The Introduction To Artificial Intelligence

Yuni Zeng yunizeng@zstu.edu.cn

2025-2026-1

Machine Learning

- Part I Brief Introduction to AI & Different AI tribes
- Part II Knowledge Representation & Reasoning
- Part III AI GAMES and Searching
- Part IV Model Evaluation and Selection
- Part V Machine Learning

What is Machine Learning?

A widely quoted and formal definition is "*A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E*"

"如果一个程序在某类任务T中，受性能指标P的度量，其性能能随着经验值E的上升而不断提升，这个程序就能从与任务T和性能指标P相关的经验值E中学习。"

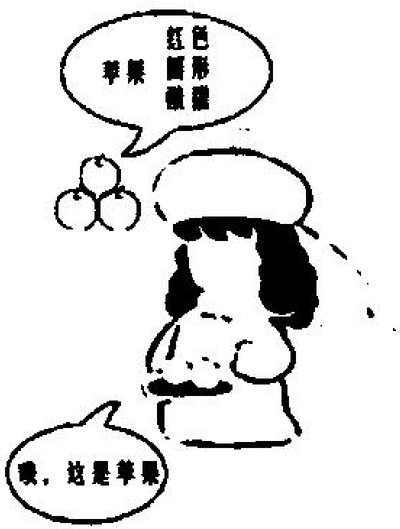
What is Machine Learning?

- **机器学习（Machine Learning，ML）**是一门专注于利用**经验**，通过计算技术来模拟和实现人类学习过程的学科。

数据

学习算法

知识和规律



What is Machine Learning?

- **Data is essential.** (**数据**是不可或缺的基石)

街道类型

建造年份

实际房价

收集数据

What is Machine Learning?

- **Data is essential. (数据是不可或缺的基石)**

收集数据

构造
数据集

特征（Feature）

标签（Label）

街道类型	建造年份	实际房价
商业街	2003	10
商业街	2023	20
普通街区	2010	10
…	…	…

样本

学习过程是否提供标签

- <u>**监督学习（Supervised Learning）**</u>
- <u>**无监督学习（Unsupervised Learning）**</u>

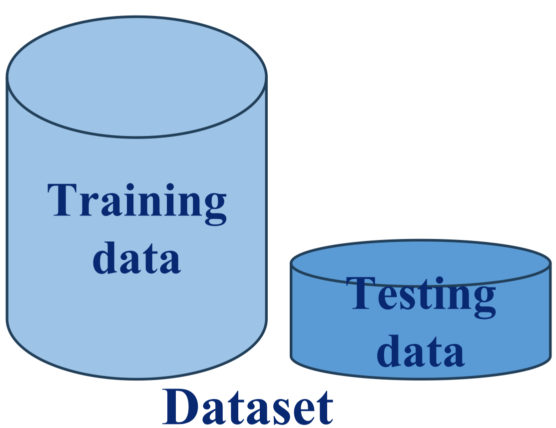
What is Machine Learning?

学习过程是否提供标签

(Learning with/without labels)

- <u>**监督学习（Supervised Learning）**</u>
- <u>**无监督学习（Unsupervised Learning）**</u>

- **监督学习：**利用一组**已知输入和对应标签**的数据集来训练模型，使模型能够学习到一个从输入到输出的映射关系
- **无监督学习：**直接对这些未标注数据进行建模分析，以实现相应的学习任务



- **训练集**的样本用来训练模型
- **测试集**用来检验模型的性能，即利用学习得到的模型参数对测试样本进行预测，并通过评价指标评测预测结果与标签的接近程度

What is Machine Learning?

学习过程是否提供标签
(Learning with/without labels)

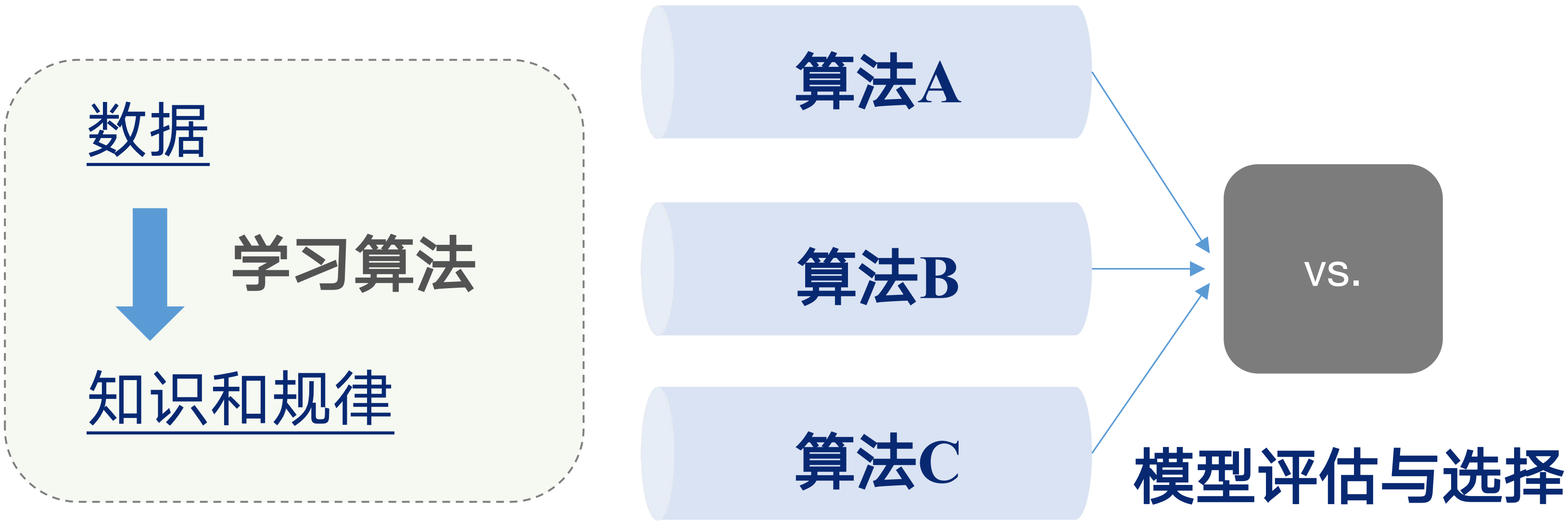
- **监督学习（Supervised Learning）**
- **无监督学习（Unsupervised Learning）**

标签的属性类型

- **分类任务（Classification Task）**
  - 线性判别分析、K近邻算法、**决策树**、支持向量机等
  - 房价值预测、气温预测、股市走势预测
- **回归任务（Regression Task）**
  - **线性回归**、逻辑回归等
  - 房价的高/低、预测明天是/否降雨、股市的涨/跌趋势

■ 实际上，分类模型与回归模型在底层逻辑上存在**共通之处**，通过适当调整，某些经典算法如逻辑回归和决策树能够灵活应用于两类任务之中。

Machine Learning



■ **目标：**选择**泛化能力强**的模型

不仅能在**训练数据集**上展现优异表现，更需在**未知新样本**上也表现出色

Machine Learning

- *1. Different ML methods*
- 2. Today's Machine learning

1. Different ML methods

Supervised learning

Unsupervised learning

1. Different ML methods

Supervised learning

Unsupervised learning

1. Different ML methods

Supervised Learning

WIKIPEDIA
The Free Encyclopedia

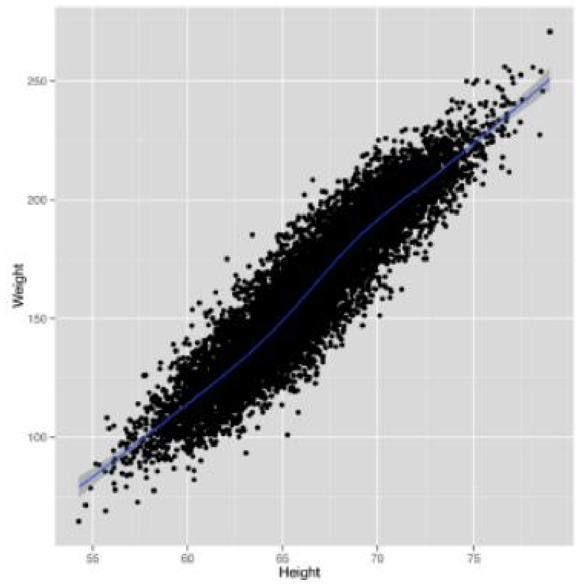
Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs.

- Regression
- Classification

1. Different ML methods

- What is regression?

Regression is to relate input variables to the output variable, to either predict outputs for new inputs and/or to interpret the effect of the input on the output.



Height is correlated with weight.

1. Different ML methods

- Two goals of regression

Prediction

wish to predict the output for a new input vector

e.g. What is the weight of a person who is 170 cm tall?

Interpretation

Understand the effect of inputs on output

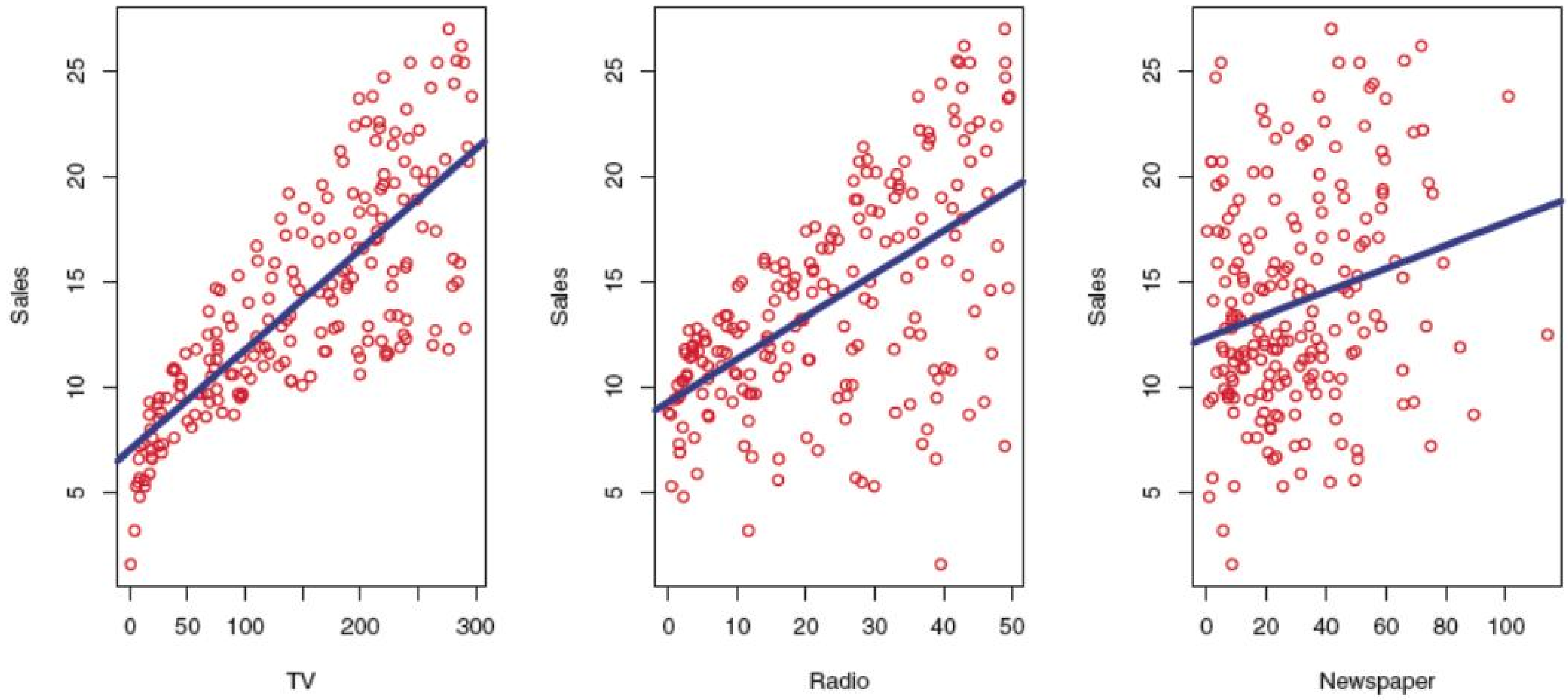
e.g. Are taller people heavier too?

For both the goals, we need to find a function that approximates the output "well enough" given inputs.

$$\boxed{y_n \approx f(x_n), \text{ for all } n}$$

1. Different ML methods

Regression --- example

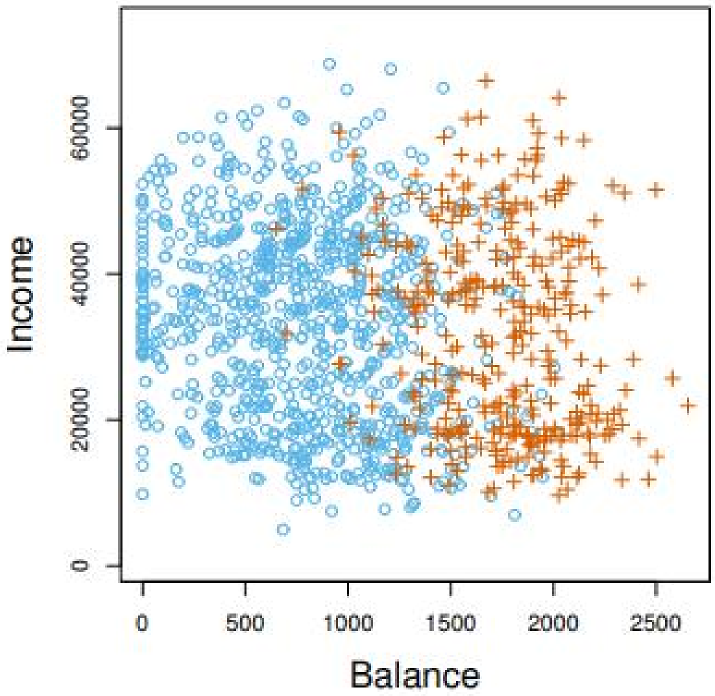


How does advertisement in TV, radio, and newspaper affect sales?

1. Different ML methods

Classification

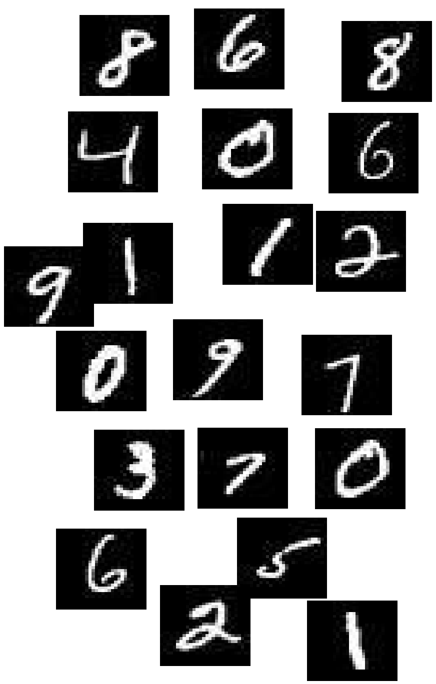
- Classification is same as regression but now y_n is binary or has finite values.
- Examples: object detection, face detection, hand-written digits recognition.



1. Different ML methods

Classification – An example

Training set



My baby, I will show you the digits today.
Let’s repeat it again…

eight.

It’s eight.

Mother knows the label for each training data.

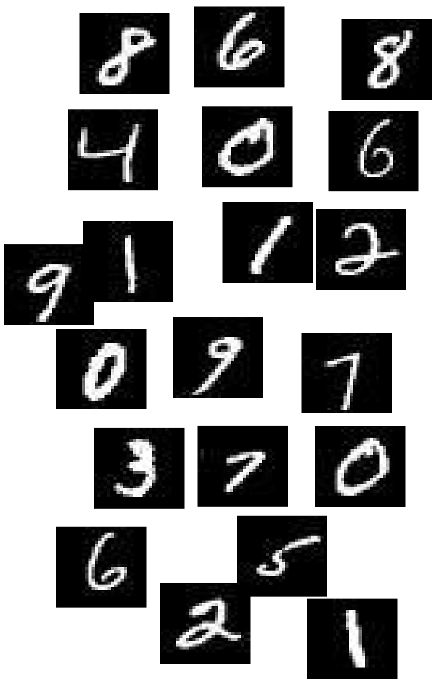
1. Different ML methods

- Classification – An example

Training now …

Training set

Let me see whether you know these digits.



Training set

Yes, I know all of them. They are eight, six, eight, four…

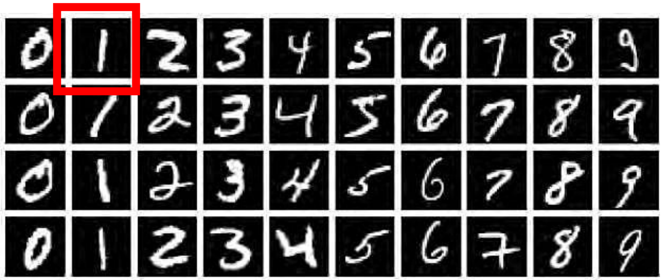
Yes! You know all the digits!

Recognition accuracy on training set:100%

1. Different ML methods

Classification – An example

The testing set: independent from the training set.



The central challenge in machine learning is that we must perform well on inputs—not just those on which our model was trained. The ability to perform well on previously unobserved inputs is called generalization.

My baby, I will test you on what you learned.

What’s this?

It’s one. √

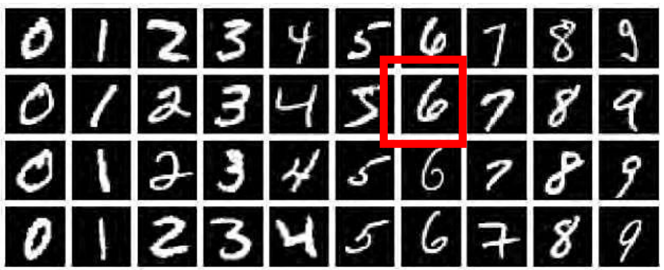
Testing now …

1. Different ML methods

Classification – An example

My baby, I will test you on what you learned.

The testing set: independent from the training set.



The central challenge in machine learning is that we must perform well on inputs—not just those on which our model was trained. The ability to perform well on previously unobserved inputs is called generalization.

What's this?

It's six. √

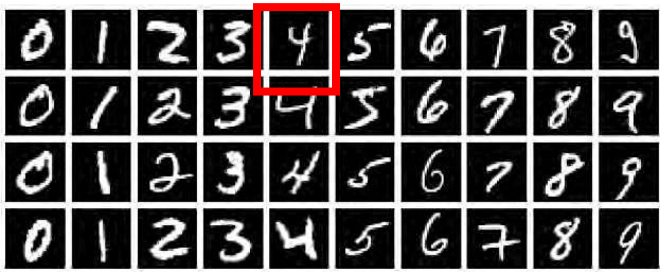
Testing now …

1. Different ML methods

Classification – An example

My baby, I will test you on what you learned.

The testing set: independent from the training set.



The central challenge in machine learning is that we must perform well on inputs—not just those on which our model was trained. The ability to perform well on previously unobserved inputs is called generalization.

It's nine. X

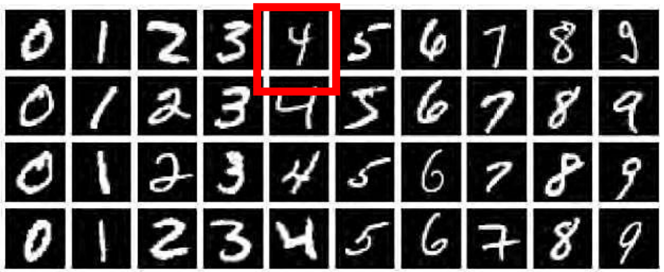
Testing now …

1. Different ML methods

Classification – An example

My baby, I will test you on what you learned.

The testing set: independent from the training set.



The central challenge in machine learning is that we must perform well on inputs—not just those on which our model was trained. The ability to perform well on previously unobserved inputs is called generalization.

Let me see.
You've answered 38 of the 40 digits correctly.
So you scored 95 points (38/40=95%).

Recognition accuracy on testing set :95%

Testing now …

1. Different ML methods

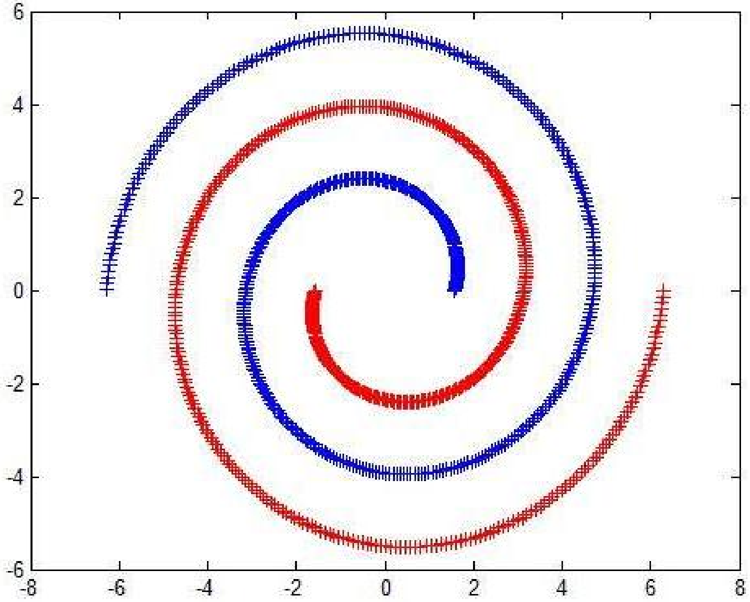
Supervised learning

Unsupervised learning

1. Different ML methods

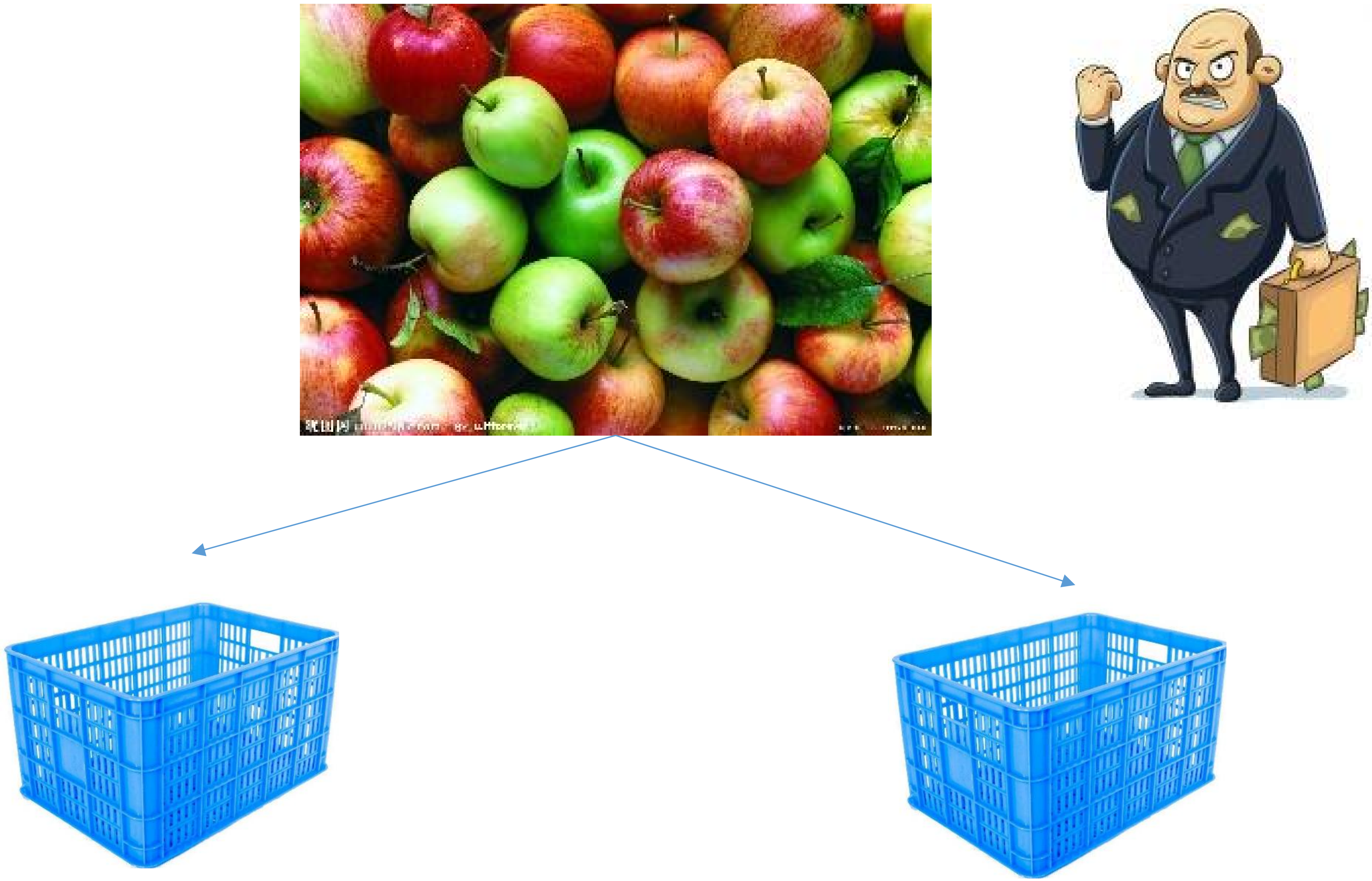
- Clustering

  - **Unsupervised machine learning** is the machine learning task of inferring a function that describes the structure of *"unlabeled" data.*



1. Different ML methods

- Clustering



1. Different ML methods

- Clustering

1. Different ML methods

- Clustering

I give you some examples.

1. Different ML methods

Supervised learning	Semi-supervised learning	Unsupervised learning

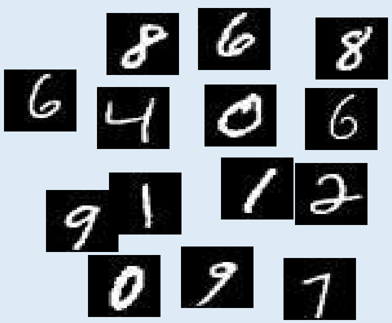
- **Semi-supervised learning** is a class of techniques that make use of unlabeled data for training.
- There are typically a small amount of labeled data with a large amount of unlabeled data

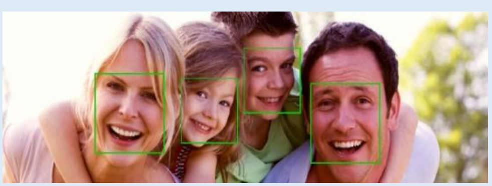
Machine Learning

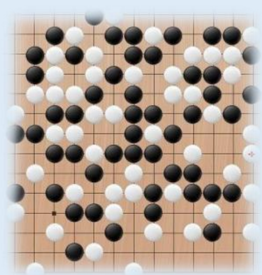
- 1. Different ML methods
- *2. Data representation*
- 3. Today's Machine learning

2. Data representation

Different types of inputs







Functions

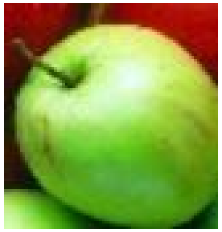
Outputs

Different tasks

2. Data representation

A_1

A_2



A_3

A_4

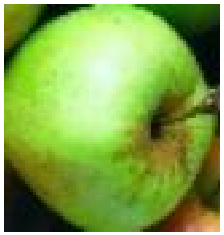
A_5

A_6

- Feature: what is feature
- Apple = [diameter, color, shape, spots, place of production]
- Dimensionality: 5

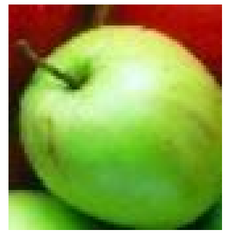
2. Data representation

Apple = [diameter, color, shape, spots, place of production]

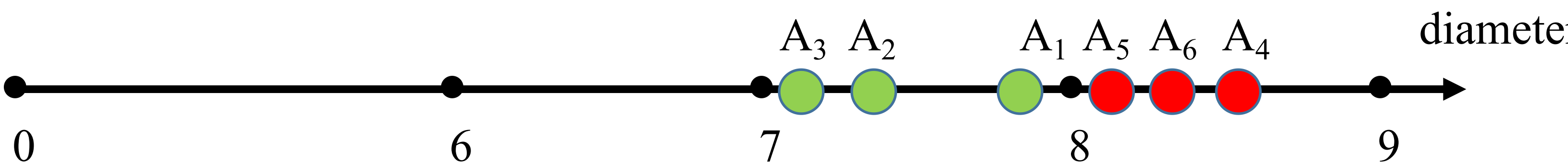


A_1= [7.8]

A_2= [7.4]



A_3= [7.1]



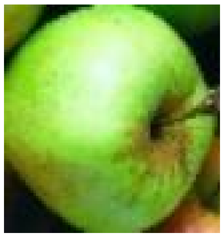
A_4= [8.5]

A_5= [8.1]

A_6= [8.3]

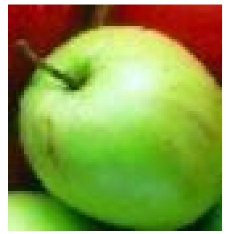
2. Data representation

Apple = [diameter, color, shape, spots, place of production]



$A_1 = \begin{bmatrix} 7.8 \\ 0.2 \end{bmatrix}$

$A_2 = \begin{bmatrix} 7.4 \\ 0.2 \end{bmatrix}$

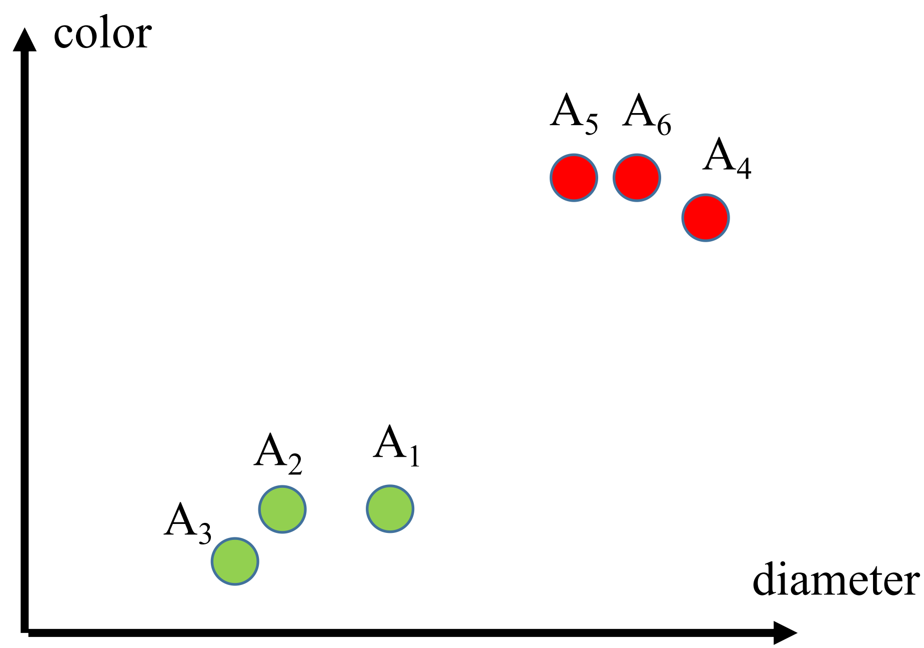


$A_3 = \begin{bmatrix} 7.1 \\ 0.1 \end{bmatrix}$

$A_4 = \begin{bmatrix} 8.5 \\ 0.7 \end{bmatrix}$

$A_5 = \begin{bmatrix} 8.1 \\ 0.8 \end{bmatrix}$

$A_6 = \begin{bmatrix} 8.3 \\ 0.8 \end{bmatrix}$

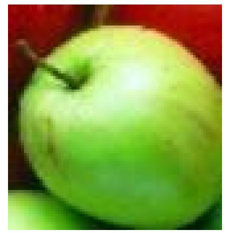


2. Data representation

Apple = [diameter, color, shape, spots, place of production]

$A_1 = \begin{bmatrix} 7.8 \\ 0.2 \\ 0.6 \end{bmatrix}$

$A_2 = \begin{bmatrix} 7.4 \\ 0.2 \\ 0.7 \end{bmatrix}$

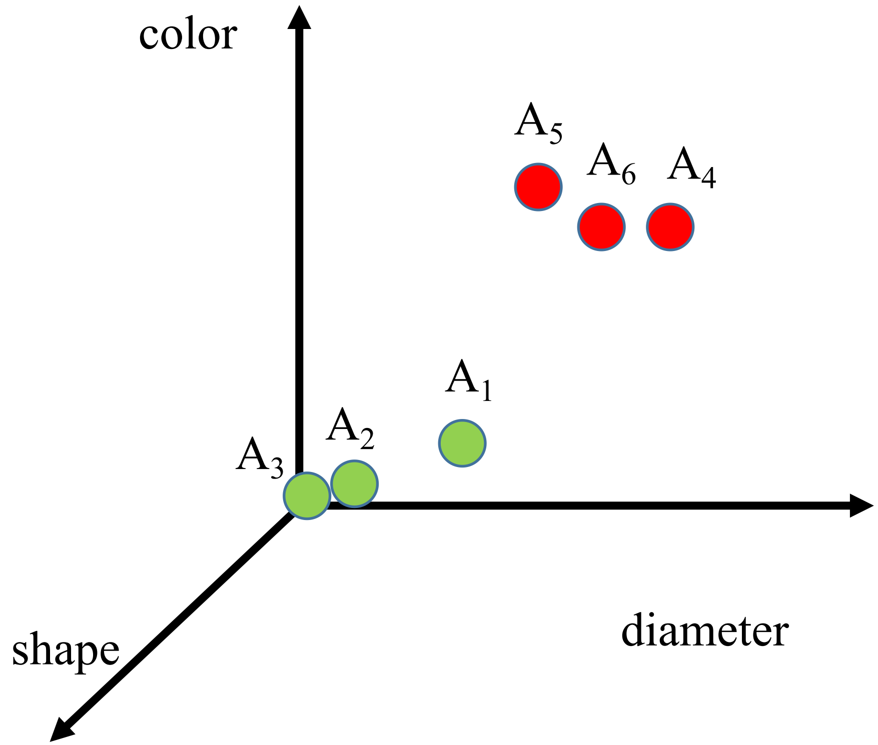


$A_3 = \begin{bmatrix} 7.1 \\ 0.1 \\ 0.6 \end{bmatrix}$

$A_4 = \begin{bmatrix} 8.5 \\ 0.7 \\ 0.7 \end{bmatrix}$

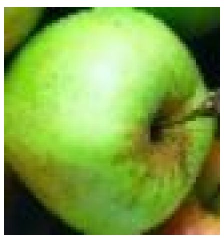
$A_5 = \begin{bmatrix} 8.1 \\ 0.8 \\ 0.7 \end{bmatrix}$

$A_6 = \begin{bmatrix} 8.3 \\ 0.8 \\ 0.8 \end{bmatrix}$

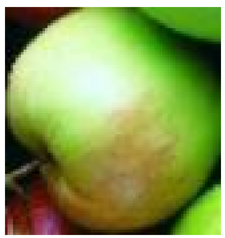


2. Data representation

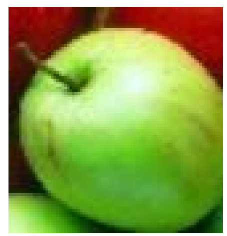
Apple = [diameter, color, shape, spots, place of production]



$$A_1 = \begin{bmatrix} 7.8 \\ 0.2 \\ 0.6 \\ 1 \\ 1 \end{bmatrix}$$



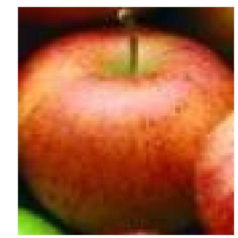
$$A_2 = \begin{bmatrix} 7.4 \\ 0.2 \\ 0.7 \\ 0 \\ 1 \end{bmatrix}$$



$$A_3 = \begin{bmatrix} 7.1 \\ 0.1 \\ 0.7 \\ 0 \\ 2 \end{bmatrix}$$

$$A_4 = \begin{bmatrix} 8.5 \\ 0.7 \\ 0.7 \\ 0 \\ 3 \end{bmatrix}$$

$$A_5 = \begin{bmatrix} 8.1 \\ 0.8 \\ 0.7 \\ 0 \\ 3 \end{bmatrix}$$



$$A_6 = \begin{bmatrix} 8.3 \\ 0.8 \\ 0.8 \\ 1 \\ 4 \end{bmatrix}$$

2. Data representation

Input: the values of the apples

$$\begin{bmatrix}7.8\\0.2\\0.6\\1\\1\end{bmatrix}\begin{bmatrix}8.1\\0.8\\0.7\\0\\3\end{bmatrix}\begin{bmatrix}7.4\\0.2\\0.7\\0\\1\end{bmatrix}\begin{bmatrix}7.1\\0.1\\0.7\\0\\2\end{bmatrix}\begin{bmatrix}8.5\\0.7\\0.7\\0\\3\end{bmatrix}\begin{bmatrix}8.3\\0.8\\0.8\\1\\4\end{bmatrix}$$

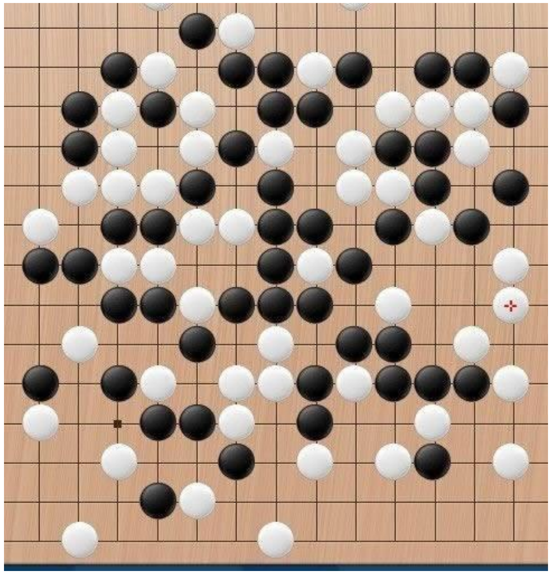
Output: the values of the apples

0 1

2. Data representation

Another example

Input: A certain state of the board



The state can be represented by a matrix.

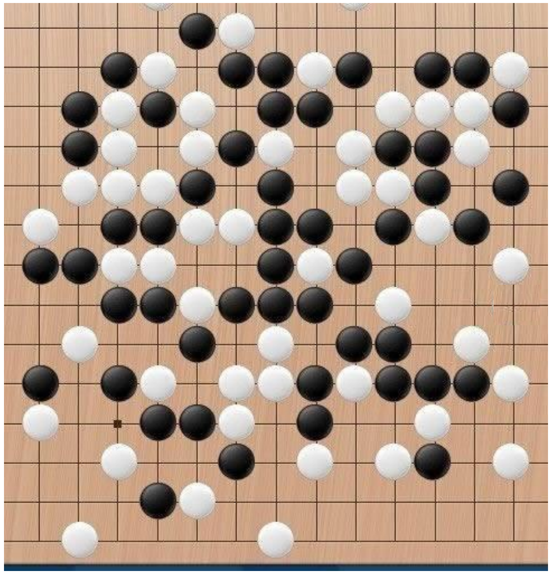
	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	-1	-1	-1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
3	-1	0	0	-1	0	0	1	0	-1	0	0	1	-1
4	0	1	0	1	-1	0	0	-1	1	1	1	0	-1
5	0	1	-1	1	0	1	-1	1	0	0	1	-1	-1
6	1	-1	1	0	-1	0	-1	1	1	0	-1	0	-1
7	-1	0	0	1	1	0	0	-1	0	1	0	-1	-1
8	0	1	1	-1	-1	0	1	0	-1	-1	-1	1	-1
9	-1	0	0	1	0	0	0	-1	1	-1	-1	1	-1
10	1	-1	-1	0	-1	1	-1	0	0	-1	1	-1	-1
11	-1	0	1	-1	1	1	0	1	0	0	0	1	-1
12	-1	-1	0	0	1	-1	0	-1	-1	1	-1	-1	-1
13	-1	1	-1	-1	0	-1	1	-1	1	0	-1	1	-1
14	-1	-1	0	1	-1	-1	-1	-1	-1	-1	-1	-1	-1
15	1	-1	-1	-1	-1	1	-1	-1	-1	-1	-1	-1	-1

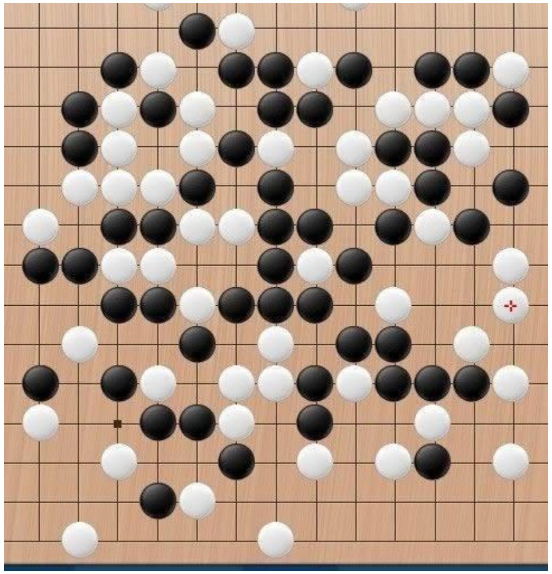
2. Data representation

- Another example

Input: A certain state of the board

Output: A new state after a move





2. Data representation

Another example

The input matrix.

	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	-1	-1	-1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
3	-1	0	0	-1	0	0	1	0	-1	0	0	1	-1
4	0	1	0	1	-1	0	0	-1	1	1	1	0	-1
5	0	1	-1	1	0	1	-1	1	0	0	1	-1	-1
6	1	-1	1	0	-1	0	-1	1	1	0	-1	0	-1
7	-1	0	0	1	1	0	0	-1	0	1	0	-1	-1
8	0	1	1	-1	-1	0	1	0	-1	-1	-1	1	-1
9	-1	0	0	1	0	0	0	-1	1	-1	-1	1	-1
10	1	-1	-1	0	-1	1	-1	0	0	-1	1	-1	-1
11	-1	0	1	-1	1	1	0	1	0	0	0	1	-1
12	-1	-1	0	0	1	-1	0	-1	-1	1	-1	-1	-1
13	-1	1	-1	-1	0	-1	1	-1	1	0	-1	1	-1
14	-1	-1	0	1	-1	-1	-1	-1	-1	-1	-1	-1	-1
15	1	-1	-1	-1	-1	1	-1	-1	-1	-1	-1	-1	-1

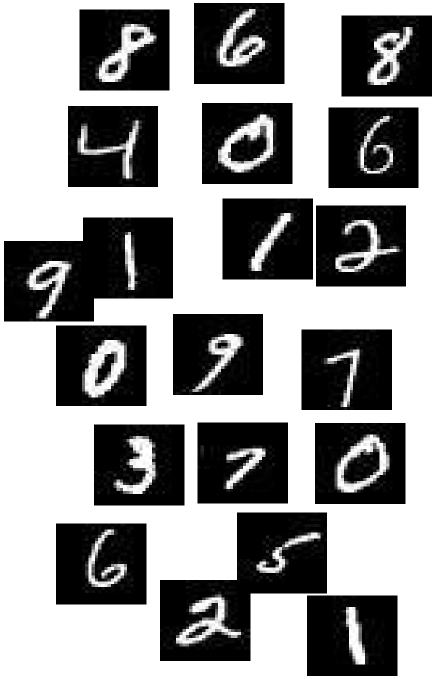
The output matrix.

	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	-1	-1	-1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
3	-1	0	0	-1	0	0	1	0	-1	0	0	1	-1
4	0	1	0	1	-1	0	0	-1	1	1	1	0	-1
5	0	1	-1	1	0	1	-1	1	0	0	1	-1	-1
6	1	-1	1	0	-1	0	-1	1	1	0	-1	0	-1
7	-1	0	0	1	1	0	0	-1	0	1	0	-1	-1
8	0	1	1	-1	-1	0	1	0	-1	-1	-1	1	-1
9	-1	0	0	1	0	0	0	-1	1	-1	-1	1	-1
10	1	-1	-1	0	-1	1	-1	0	0	-1	1	-1	-1
11	-1	0	1	-1	1	1	0	1	0	0	0	1	-1
12	-1	-1	0	0	1	-1	0	-1	-1	1	-1	-1	-1
13	-1	1	-1	-1	0	-1	1	-1	1	0	-1	1	-1
14	-1	-1	0	1	-1	-1	-1	-1	-1	-1	-1	-1	-1
15	1	-1	-1	-1	-1	1	-1	-1	-1	-1	-1	-1	-1

2. Data representation

- 3rd example

Input: Images of size 28*28

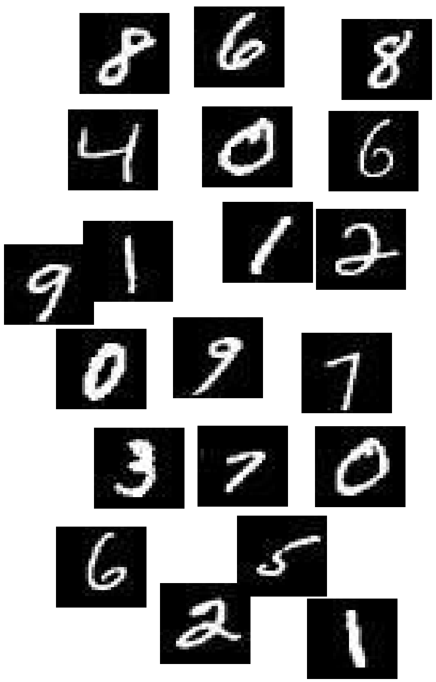


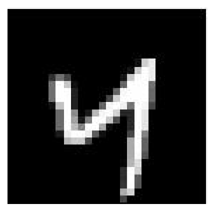
Output: Recognition results

8, 6, 8, 4, 0, 6…

2. Data representation

- 3rd example





Matrix of size 28*28

		5	6	7	8	9	10	11	12	13	14	15	16	17
6	0	0	0	0	0	0	0	0	0	0	0	0	0	
7	0	0	0	0	0	0	0	0	0	0	0	0	0	
8	0	0	0	0	0	0	0	0	0	0	0	0	0	
9	0	0	0	0	0	0	0	0	0	0	0	0	0	
10	0	0	0	49	112	69	0	0	0	0	0	0	0	
11	0	0	0	112	254	197	14	0	0	0	0	0	0	
12	0	0	0	112	254	254	32	0	0	0	0	0	99	
13	0	0	0	112	254	254	32	0	0	0	0	69	209	
14	0	0	0	17	195	254	32	0	0	0	100	245	254	
15	0	0	0	0	106	254	139	0	25	183	244	254	211	
16	0	0	0	0	106	254	162	25	128	254	254	200	78	
17	0	0	0	0	106	254	186	129	254	254	170	15	0	
18	0	0	0	0	27	236	254	254	236	91	15	0	0	
19	0	0	0	0	0	182	202	202	73	0	0	0	0	
20	0	0	0	0	0	0	0	0	0	0	0	0	0	
21	0	0	0	0	0	0	0	0	0	0	0	0	0	
22	0	0	0	0	0	0	0	0	0	0	0	0	0	
23	0	0	0	0	0	0	0	0	0	0	0	0	0	
24	0	0	0	0	0	0	0	0	0	0	0	0	0	

Gray value: 0~255

2. Data representation

- 3rd example

The input matrix.

		5	6	7	8	9	10	11	12	13	14	15	16	17
6	0	0	0	0	0	0	0	0	0	0	0	0	0	
7	0	0	0	0	0	0	0	0	0	0	0	0	0	
8	0	0	0	0	0	0	0	0	0	0	0	0	0	
9	0	0	0	0	0	0	0	0	0	0	0	0	0	
10	0	0	0	49	112	69	0	0	0	0	0	0	0	
11	0	0	0	112	254	197	14	0	0	0	0	0	0	
12	0	0	0	112	254	254	32	0	0	0	0	0	99	
13	0	0	0	112	254	254	32	0	0	0	0	69	209	
14	0	0	0	17	195	254	32	0	0	0	100	245	254	
15	0	0	0	0	106	254	139	0	25	183	244	254	211	
16	0	0	0	0	106	254	162	25	128	254	254	200	78	
17	0	0	0	0	106	254	186	129	254	254	170	15	0	
18	0	0	0	0	27	236	254	254	236	91	15	0	0	
19	0	0	0	0	0	182	202	202	73	0	0	0	0	
20	0	0	0	0	0	0	0	0	0	0	0	0	0	
21	0	0	0	0	0	0	0	0	0	0	0	0	0	
22	0	0	0	0	0	0	0	0	0	0	0	0	0	
23	0	0	0	0	0	0	0	0	0	0	0	0	0	
24	0	0	0	0	0	0	0	0	0	0	0	0	0	

The output **labels**.

8, 6, 8, 4, 0, 6…

Machine Learning

- 1. Different ML methods
- 2. Data representation
- *3. Today's Machine learning*

Machine Learning

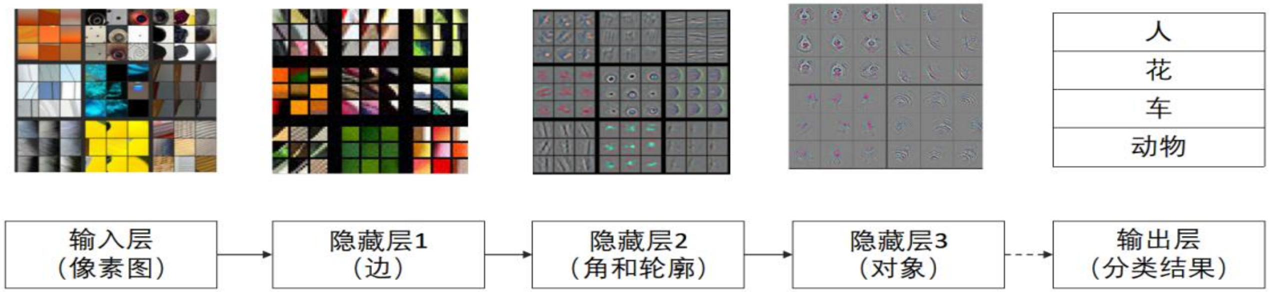
当代机器学习简介

- **深度学习：** 作为引领人工智能浪潮的核心驱动力，它以其强大的特征表征能力，在众多领域取得了突破性进展
- **强化学习：** 模拟了生物体在环境中通过试错学习最优行为的过程，为智能体自主决策与适应复杂环境提供了可能
- **图学习：** 专注于处理图结构数据，解锁了社交网络、生物信息学等领域中的隐藏价值
- **迁移学习：** 利用源域知识辅助目标域任务，极大地降低了新任务的学习成本.......

Machine Learning

■ 深度学习

深度学习（Deep Learning）是机器学习的一个子领域，模仿人脑的结构与功能，使用神经网络从数据中自动提取特征。它特别擅长处理大规模、复杂的非结构化数据，如图像、视频、声音和文本。



- **神经网络是深度学习的核心**，通常由多个层次组成，包括**输入层、隐藏层和输出层**。每一层中的节点通过权重和偏置进行连接，模拟人脑中的神经元结构。
- 深度学习中的“深度”指的是网络的**层数**。通过多个隐藏层，模型能够学习数据中的多层次抽象特征。

Machine Learning

■ 深度学习

与传统机器学习的不同

- 传统机器学习依赖人工设计的特征，而深度学习能够自动从数据中学习出适合任务的特征，无需人为干预
- 深度学习在处理非结构化数据（如图像、音频和自然语言）方面具有显著优势，依赖于大规模计算资源（如GPU）来实现

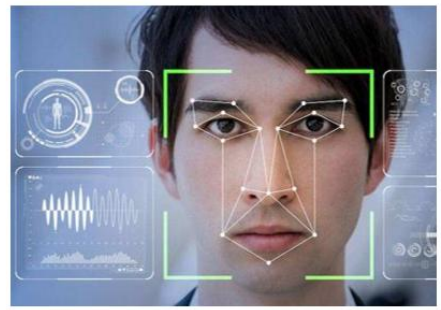
提出	陷入低谷	复兴	再度遇冷	大爆发
1943年，MP模型作为首个基于简单逻辑运算的人工神经网络模型被提出，**拉开了深度学习的序幕**	模型无法处理“异或”问题，且当时的计算能力不足	1983年，Hopfield网络和玻尔兹曼机的提出标志着神经网络的复兴，反向传播算法开始吸引越来越多的目光。 1989年，杨立昆（Yann LeCun）等人将其应用于卷积神经网络，实现手写体数字识别	统计学习理论和支持向量机等机器学习方法的崛起	• AlexNet • ResNet • Bert • GPT

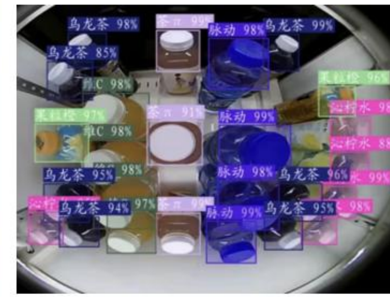
Machine Learning

■ 深度学习

- ➢ 计算机视觉


（a）自动驾驶应用


（b）人脸识别应用


（c）电子零售应用

- ➢ 语音识别、自然语言处理

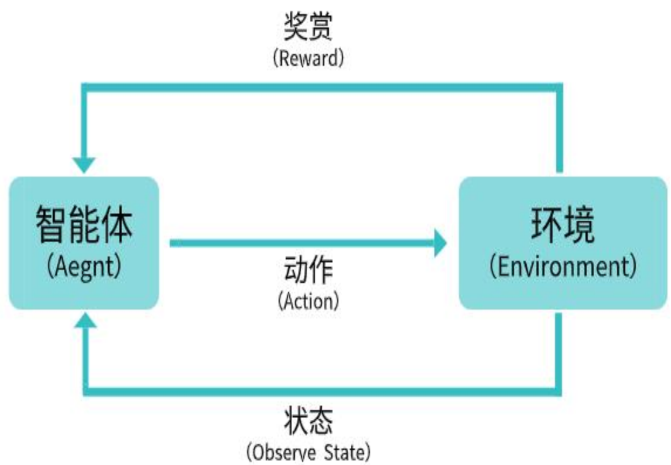
（a）百度人工智能音箱

（b）智能家居场景

Machine Learning

□ 强化学习

强化学习（Reinforcement Learning，RL）是机器学习的一个分支，专注于如何通过试错的方式让智能体在环境中学会执行任务。智能体通过与环境的不断交互，获得反馈（奖励或惩罚），以优化其决策策略



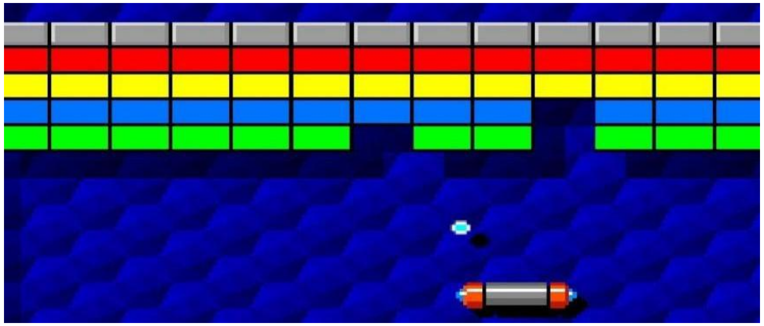
框架：

- □ **智能体（Agent）**：强化学习中的决策者，它根据环境状态采取行动。
- □ **环境（Environment）**：智能体所在的环境，给智能体提供状态和奖励。
- □ **奖励（Reward）**：智能体根据其动作获得的反馈，目标是最大化累积奖励。
- □ **策略（Policy）**：智能体根据状态选择动作的规则，是强化学习中的核心。

Machine Learning

强化学习

- **游戏领域**
  - **AlphaGo**：通过强化学习，AlphaGo学会了超越人类的围棋策略。这一成就展示了强化学习在复杂策略游戏中的强大潜力。
  - **雅达利游戏**：强化学习算法被广泛应用于经典雅达利游戏，通过不断探索和优化策略，智能体可以在游戏中表现出色。
- **机器人控制**
  - **自主导航**：通过强化学习，机器人可以学会在复杂环境中自主导航，调整路线以避免障碍物。
  - **机械臂操作**：在工业生产中，强化学习帮助机器人学会精准的抓取和操作物体，提升生产效率。



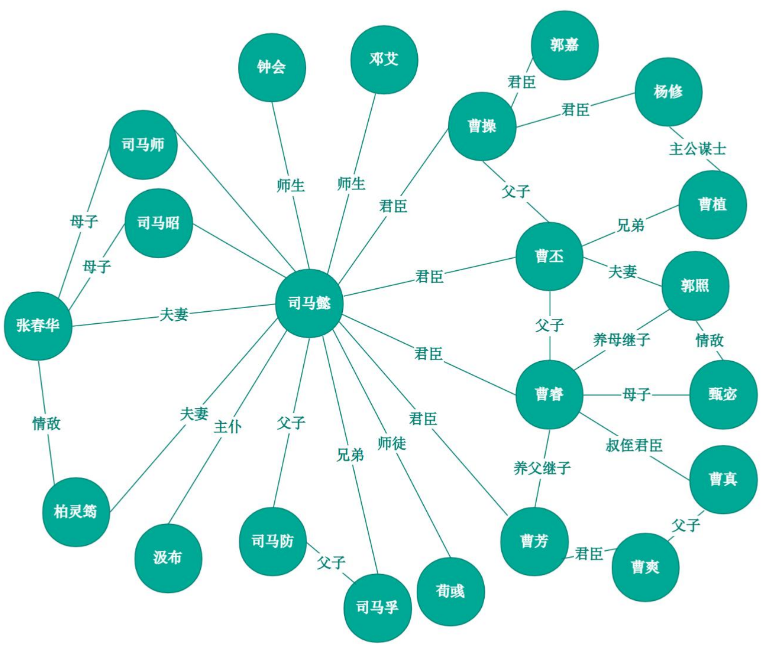
Machine Learning

□ 强化学习

- **金融领域**
  - □ **算法交易**：强化学习被用来优化交易策略，通过实时市场数据，智能体能够动态调整投资组合以获得最大收益。
  - □ **风险管理**：在金融风险预测中，强化学习帮助智能系统学习市场变化，优化决策以降低投资风险。
- **自动驾驶**：自动驾驶汽车通过强化学习技术，能够学习在复杂路况中的最佳驾驶策略，如避开障碍物、优化驾驶路径。
- **医疗领域**
  - □ **个性化治疗**：强化学习可以帮助医生根据病人的病情和历史数据制定个性化治疗方案，以实现最佳治疗效果。
  - □ **手术机器人控制**：通过强化学习，手术机器人能够在复杂的手术场景中实时学习并调整操作策略。

■ 图神经网络

图神经网络（Graph Neural Networks，GNN）是用于处理图结构数据的深度学习模型。图数据由节点和边构成，常用于描述复杂关系，如社交网络、分子结构、知识图谱等。



图的结构

- **节点（Node）**：代表图中的个体或对象，如社交网络中的用户。
- **边（Edge）**：代表个体之间的关系，如社交网络中的朋友关系。边可以有权重，表示关系的强度。

GNN通过消息传递机制，每个节点从其邻居节点接收信息，并结合自身特征更新状态。这种信息传递和状态更新的过程，使得GNN能够捕捉图中复杂的节点间关系。

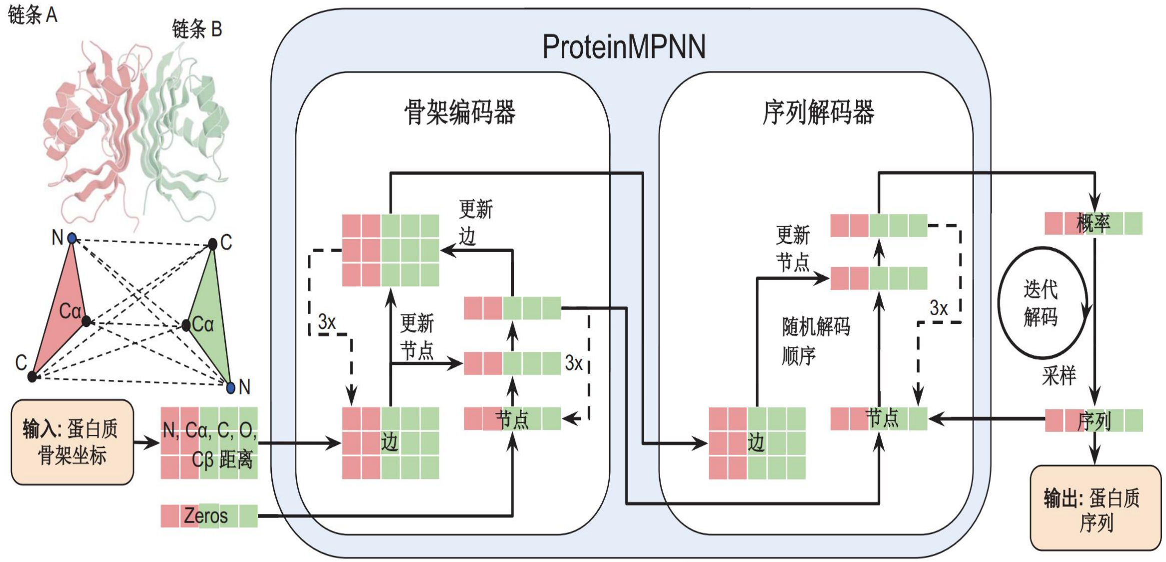
■ 图神经网络

常见的图神经网络模型

- **图卷积网络（GCN）：**通过卷积操作对图数据进行处理，适用于节点分类和图嵌入任务。
- **图注意力网络（GAT）：**引入注意力机制，根据每个邻居节点的重要性为其赋予不同的权重，适用于处理异构图数据。
- **图时序网络（GNN with time series）：**用于处理动态图或时间依赖的图数据，适用于交通预测、社交网络演化等领域。

Machine Learning

■ 图神经网络



图神经网络的优势

- **处理非欧几里得数据**：传统深度学习处理的是图像、文本等欧几里得数据，而GNN可以处理复杂的非规则图结构数据。
- **高效信息聚合**：GNN通过消息传递机制，能够聚合图中不同节点的信息，提升对节点及图整体的理解能力。
- **广泛的应用场景**：GNN在社交网络分析、生物信息学、推荐系统等领域表现出色。

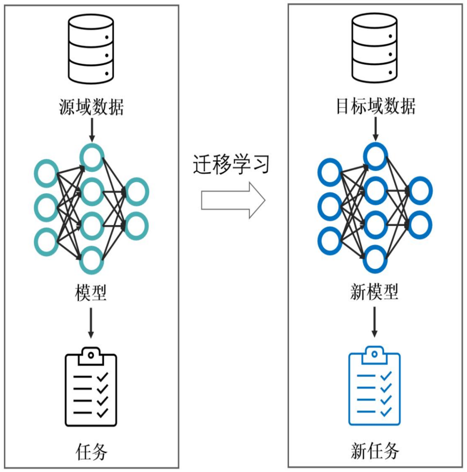
Machine Learning

■ 迁移学习

迁移学习（Transfer Learning）是一种机器学习方法，旨在将源领域（Source Domain）中学到的知识应用到目标领域（Target Domain）。当目标领域中的数据不足时，迁移学习可以有效减少对大规模数据的依赖。

核心概念:

- **源领域（Source Domain）**：模型首先在源领域进行训练，通常源领域有大量标注数据。
- **目标领域（Target Domain）**：目标领域的数据通常较少，迁移学习的目的是将源领域中学到的知识迁移到目标领域。
- **特征迁移**：在源领域中学到的通用特征可以迁移到目标领域，帮助模型在新任务中获得良好表现。



Machine Learning

■ 迁移学习

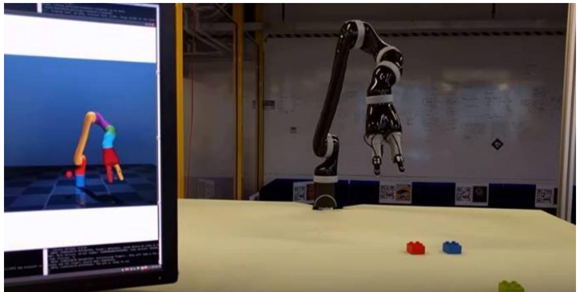
- **计算机视觉**
  - **图像分类与物体检测**：通过在大型数据集（如ImageNet）上预训练模型，迁移学习将这些模型迁移到医学影像、自动驾驶中的图像分类任务，极大提升了模型性能。
  - **图像分割**：迁移学习帮助在医学影像中进行精细的器官分割任务，减少了对标注数据的需求。
- **自然语言处理**
  - **文本生成与翻译**：BERT、GPT等语言模型在大规模文本数据上进行预训练，迁移学习使得这些模型在小规模数据集上的文本生成、翻译任务中表现优异。
  - **情感分析**：通过在通用文本数据上预训练的模型，迁移学习可以轻松应用于社交媒体、客户反馈中的情感分析。
- **语音识别与生成**
  - **口音适配**
  - **语音生成**

Machine Learning

■ 迁移学习

- 自动驾驶
  - **道路场景识别**：迁移学习帮助自动驾驶系统从不同城市的道路数据中学习，共享环境识别知识，提升自动驾驶车辆在全球范围内的表现。
  - **多场景适配**：通过在多种驾驶场景上进行预训练，自动驾驶车辆可以适应不同天气、光照和交通环境。

- 医疗健康
  - **疾病检测与诊断**：通过在公开医学数据集上预训练的模型，迁移学习应用于特定医院的数据集，提升了疾病检测的准确性。
  - **药物研发**：迁移学习在药物化学反应预测中发挥了重要作用，加速了药物发现过程。



Machine Learning

- 1. Different ML methods
- 2. Data representation
- 3. Today’s Machine learning