Foundation of Artificial Intelligence

人工智能基础

李翔宇

软件学院

Email: lixiangyu@bjtu.edu.cn

Outline

V. Machine Learning（10h）

- ◆ Theory（5h）
  - (1) Supervised learning:
    - KNN，Support Vector Machine (SVM)
  - (1) Unsupervised learning: K-means
  - (3) Reinforcement Learning (RL)
- ◆ Quiz---1h
- ◆ Practice（4h）
  - ➢ Iris classification（SVM，multiple classes）
  - ➢ Boston house price prediction（Linear Regression）

Machine Learning

5.1 Perspectives about Machine Learning

5.1.1 What is Machine Learning

5.1.2 History of Machine Learning

5.1.3 Why Different Perspectives

5.1.4 Three Perspectives on Machine Learning

5.1.5 Applications and Terminologies

What is Machine Learning 什么是机器学习

Machine learning is a branch of artificial intelligence, is the key to intelligence. Its goal is to construct the systems that can **learn** from data and make **predictions** on data.

Machine learning is the study of **algorithms** and mathematical **models** that computer systems use to progressively improve their performance on a specific task.

Machine learning algorithms build a mathematical model of **sample data**, known as "**training data**", in order to make **predictions** or **decisions** without being explicitly programmed to perform the task.

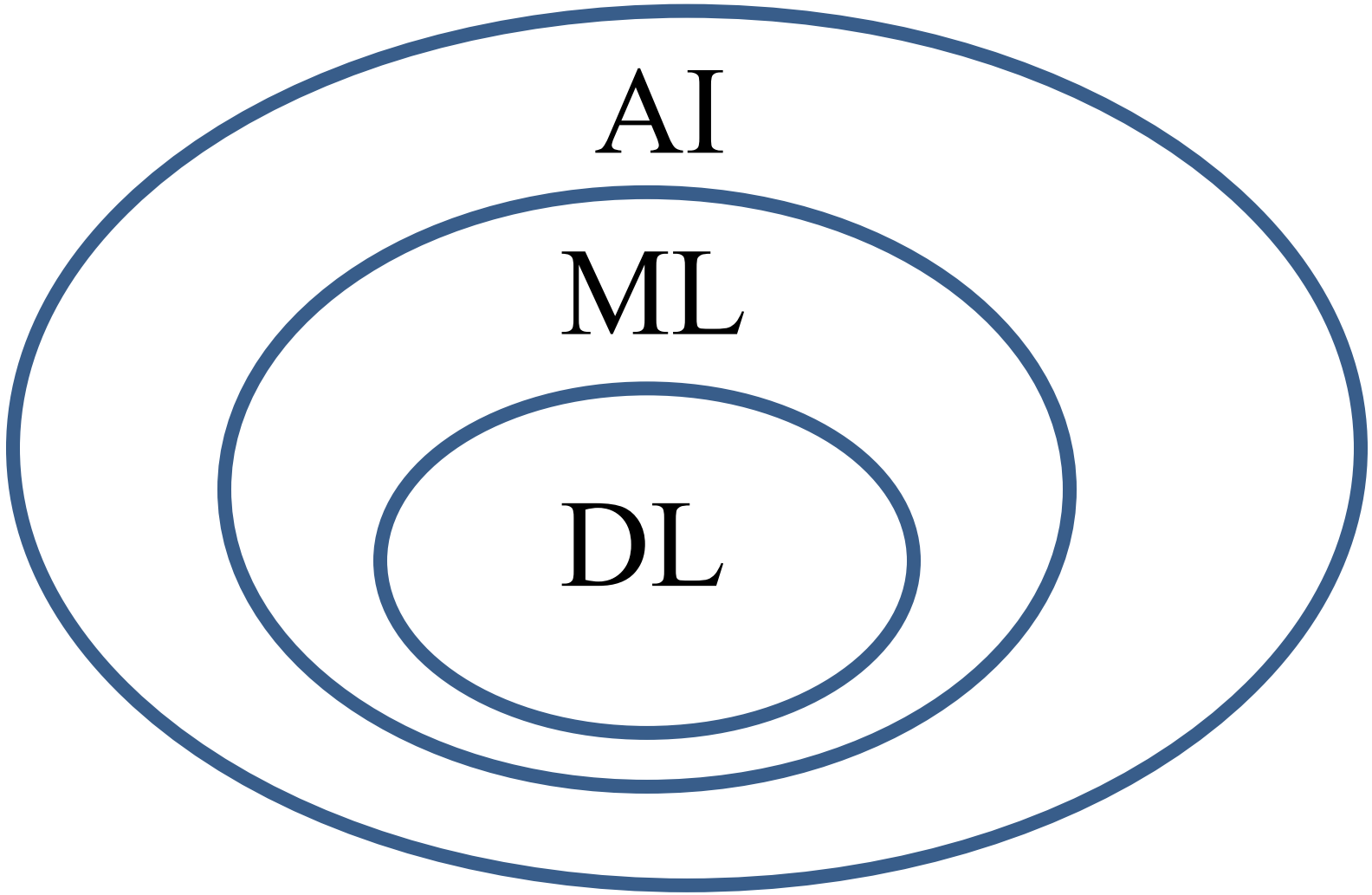
Wikipedia

What is Machine Learning ?

- Machine learning is a multi-domain interdisciplinary subject, involving probability theory, statistics, approximation theory, convex analysis, algorithm complexity theory and so on.
- It focuses on how computers simulate or implement human learning behaviors in order to acquire new knowledge or skills and reorganize existing knowledge structures to improve their performance.

Baidu encyclopedia

Relationship among AI、Machine Learning (ML)、and Deep Learning (DL)



Relations to Other Disciplines
与其他学科的关系

Statistical Learning 统计学习

- a machine learning framework drawing from statistics.
  取自于统计学的机器学习框架。

Pattern Recognition 模式识别

- the recognition of patterns in data. (≈ machine learning + data patterns)
  识别数据中的模式。(≈ 机器学习 + 数据模式)

Data Mining 数据挖掘

- the discovery of unknown properties in data. (≈ machine learning + database)
  发现数据中的未知特性。(≈ 机器学习 + 数据库)

Computer Vision 计算机视觉

- to extract information from images. (≈ machine learning + image processing)
  从图像中提取信息。(≈ 机器学习 + 图像处理)

Relationship between ML and other Fields

- **PR** ≈ ML. PR from industry, ML from CS.
- **DM**=ML+DB. Most DM algorithm is the optimization of ML algorithms in DB.
- **SL** is a part of ML. SL focus on mathematical theory research， ML focus on practice.
- **CV**=ML+IP. Image Processing provides input for the ML model, ML provide learning algorithms and output the vision result.
- **SR**=Speech Processing+ML. SR provides input for the ML model. SR and NLP are usually combined to be used.
- **NLP**=Text Processing+ML.

模式识别

Pattern Recognition

计算机视觉

Computer Vision

Data Mining

数据挖掘

Speech Recognition

语音识别

Statistic Learning

统计学习

机器学习

Machine Learning

Natural Language Processing (NLP)

自然语言处理

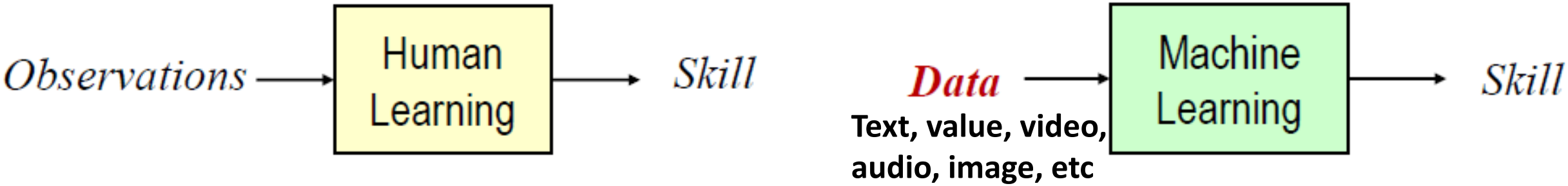
Artificial Intelligence vs. Machine Learning

◆Human Learning

Human being acquire skill with **experience** accumulated from **observations**.

◆Machine Learning

Machine acquire skill with **experience** accumulated / computed from **data**.

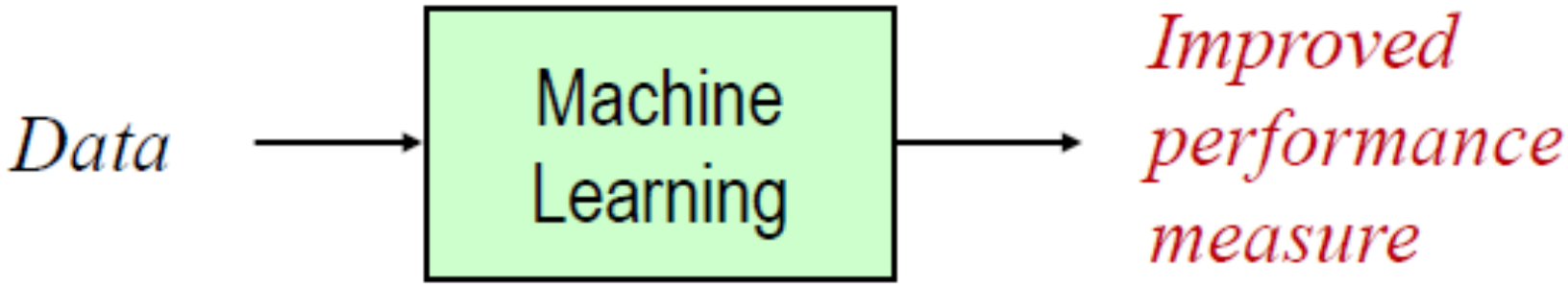


Machines mimic human's study activities.

What is Skill in Machine Learning

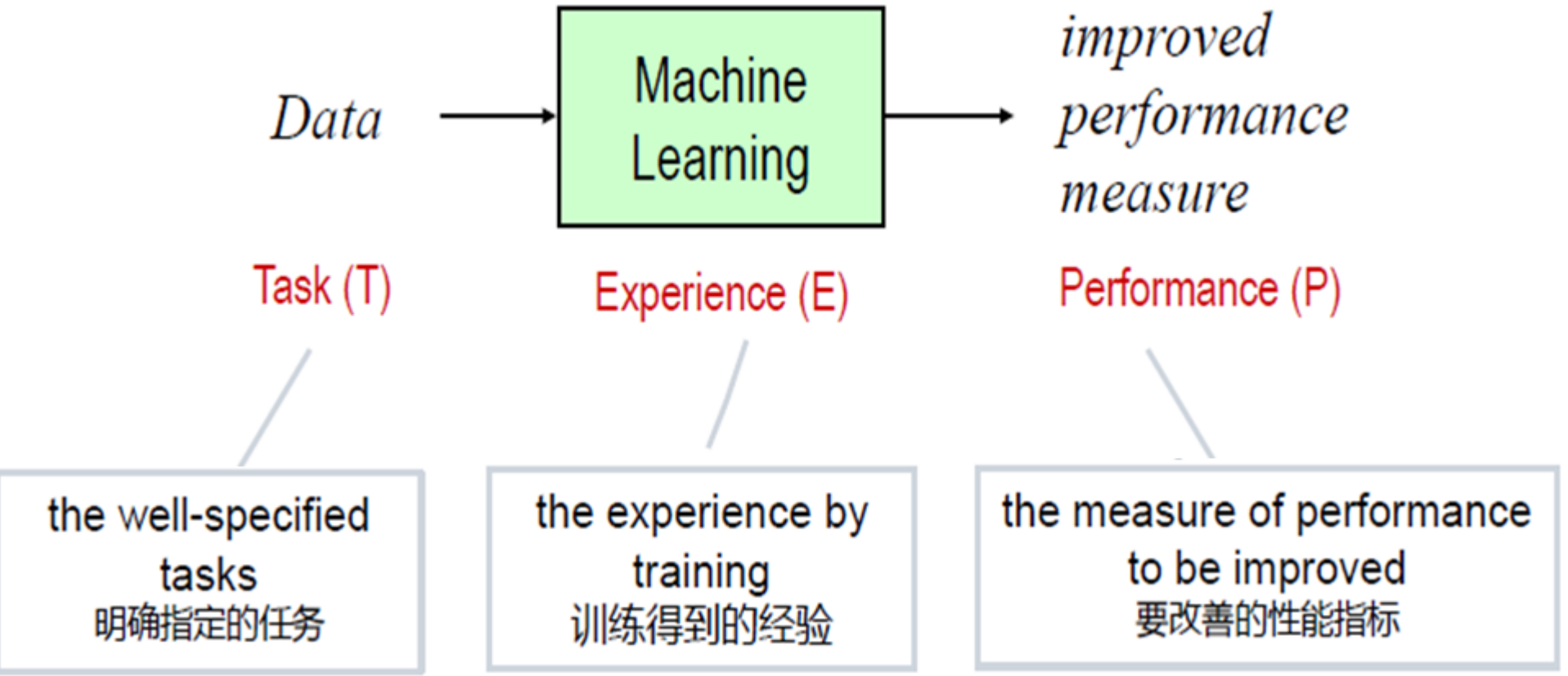
- **Skill**
  Skill is used to improve some performance measure. (e.g. prediction accuracy)

- **Why Use Machine Learning to acquire skill**
  Machine learning can improve some **performance measure** with **experience** computed from data.



Three Key Elements in the Formal Definition

To have a well-defined learning problem, we must identity those three features:



What task is to be completed?

What experience can be obtained from the data?

How to evaluate the performance?

(1) Determine the task and collect training data.
(2) Obtain the experience from the data
(3) Give the result according to the experience and evaluate the performance of result.

*Example*1: A handwriting recognition problem

Data → Machine Learning → *improved performance measure*

Task (T)

Experience (E)

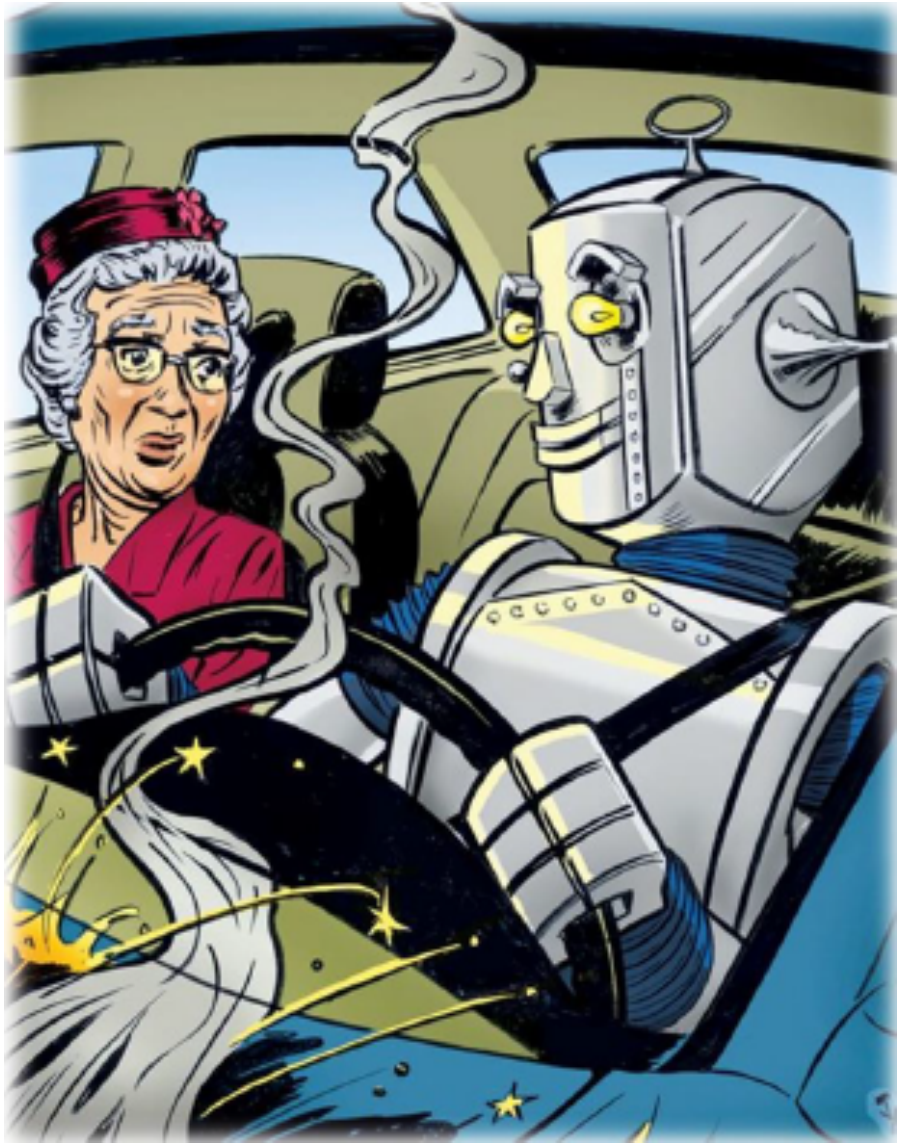
Performance (P)

recognizing and classifying handwritten words within images
对图像内的手写文字进行识别和分类

a database of handwritten words with given classifications
具有给定分类的手写文字数据库

percent of words correctly classified
正确分类文字的百分比

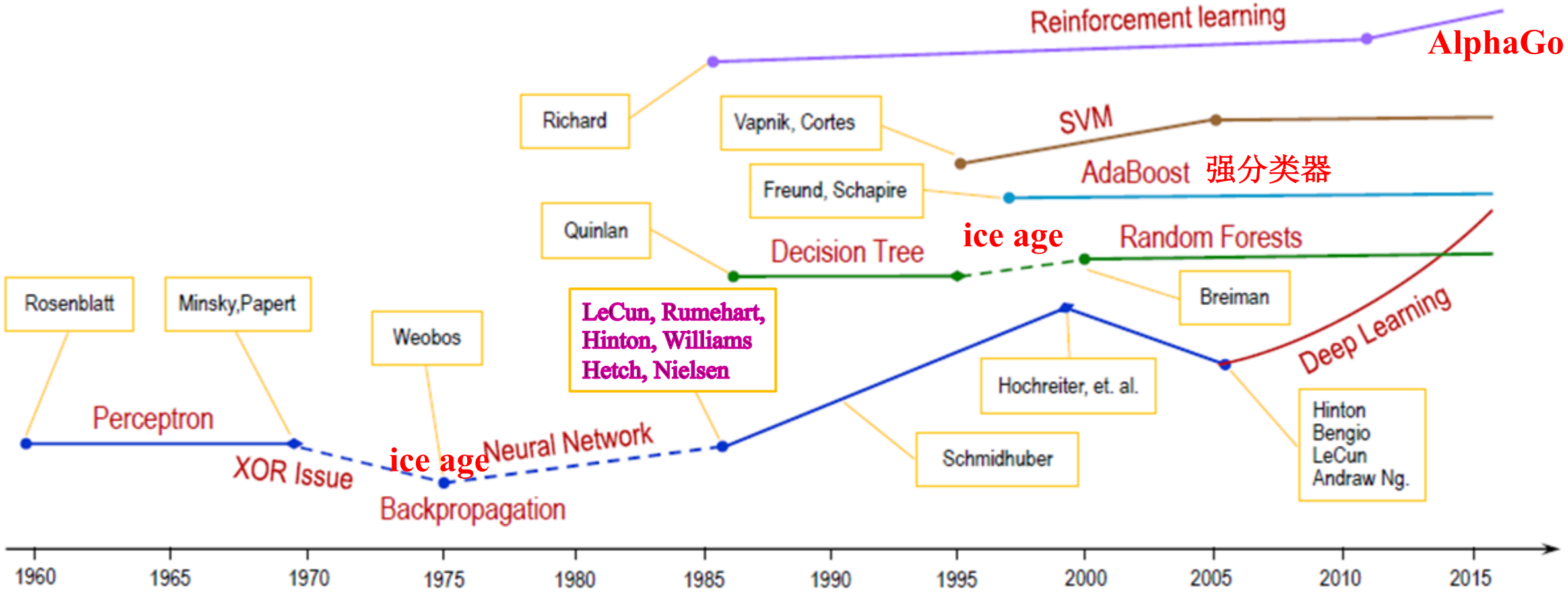
*Example*2: A robot driving problem



- **Task (T)**:
  driving on public four-lane highways using vision sensors
- **Experience (E)**:
  a sequence of images and steering commands recorded while observing a human driver
- **Performance (P)**:
  average distance traveled before an error (as judged by human overseer).

History of Machine Learning

Timeline of Machine Learning



Perceptron an algorithm of binary classifiers

XOR Issue perceptrons can't process XOR circuit

Backpropagation a common method of training neural networks

Neural Network learning algorithms inspired by biological neural networks

Reinforcement Learning concerned with how an agent take actions in an environment

Decision Tree a decision support tool that uses a tree-like graph

SVM support vector machine

AdaBoost adaptive boosting algorithm

Random Forests a multitude of decision trees

Deep Learning deep networks, hierarchical learning

Three Parties of Machine Learning

Connectionism 联结主义

also called Bionicsism, Physiologism.
亦称仿生主义、生理学派

e.g., Perceptron, Artificial Neural Network, Deep learning.

Symbolicism 符号主义

also called Logicism, Psychologism, Computerism.
亦称逻辑主义、心理学派、计算机主义

e.g., Association Rules, Decision tree, Random Forests

Behaviorism 行为主义

also called Actionism, Evolutionism, Cyberneticsism
亦称行动主义、进化主义、控制论学派

e.g., Reinforcement learning

Why Different Perspectives

Difficulty in Understanding Machine Learning

- How many learning algorithms
  - There are many algorithms for machine learning.
  - Literally thousands available, and hundreds more published each year.

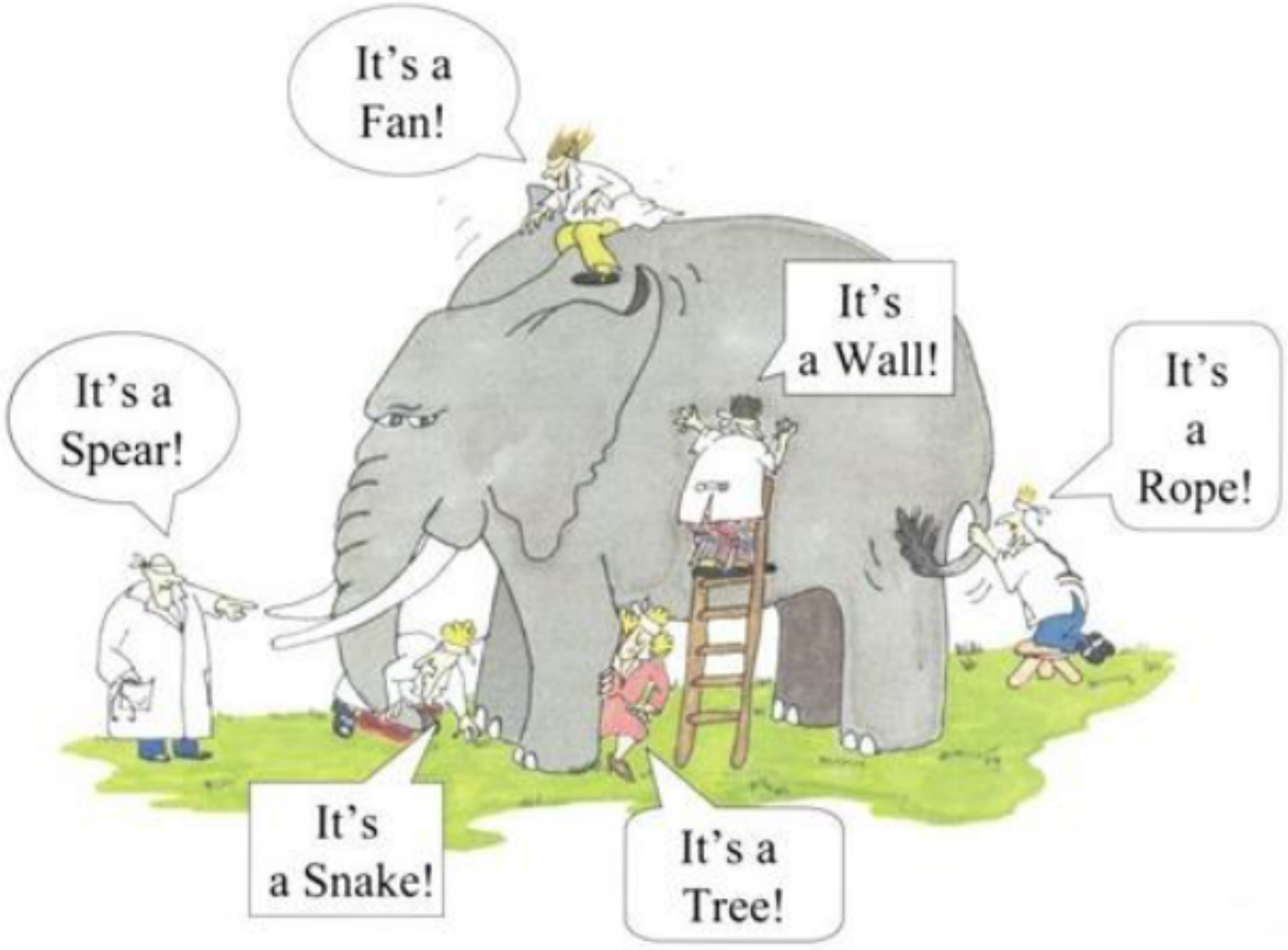
- Which algorithm should we choose
  - Suppose we have an application that machine learning might be good for, so we need an appropriate algorithm for learning from data.

  - The problem we faced is how to choose one of machine learning algorithms.

Difficulty in Understanding Machine Learning

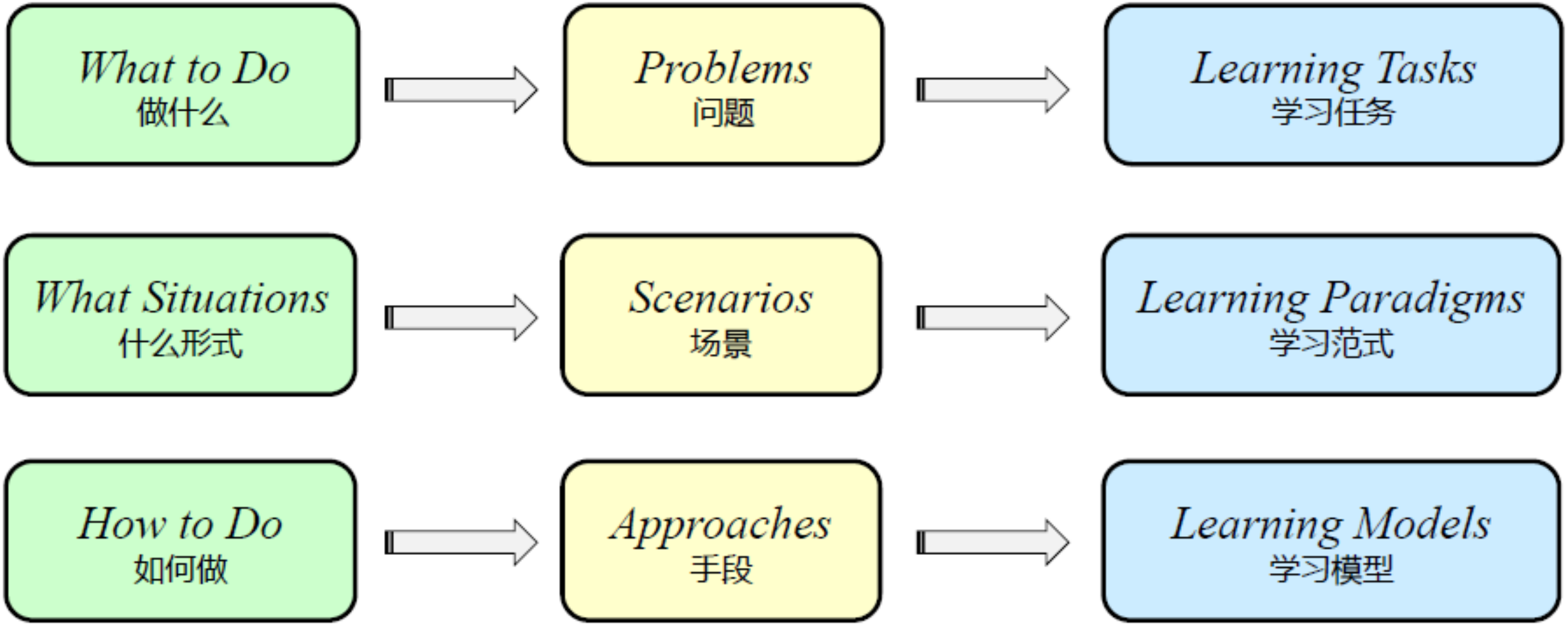
- What is the difficulty
  - Without a category of machine learning, how to determine which algorithm could be used?
  - The categorization relates our perspective on choosing machine leaning.
- Is one perspective enough To outlook on most of machine learning algorithms, one perspective is so hard.
  - We should look from multiple perspectives to have a full view of machine learning.

The Three Perspectives



Maybe "Blind Men and an Elephant"

Three Perspectives of Machine Learning
机器学习的三个视角



Three Perspectives of Machine Learning

Perspectives	Description 描述
Learning Tasks	Denoting the general **problems** that can be solved by machine learning.
Learning Paradigms	Denoting the **typical scenarios** that are happened in machine learning.
Learning Models	Denoting the **approaches** that can fulfil a learning task.

Three Perspectives on Machine Learning

What are Learning Tasks

- **What are Learning Tasks**
  The learning tasks are used to denote the general problems that can be solved by learning with desired output.

- **Why do we need to Study Learning Tasks**
  Various types of problems arising in applications:

  - computer vision,
  - pattern recognition,
  - natural language processing,
  - ……..

Typical Tasks in Machine Learning

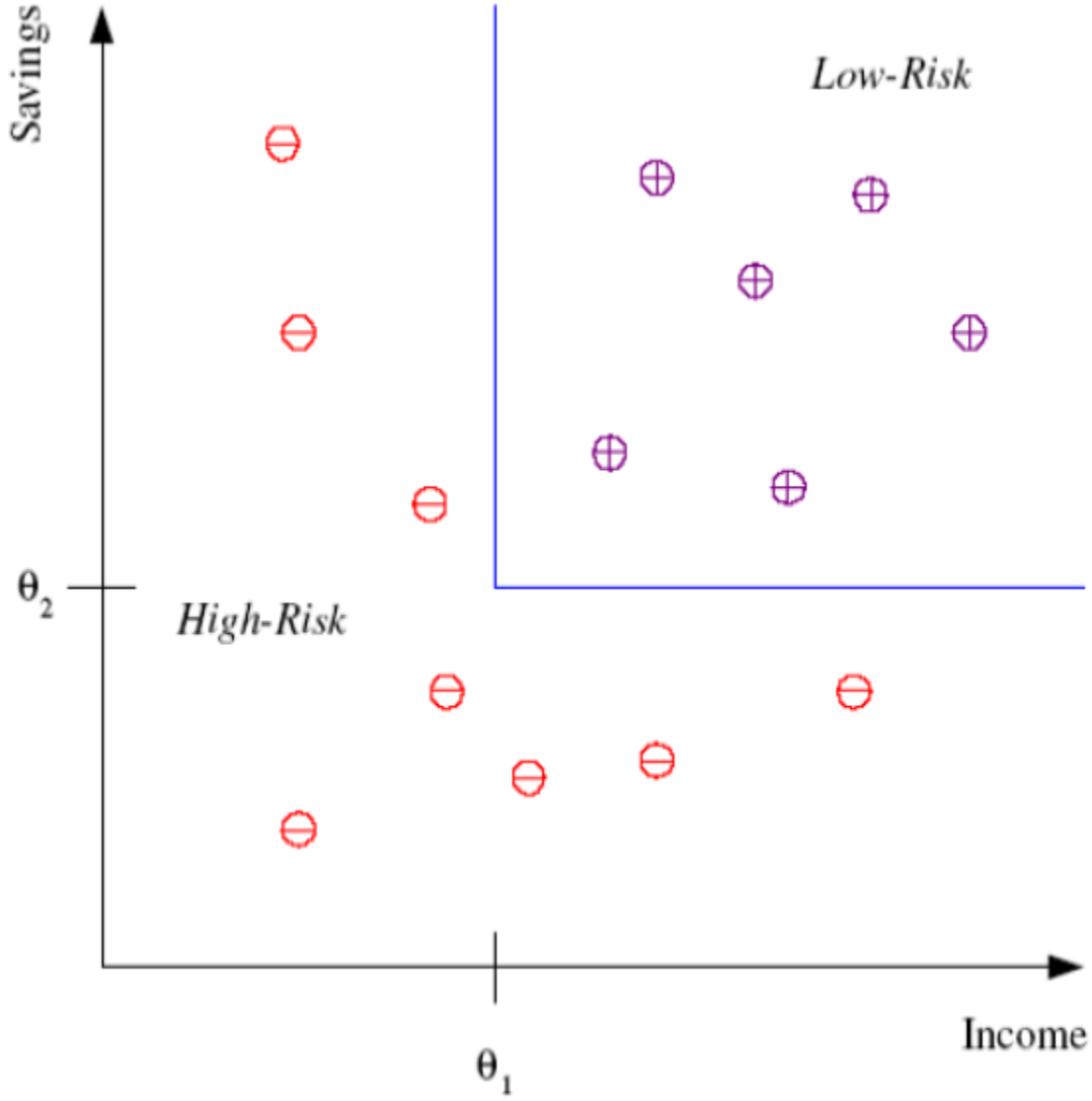
Tasks 任务	Brief Statements 简短描述	Typical algorithm 典型算法
Classification 分类	Inputs are divided into two or more known classes. 将输入划分成两个或多个类别。	SVM 支撑向量机
Regression 回归	Outputs are continuous values rather than discrete ones. 输出是连续值而不是离散的。	Bayesian linear regression 贝叶斯线性回归
Clustering 聚类	Inputs are divided into groups which are not known beforehand. 输入被划分为若干个事先未知的组。	*k*-means *k*-均值
Ranking 排名	Data transformation in which values are replaced by their rank. 用它们的排名来代替值的数据转换。	PageRank 网页排名
Density estimation 密度估计	Find the distribution of inputs in some space. 寻找某个空间中输入的分布。	Boosting Density Estimation 增强式密度估计
Dimensionality reduction 降维	Simplify inputs by mapping them into a lower dimensional space. 通过将输入映射到低维空间来将其简化。	Isomap 等距特征映射
Optimization 优化	Find the best solution from all feasible solutions 从所有可能的解中寻找最优解。	Q-learning Q-学习

Case study: Credit scoring 信用评分

- **Two classes**: Low-risk and high-risk customers.

- A customer will be classified into one class according to his information.
- After training with past data, a classification rule learned may be:

IF *income* > θ_1 AND *savings* > θ_2
THEN *low-risk*
ELSE *high-risk*



Learning Paradigms

What are Learning Paradigms

The Learning Paradigms are used to denote the typical scenarios that are happened in machine learning.

How to Distinguish Learning Paradigms

by the *scenarios* or *styles* in machine learning about

- how it learns from data,
- how it interactives with environment.

Learning Paradigms in Machine Learning

Paradigms 范式	Brief Statements 简短描述	Typical Algorithm 典型算法
Supervised 有监督	The algorithm is trained by a set of labeled data, and makes predictions for all unseen points. 算法采用一组标注数据进行训练，再对所有的未知点做出预测。	Support vector machines 支撑向量机
Unsupervised 无监督	The algorithm exclusively receives unlabeled data, and makes predictions for all unseen points. 算法仅接收未标注的数据，再对所有的未知点做出预测。	*k*-means *k*-均值
Reinforcement 强化	The algorithm interacts with environment, and receives an reward for each action. 算法与外部环境交互，每个动作得到一个回报。	Q-learning

Learning Models

What are Learning Models

The **learning models** are used to denote the approaches that can fulfil a learning task.

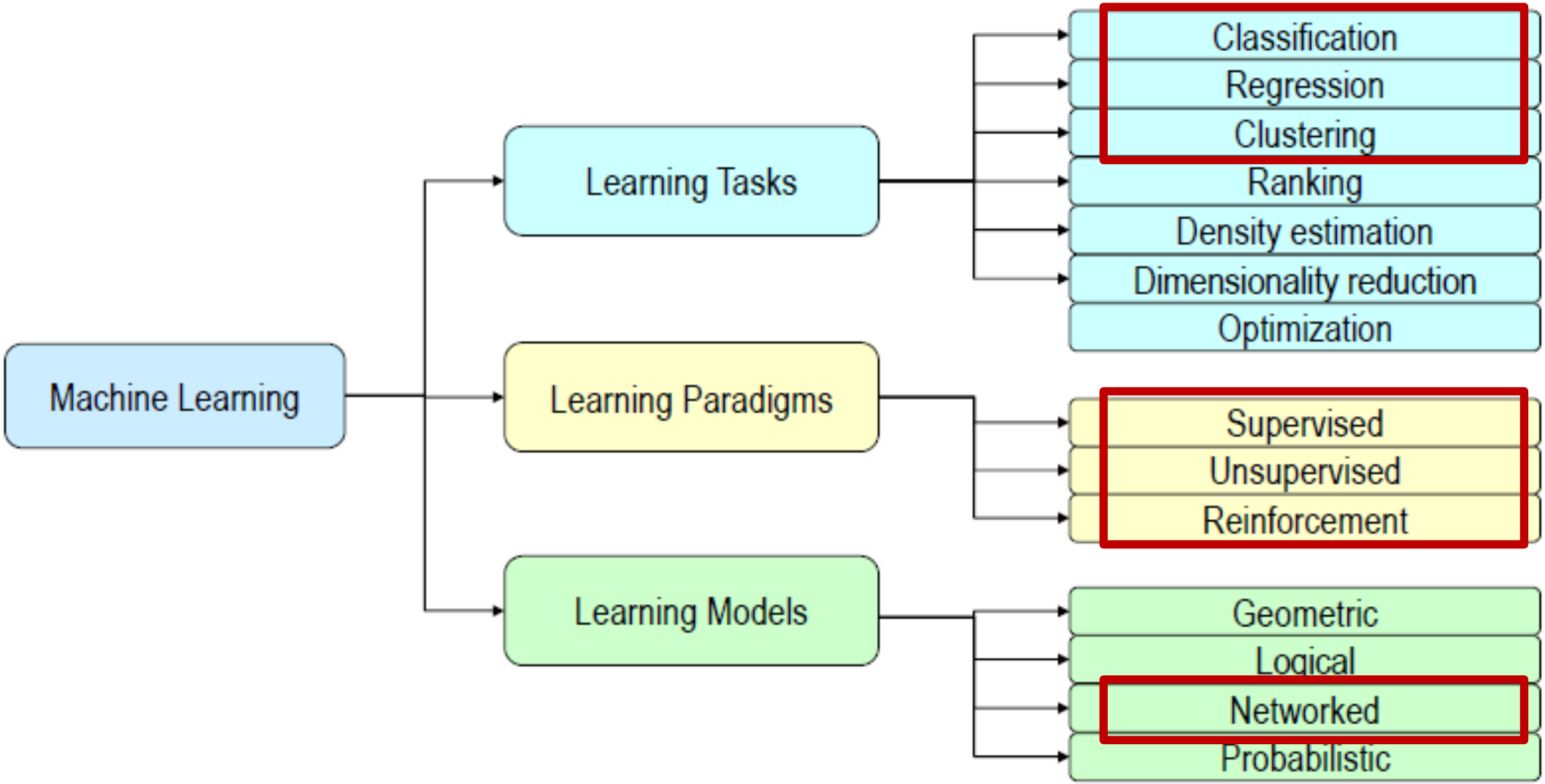
Why Study Learning Models

The result of machine learning is heavily dependent on the choice of an approach for solving the learning task.

Typical Models for Machine Learning

Models 模型	Brief Statements 简短描述	Sub-models 子模型	Typical Algorithm 典型算法
Geometric 几何	Use geometric models such as line, plane, distance or manifold to construct learning algorithms. 采用线、面、距离或流行等几何图形模型来构建学习算法。	Line 线	Linear Regression 线性回归
		Plane 面	SVM 支撑向量机
		Distance 距离	k-NN k-近邻
		Manifold 流行	Isomap 等距映射
Logical 逻辑	Use logical models to construct learning algorithms. 采用逻辑模型来构建学习算法。	Logic 逻辑	Inductive Logic Program. 归纳逻辑编程
		Rule 规则	Association Rule 相关规则
Networked 网络	Use networked models to construct learning algorithms. 采用网络模式构建机器学习算法。	Shallow 浅层	Perceptron 感知机
		Deep 深层	CNN 卷积神经网络
Probabilistic 概率	Use probabilistic models to denote the conditional dependence between random variables. 采用概率模式来表示随机变量之间的条件相关性。	Bayes 贝叶斯	Bayesian Network 贝叶斯网络
		Generative 生成	Probabilistic Program. 概率规划
		Statistic 统计	Linear Regression 线性回归

The Three Perspectives



Applications and Terminologies

Application Fields of Machine Learning 机器学习的应用领域

- Machine Perception 机器感知
- Computer Vision 计算机视觉
- Video Analysis 视频分析
- Pattern Recognition 模式识别
  - Face/Speech/Fingerprint Recognition 人脸/语音/指纹识别
  - Optical Character Recognition (OCR) 光学字符识别 (OCR)
  - Handwriting Recognition 手写体识别
- Game Playing 玩游戏
- Natural Language Processing 自然语言处理
- Information Retrieval 信息检索

Application Fields of Machine Learning

- Text or Document Classification (e.g. Spam Email Detection) — 文本与文档分类（例如垃圾邮件检测）
- Recommender Systems — 推荐系统
- Ad Placement — 广告配置
- Credit Scoring — 信用评分
- Fraud Detection — 欺诈检测
- Stock Trading — 股票交易
- Drug Design — 新药设计
- Medical Diagnosis — 医学诊断
- Robotics — 机器人学

Some Terminologies in Machine Learning

- **Samples**
  Items or instances of data used for learning or evaluation.

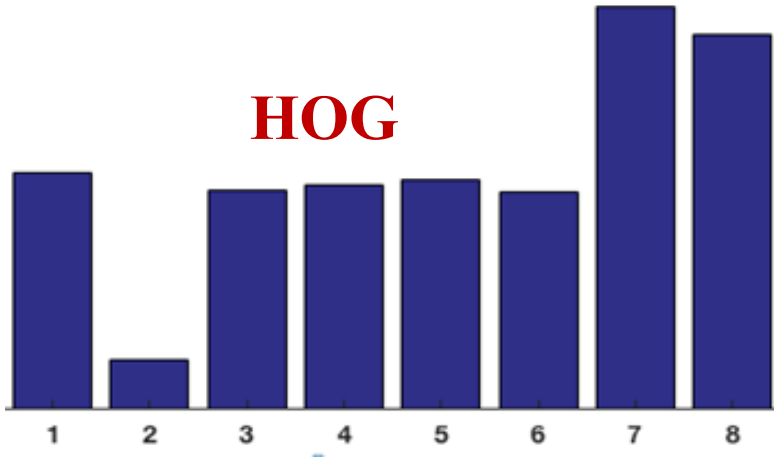
- **Features**
  The set of attributes, often represented as a vector associated to a sample:

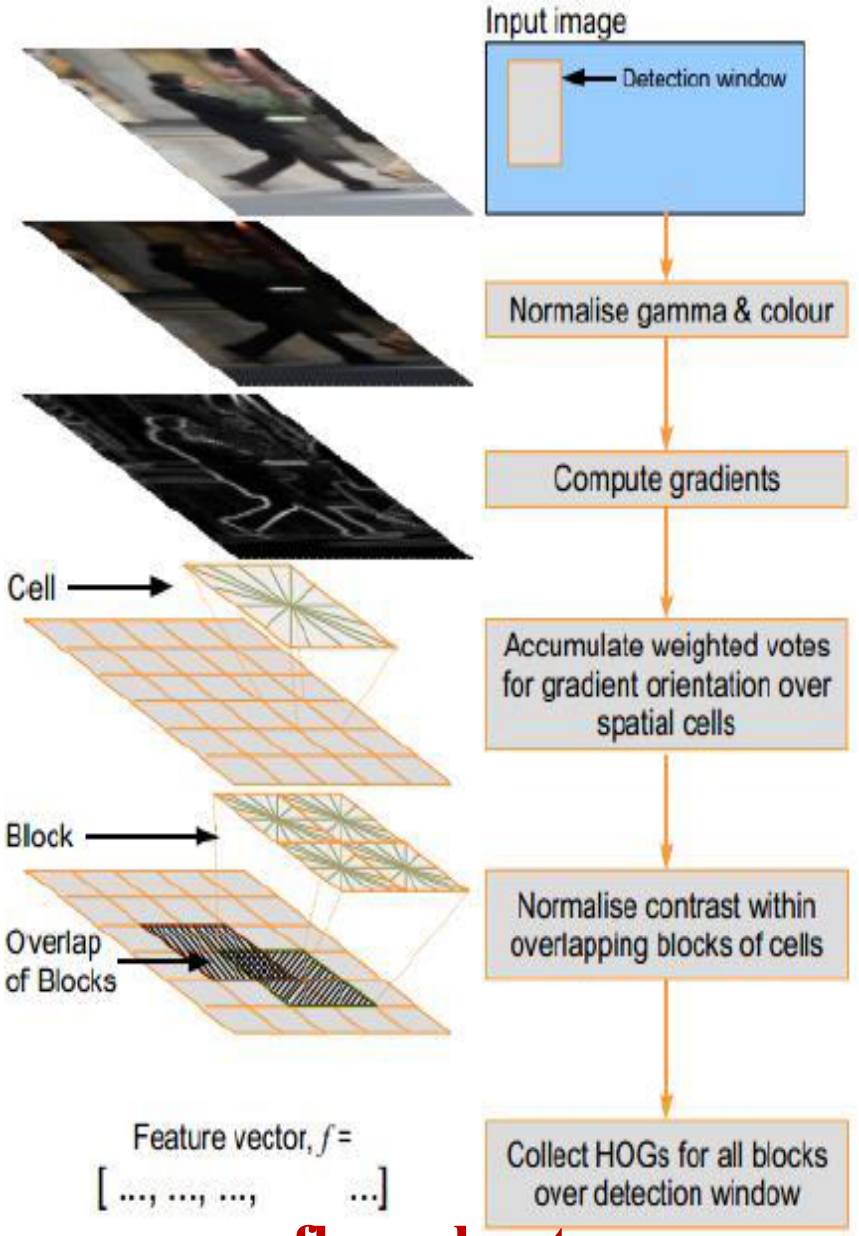
  - Handcrafted features:
    e.g., SIFT, HOG, SURF, LBP, GLOH, LESH, CENTRIST.
  - Learned features:
    e.g., visual features obtained by convolutional neural network.

Some Terminologies in Machine Learning

◆ Handcrafted Features /Designed Features

- ➢ It is the feature that researchers design.
- ➢ e.g. HOG (Histogram of Oriented Gradients)
- ➢ By distribution of intensity gradients or edge directions.
- ➢ 64×128 detection window.



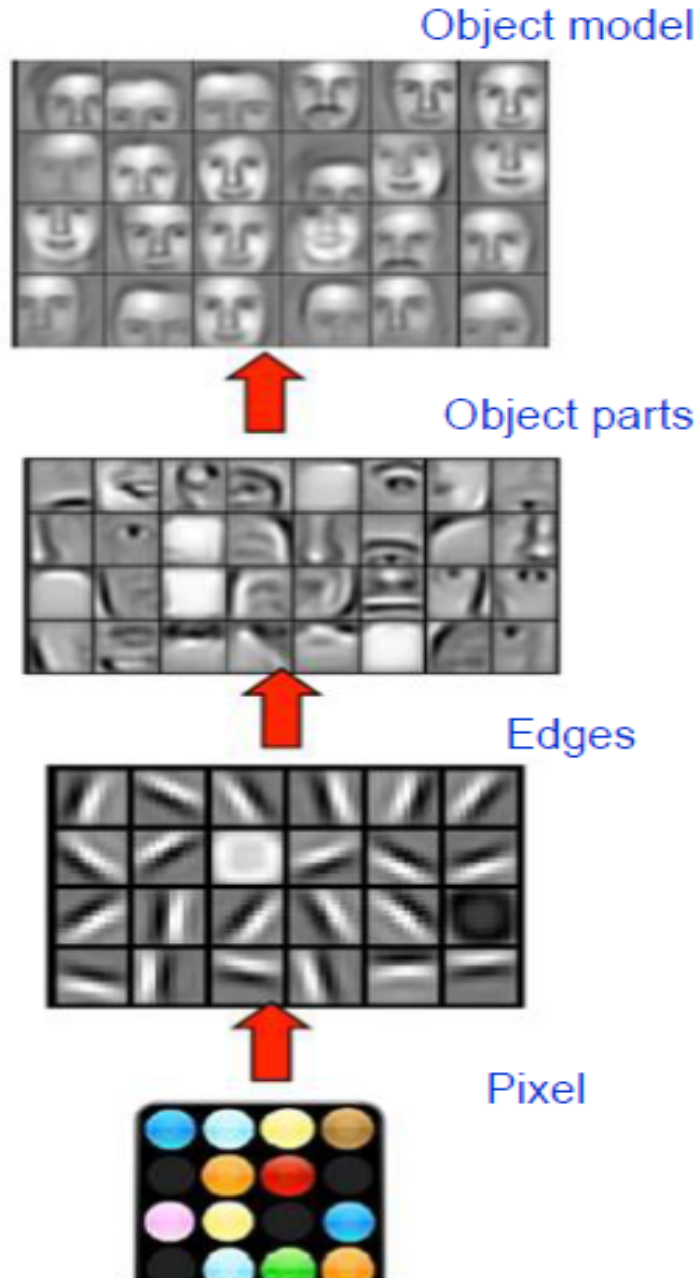


flow chart

Some Terminologies in Machine Learning

◆ Learned Features

- Humans can learn to see efficiently. Because brains are deep, with many layers of processing.
- Some algorithms for such deep architectures, can produce visual features automatically from raw data.
- Feature learning also be called representation learning.
- Understanding deep learning will enable us to build more intelligent machines for visual recognition.



Some Terminologies in Machine Learning

- **Labels**
  - Values or categories assigned to samples.
  - In classification problems, samples are assigned specific categories.
  - In regression problems, items are assigned real-valued labels.
- **Training sample**
  - Samples used for training learning algorithm.
  - In spam problem, the training sample consist of a set of email samples along with their associated labels.

Some Terminologies in Machine Learning

- **Validation sample**
  - **Validation samples** are the labeled data used to tune the parameters of a learning algorithm.
  - Learning algorithms typically have one or more free parameters, and validation sample is used to select appropriate values for these model parameters.
- **Test sample**
  - Samples used to evaluate the performance of a learning algorithm.
  - These predictions are then compared with the labels of the test sample to measure the performance of the algorithm.

Some Terminologies in Machine Learning

- **Loss function**
  - To measure the difference, or loss, between a predicted label and a true label.
  - Denote the set of all labels as Y and the set of possible predictions as Y', a loss function L is a mapping:

$$L: \ Y \times Y' \rightarrow \mathbb{R}+$$

- **Hypothesis set**
  - A set of functions mapping features to the set of labels Y.
  - For example, the following are a set of functions mapping email features to Y:

$$Y=\{\text{spam, non}-\text{spam}\}.$$

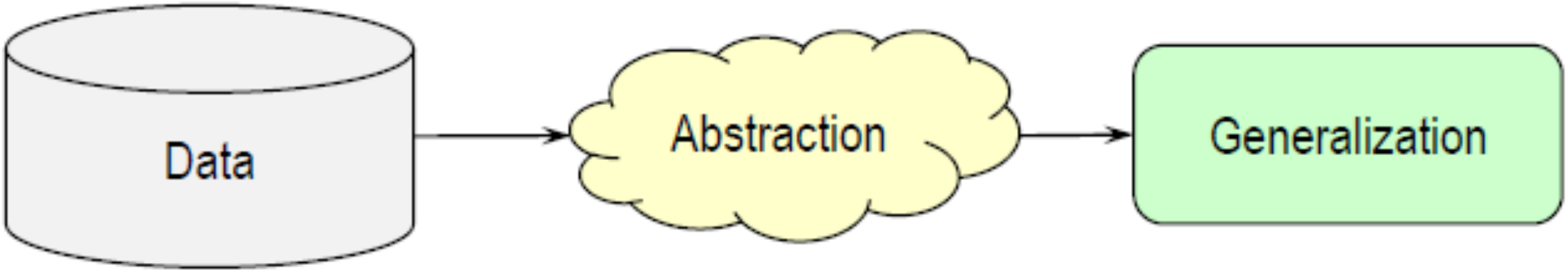
Some Terminologies in Machine Learning

- **Abstraction**

  It involves the translation of data into broader representations.

- **Generalization**

  It describes the process of turning abstracted knowledge into a form that can be utilized for action. It is also the ability of a learning algorithm to perform accurately on unseen samples after having experienced a learning data set.



Some Notations in This Course 本课程的一些符号

Notation 符号	Description 说明	
$\mathbb{R}$	Set of real numbers	实数集
$\mathbb{R}_+$	Set of non-negative real numbers	非负实数集
$\mathbb{R}^n$	Set of n-dimensional real-valued vectors	n维实值向量集
$[a, b]$	Closed interval between a and b	a和b之间的闭区间
(a, b)	Open interval between a and b	a和b之间的开区间
$\mathbb{N}$	Set of natural numbers, i.e., $\{0, 1, \ldots\}$	自然数集，即：$\{0, 1, \ldots\}$
S	An arbitrary set	任意集合
X	Input space	输入空间
Y	Target space	目标空间
H	hypothesis set	假设集

Summary

- Machine learning is to study some algorithms that can learn from and make predictions on data.
- The different perspectives are aimed to try to have a taxonomy on the algorithms of machine learning, for being easy to understand machine learning.
- Three perspectives on machine learning are proposed in this chapter, those are learning tasks, Learning Paradigms and learning models.