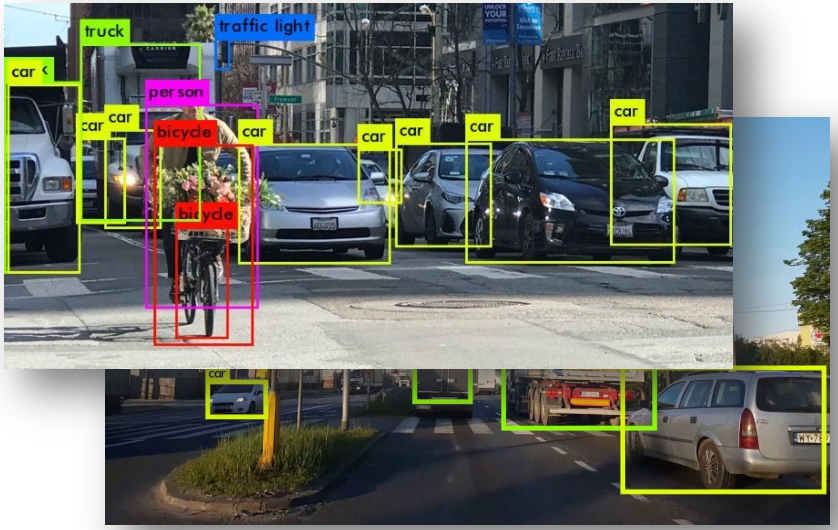


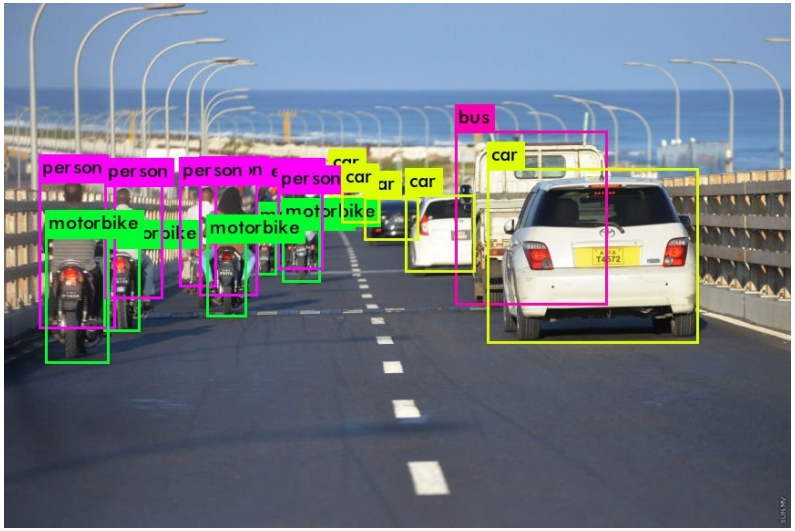


Limitation of Traditional Paradigm (2)

- Traditional paradigm
  - Program all the instructions
- New paradigm
  - Provide examples, and let machine learn

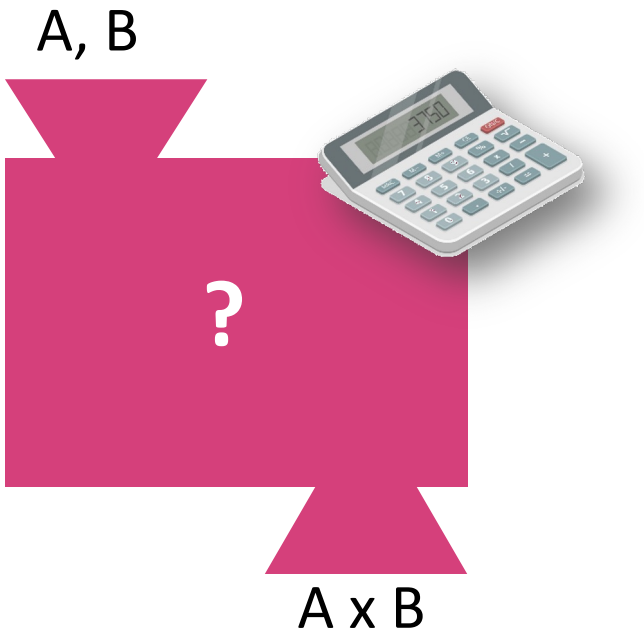






New Paradigm: Machine Learning

= Letting the machine learn the function from dataset



f(5, 2) = 10

f(3, 7) = 21

f(4, 4) = 16

f(2, 3) = 6

f(8, 7) = 56

f(9, 3) = 27

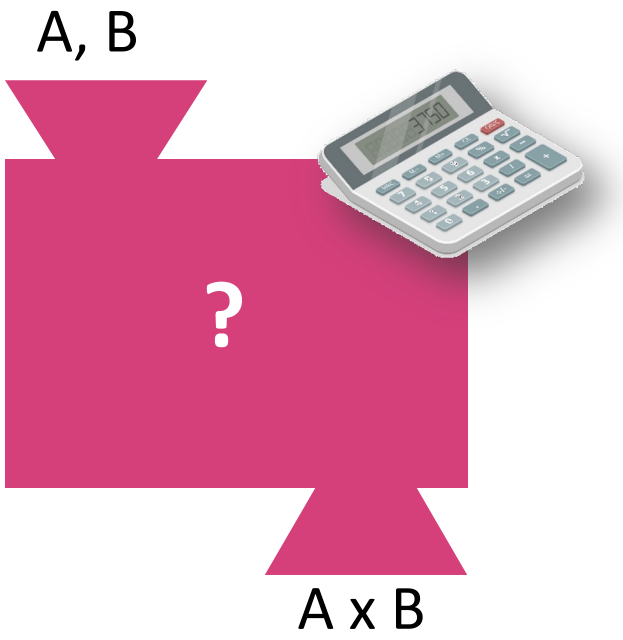
…

Dataset

: a set of samples of mapping between input and output

New Paradigm: Machine Learning

= Letting the machine learn ~~the function~~ from dataset pattern

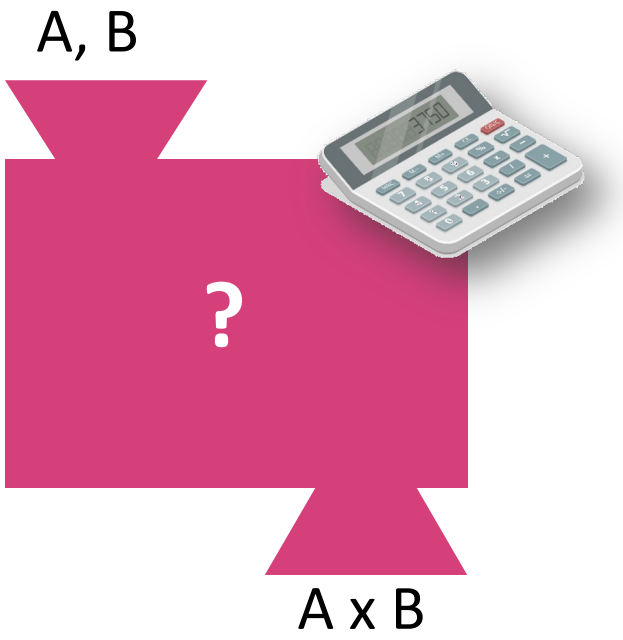


B \ A	1	2	3	4	5	6	7	8	9
1			3	4		6			9
2	2		6		10		14		18
3	3			12		18		24	27
4	4	8	12	16		24	28		
5					25		35	40	
6	6	12	18		30	36		48	54
7	7		21	28	35		49	56	
8		16		32	40		56		72
9	9	18		36		54	64		81

Pattern = Multiplication
(a simple explanation)

New Paradigm: Machine Learning

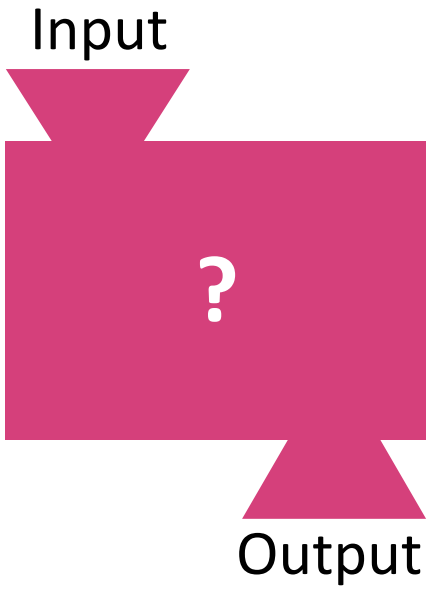
= Letting the machine learn ~~the function~~ from dataset pattern



B \ A	1	2	3	4	5	6	7	8	9
1	1	2	3	4	5	6	7	8	9
2	2	4	6	8	10	12	14	16	18
3	3	6	9	12	15	18	21	24	27
4	4	8	12	16	20	24	28	32	36
5	5	10	15	20	25	30	35	40	45
6	6	12	18	24	30	36	42	48	54
7	7	14	21	28	35	42	49	56	64
8	8	16	24	32	40	48	56	64	72
9	9	18	27	36	45	54	64	72	81

Computer Science vs. Machine Learning

Traditional Computer Science
: **We find** pattern; and translate it to machine



Machine Learning
: **A Machine finds** pattern from dataset

Framework of Machine Learning

- Dataset and … ?

Age, Weight

Heart disease?

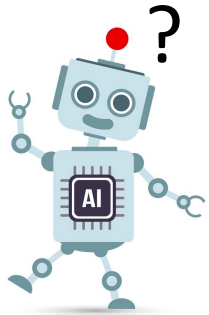
f(42-year, 60 kg) = No

f(8-year, 44 kg) = No

f(75-year, 62 kg) = Yes

f(24-year, 104 kg) = Yes

…



Framework of Machine Learning

- Dataset and model about function

Age, Weight

Heart disease?

f(42-year, 60 kg) = No

f(8-year, 44 kg) = No

f(75-year, 62 kg) = Yes

f(24-year, 104 kg) = Yes

…

Obs1. overweight → heart disease

Obs2. old → heart disease

(Pattern)

Framework of Machine Learning

- Dataset and model about function

e.g., **linear model**

$$A \times \text{Age} + B \times \text{Weight} > C?$$

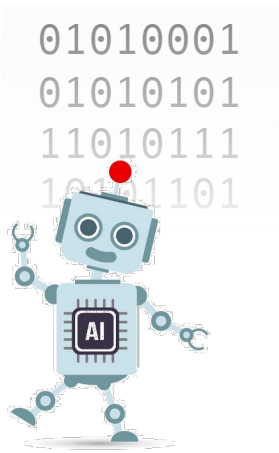
f(42-year, 60 kg) = No

f(8-year, 44 kg) = No

f(75-year, 62 kg) = Yes

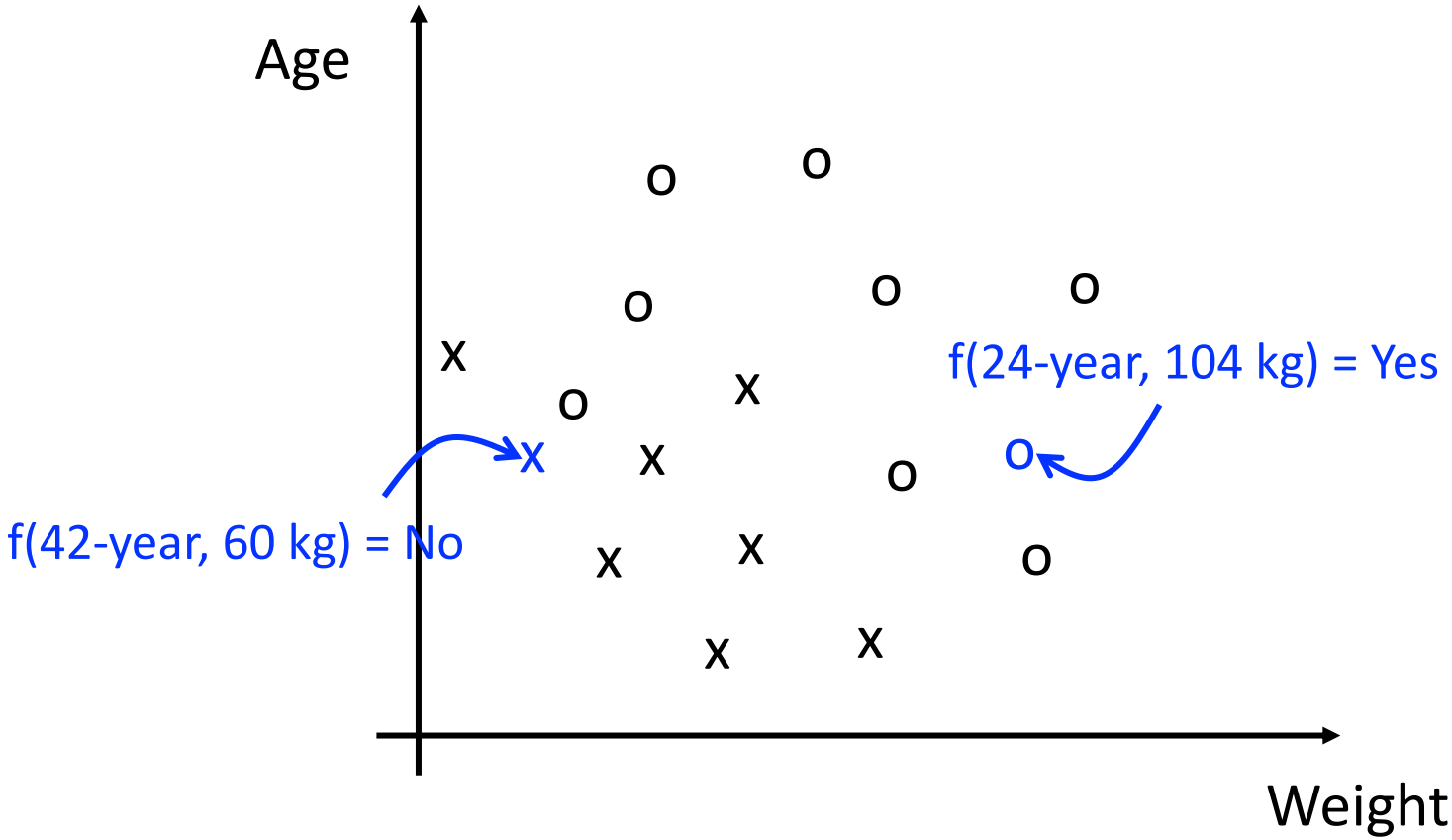
f(24-year, 104 kg) = Yes

…



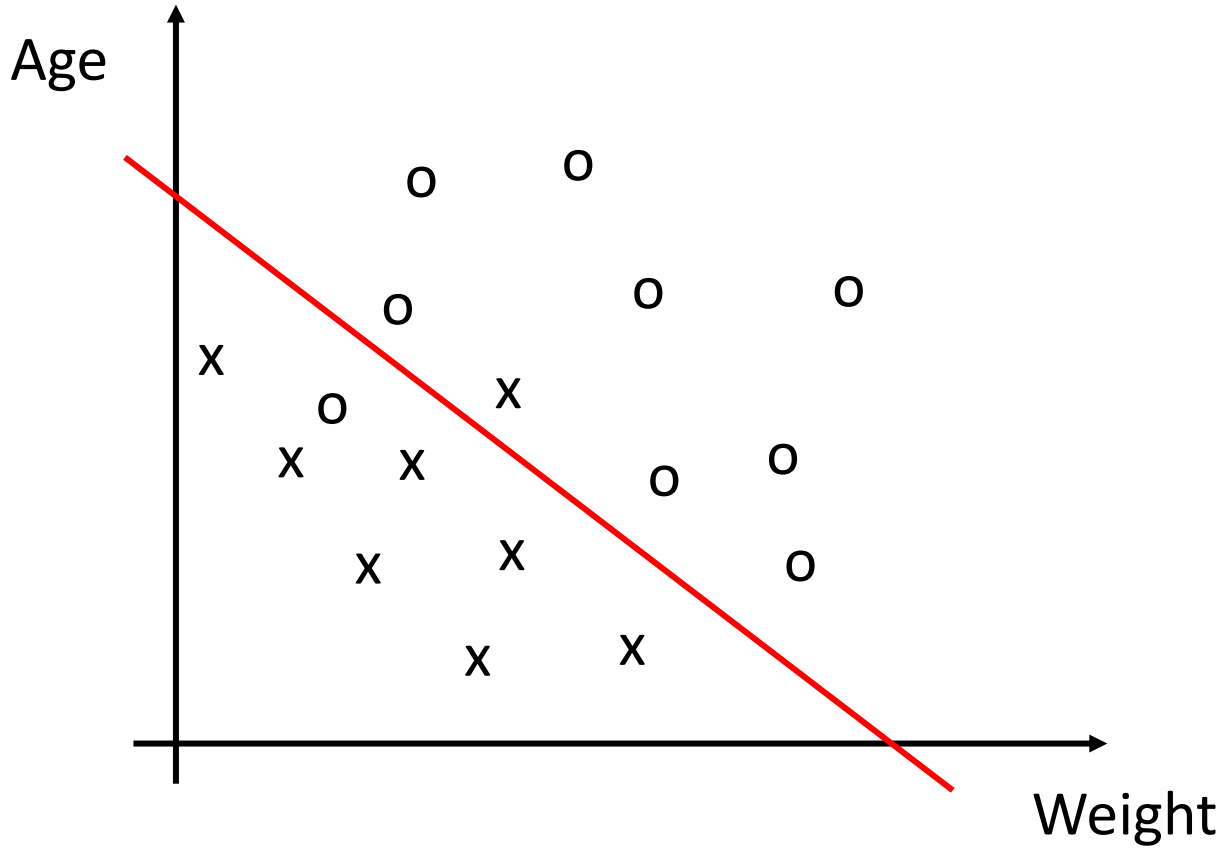
Graphical Interpretation of Linear Model

A x Age + B x Weight > C?



Graphical Interpretation of Linear Model

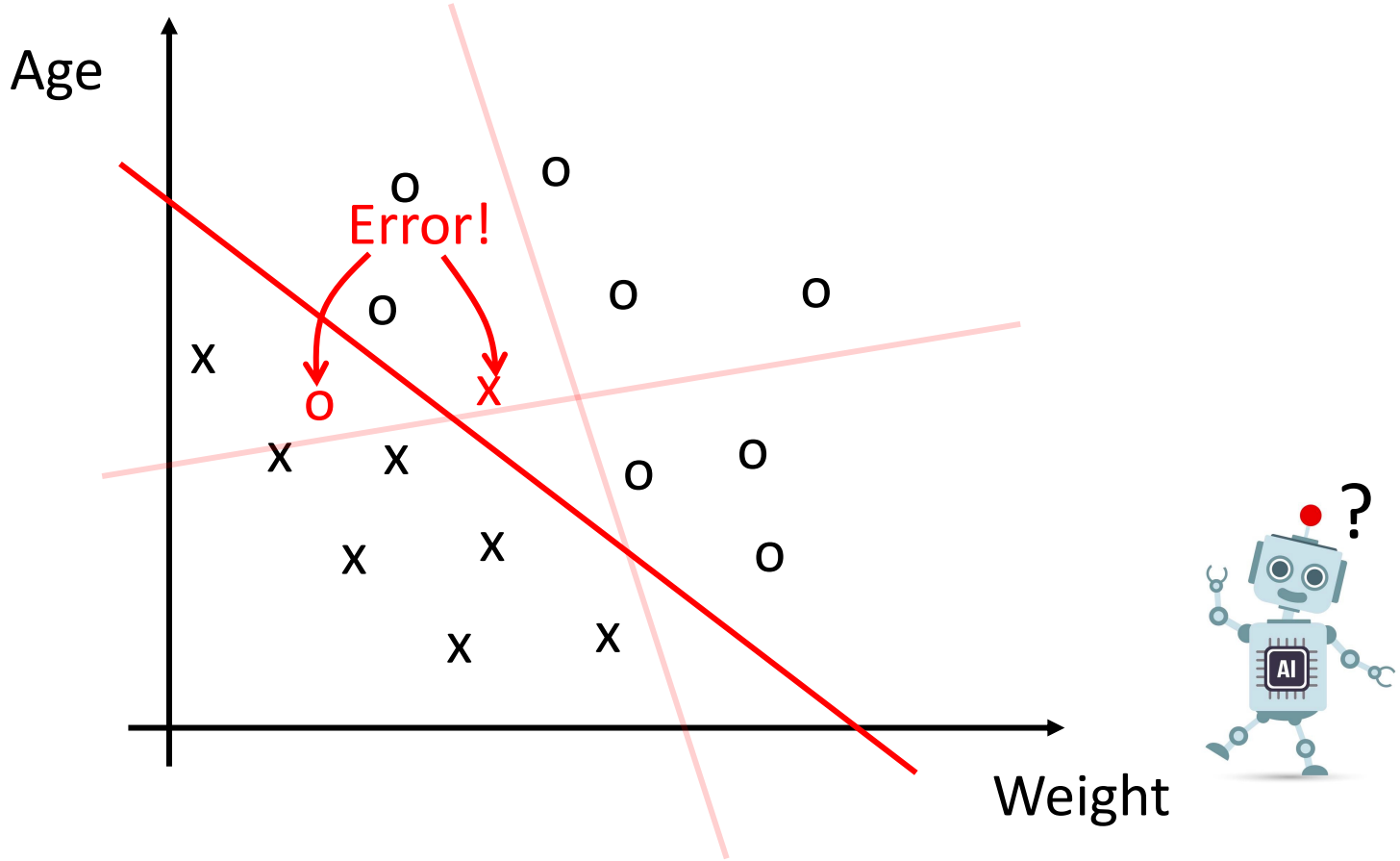
A x Age + B x Weight > C?



Graphical Interpretation of Linear Model

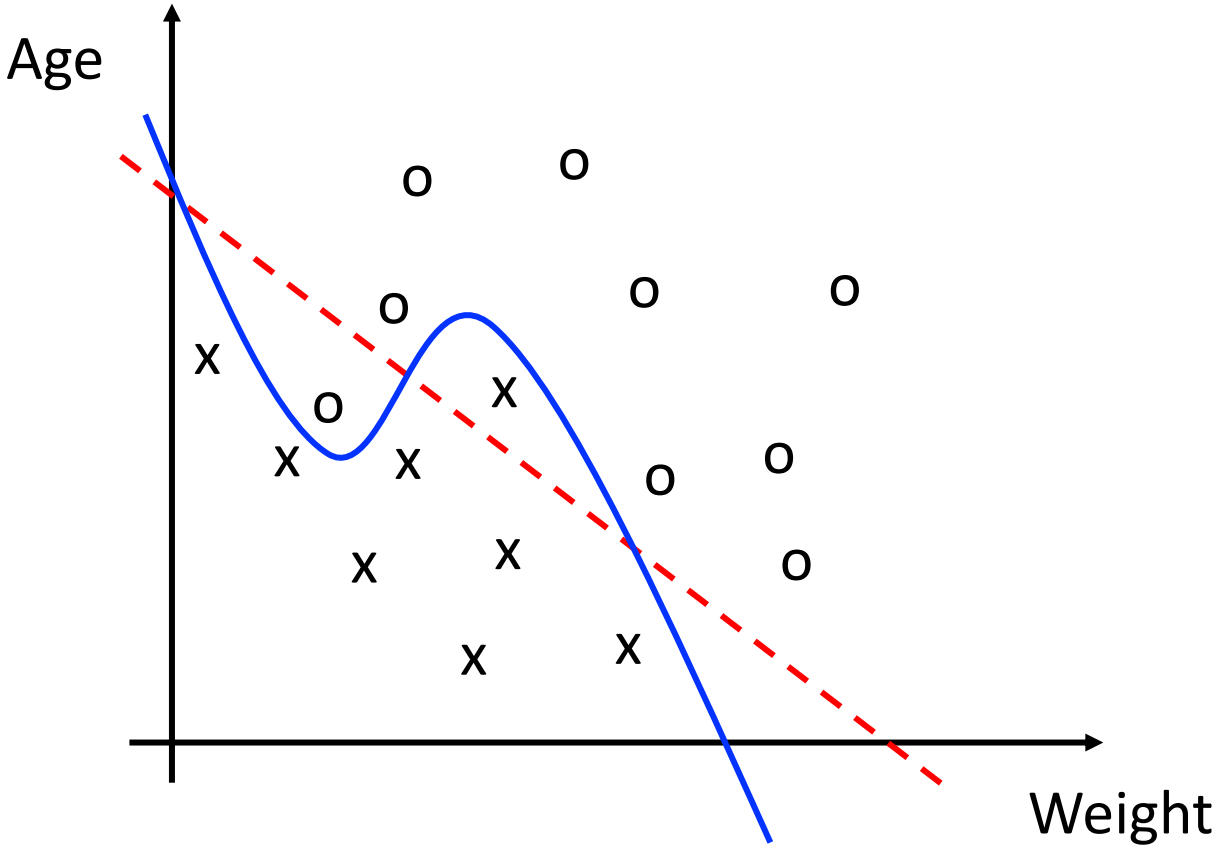
- Learning seeks best function minimizing error

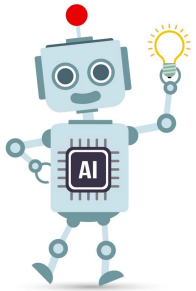
0.47 x Age + 0.92 x Weight > 90?



Better Model?

$A \times Age + B_3 \times Weight^3 + B_2 \times Weight^2 + B_1 \times Weight > C$?



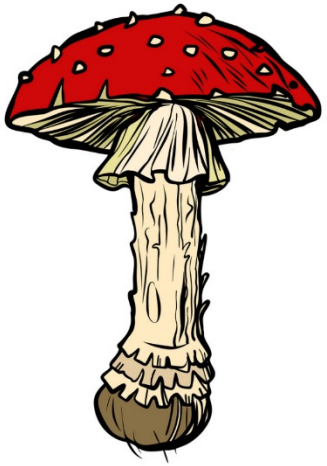


An Introduction to Machine Learning

- What is Machine Learning?
- Three Types of Machine Learning
- Course Overview

Supervised Learning

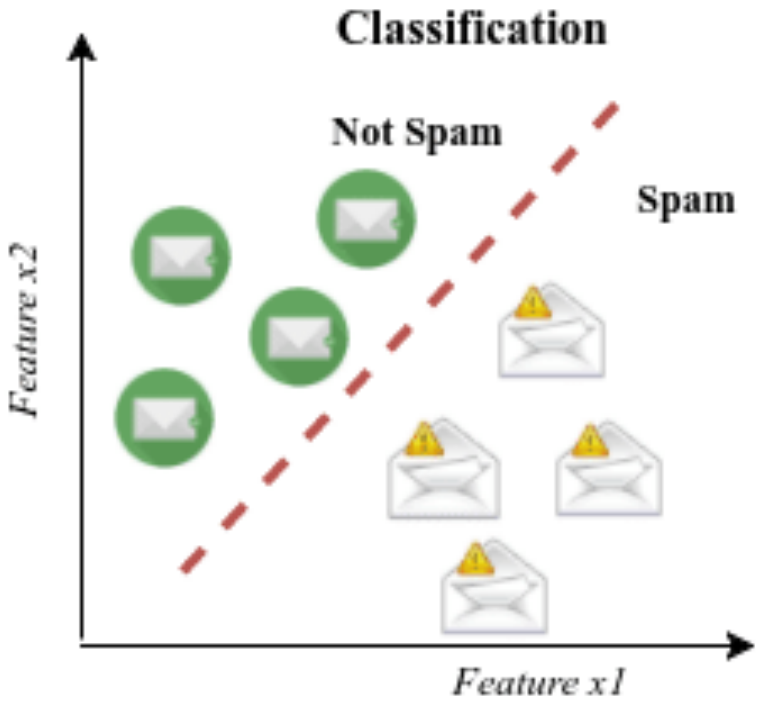
: Dataset is labeled in advance



Edible?
Poisonous?

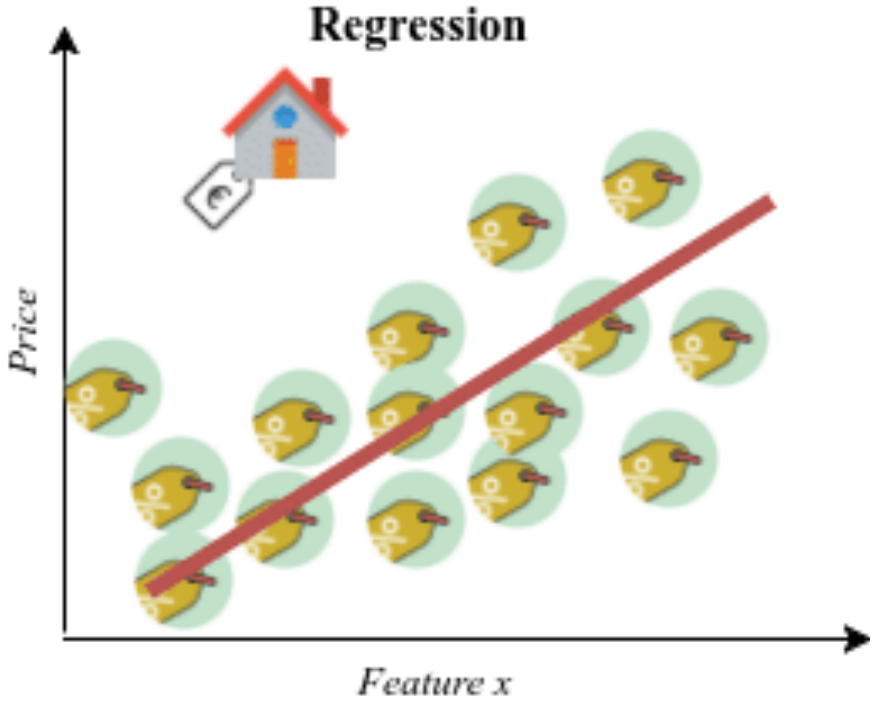
Classification, e.g., Spam Filtering

: to predict categorical labels



Regression, e.g., House Price Prediction

: to predict numerical values

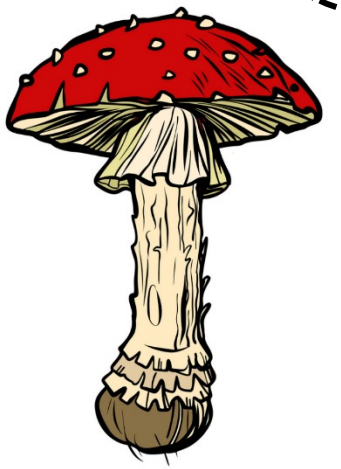


Unsupervised Learning

: Dataset has no labels

Cluster 1

Cluster 2



Which cluster?

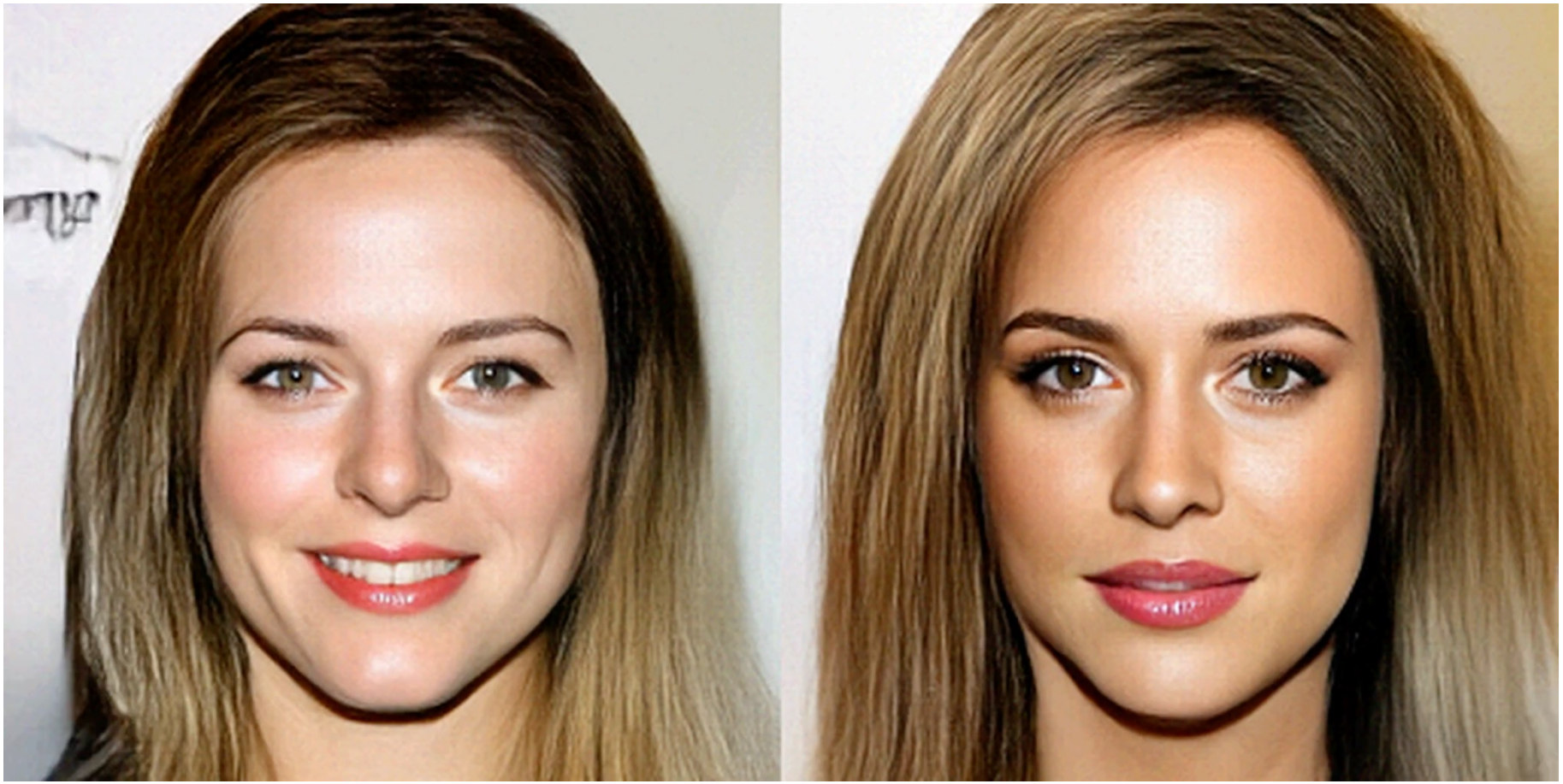
Generative Model

Training data ~ $p_{data}(x)$

Generated samples ~ $p_{model}(x)$

CycleGAN. Zhu et al. 2017.

Generative Model



Reinforcement Learning

: Label is a real-valued reward signal (possibly delayed)

Amanita (toxic): (+2; -5)

t	0	1	2	3
Action				
Reward	2	-4	2.5	-3

Shiitake (edible): (+1; +0.5)

(instant. reward; reward after 1 day)

Reinforcement in Animal Psychology

- Training animals:
  - Positive reinforcements: pleasure and food
  - Negative reinforcements: pain and hunger

Reinforcement learning (RL) performs the same with computers!

AI Playing Atari Games (2013)

AlphaGo vs. Sedol Lee (2016)

AlphaGo defeated the legendary Go player, Sedol Lee

An Introduction to Machine Learning

- What is Machine Learning?
- Three Types of Machine Learning
- Course Overview

Tentative Schedule

Categories
Supervised Learning
Unsupervised Learning
Reinforcement Learning
Learning Theory and Other Topics
Graphical Model
Deep Learning
No class

Index	Week	Topic
1	1-1	Introduction to Machine Learning
2	1-2	Review on Probability Theory
3	2-1	Generative Model and Parameter Estimation (1)
4	2-2	Generative Model and Parameter Estimation (2)
5	3-1	Linear Regression (1)
6	3-2	Linear Regression (2) / Logistic Classification (1)
7	4-1	Logistic Classification (2)
8	4-2	Support Vector Machine and Optimization (1)
9	5-1	Support Vector Machine and Optimization (2)
10	5-2	Multi-class SVM and Kernel Method
11	6-1	Neural Network Model and Neural Network Training (1)
12	6-2	Neural Network Model and Neural Network Training (2)
13	7-1	Neural Network Model and Neural Network Training (3)
14	7-2	Network Model and Neural Network Training (4) / Midterm
15	8-1	**Midterm Exam Week**
16	8-2	**Midterm Exam Week**
17	9-1	Some Insights on Model Selection
18	9-2	Mideterm Exam Review & Graphical Model
19	10-1	Graphical Model
20	10-2	Belief Propagation (BP)
21	11-1	Graph Construction
22	11-2	Clustering
23	12-1	EM algorithm
24	12-2	Principal Components Analysis
25	13-1	Non-linear PCA and Autoencoder
26	13-2	VAE and GAN
27	14-1	VAE and GAN
28	14-2	MDP and Dynamic Programming
29	15-1	RL (2)
30	15-2	RL (2)
31	16-1	**Final Exam Week**
32	16-2	**Final Exam Week**

Questions?