

# Artificial Intelligence



School of Electronic and Computer Engineering  
Peking University

Wang Wenmin



## Contents:

- ☐ Part 1. Basics
- ☐ Part 2. Searching
- ☐ Part 3. Reasoning
- ☐ Part 4. Planning
- ☐ Part 5. Learning

### Contents:

- ☐ 9. Perspectives about Machine Learning
- ☐ 10. Tasks in Machine Learning
- ☐ 11. Paradigms in Machine Learning
- ☐ 12. Models in Machine Learning

# Supervised Learning Paradigm



School of Electronic and Computer Engineering  
Peking University

Wang Wenmin

## Objectives 教学目的

- In this chapter we will discuss in detail about the paradigms that have been proposed in machine learning.  
这一章我们详细讨论针对机器学习所提出的一些范式。

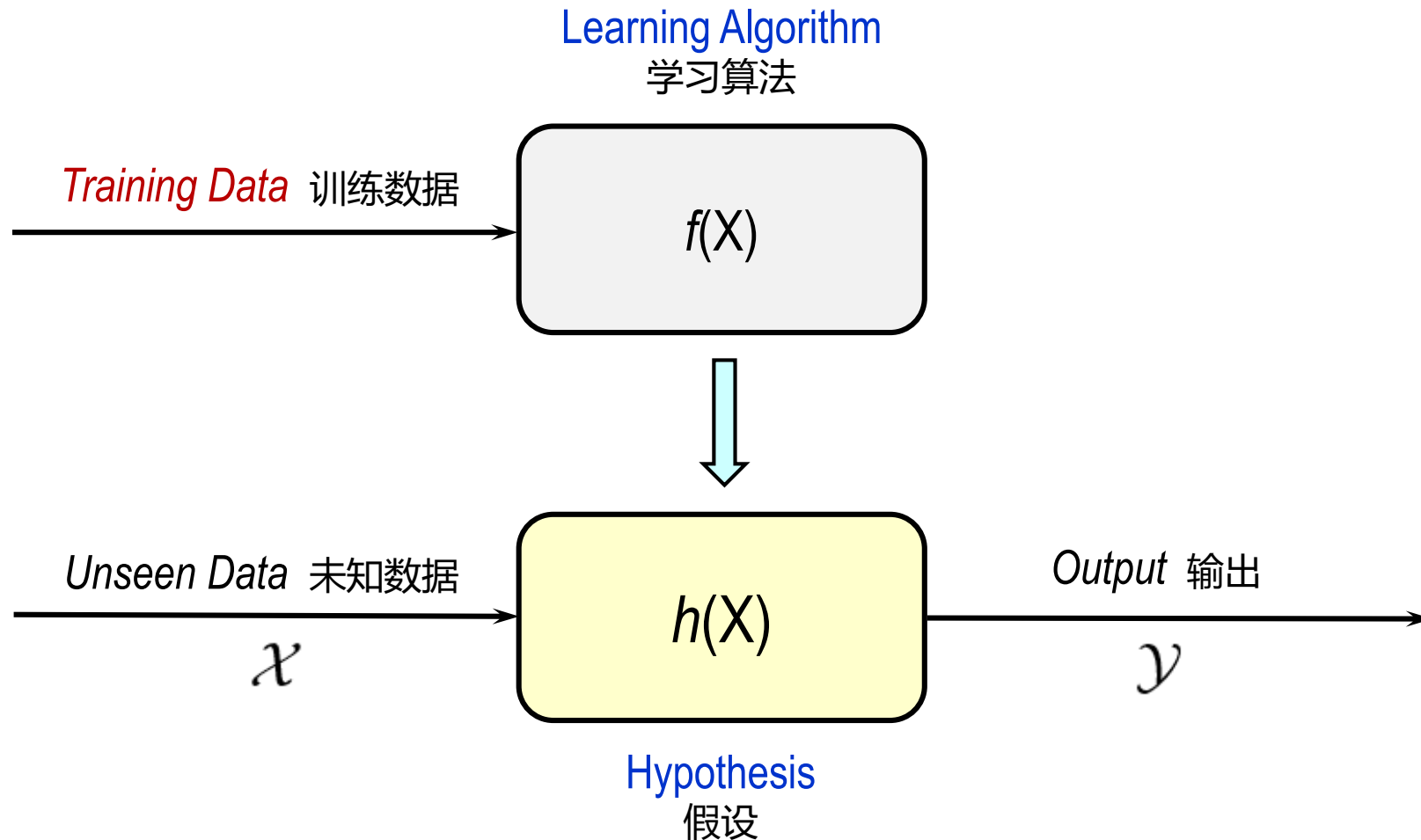
## What are Learning Paradigms 什么是学习的范式

- The learning paradigms are used to denote the typical scenarios that are happened in machine learning.  
学习范式用于表示机器学习中发生的典型场景。

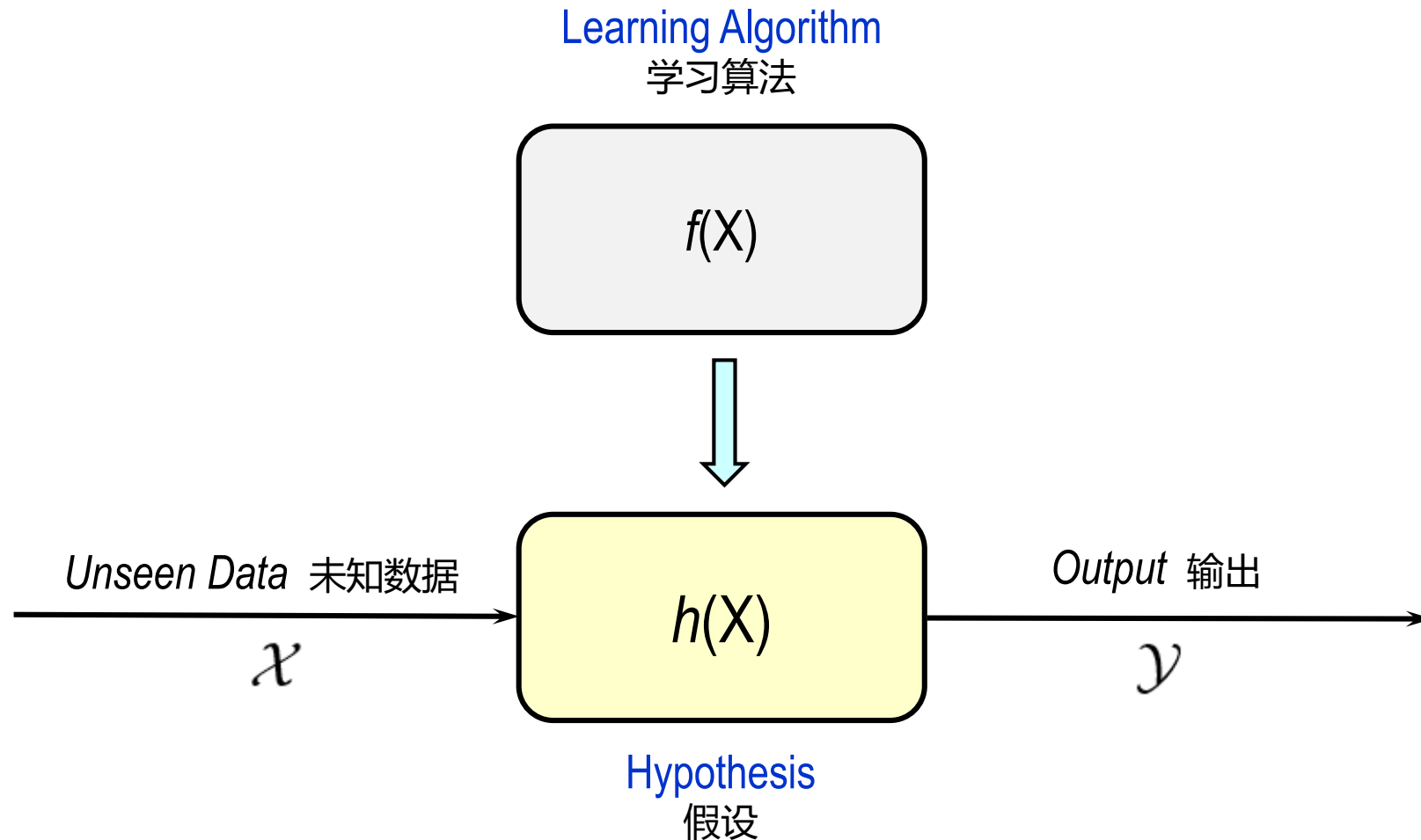
## Why Study Learning Paradigms 为什么要研究学习的范式

- Designing an algorithm to solve a learning task may take a different paradigm, such as based on its experience or the interaction with its environment.  
设计一种解决学习任务的算法可能会采用不同的范式，例如基于其经验、或者与其环境的交互。
- Why study the learning paradigms is because it can force you to think about a most appropriate paradigm for the learning task in order to get the best result.  
研读学习范式的意义在于，它可以使你考虑一个最适合该学习任务的范式，以获得最好的结果。

# How Does Machine Learning Work 机器学习是如何工作的

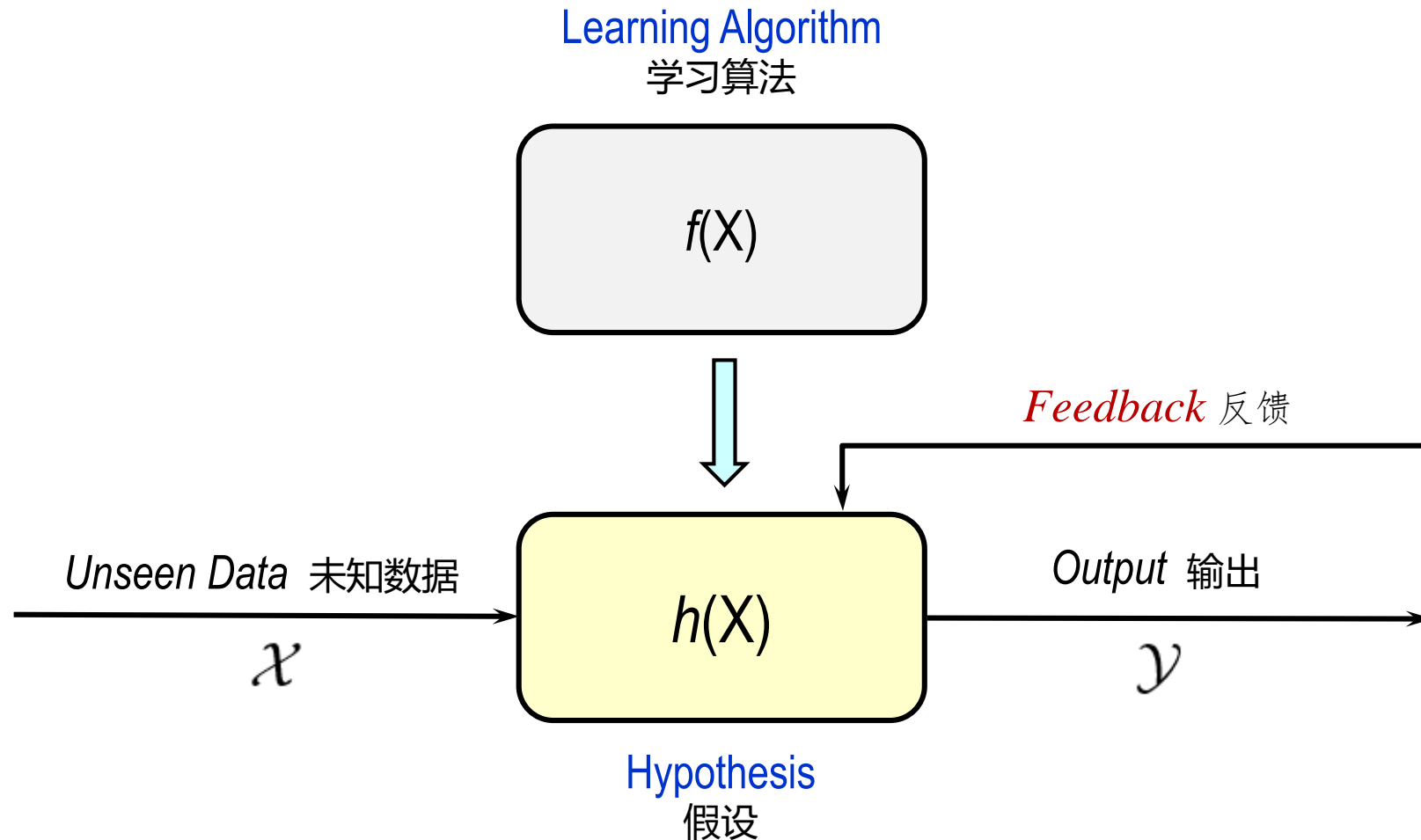


# How Does Machine Learning Work 机器学习是如何工作的





# How Does Machine Learning Work 机器学习是如何工作的



# Typical Paradigms in Machine Learning 机器学习中的典型学习范式

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

## Contents:

- ☐ 11.1. Supervised Learning Paradigm
- ☐ 11.2. Unsupervised Learning Paradigm
- ☐ 11.3. Reinforcement Learning Paradigm
- ☐ 11.4. Relations and Other Paradigms

## What is Supervised Learning 什么是有监督学习

- The agent receives a set of labeled examples as training data, and makes predictions for all unseen points.

智能体接收一组标注的样本作为训练数据，然后对所有的未知点进行推测。

- This approach attempts to generalize a function or mapping from inputs to outputs by training, which can then be used speculatively to generate an output for previously unknown data.

这种方式试图生成从输入到输出的函数或映射，然后可以将其用于对预先未知的数据生成输出。

*It is a way of “teaching” the learning algorithm, like that a “teacher” gives the classes (courses).*

这是一种“教”学习算法的方式，就像“老师”讲授课程那样。

## What is Supervised Learning 什么是有监督学习

- The training data in supervised learning:  
有监督学习中的训练数据：
  - each training data has a known label as an input data,  
每个训练数据具有一个已知标注作为输入数据，
  - the label is a pair consisting of an input object and a desired output value  
标注是由输入对象和预期输出值组成的对  
(such as spam/not-spam, or a stock price at a time)  
( 例如垃圾与非垃圾邮件、或某时刻股票价格 )。
- An hypothesis function after training:  
训练后的假设函数：
  - can be used for mapping new unseen data.  
可用于映射新的未知数据。

### Contents:

- ☐ 11.1.1. Overview of Supervised Learning
- ☐ 11.1.2. Suitable Learning Tasks
- ☐ 11.1.3. Formal Description
- ☐ 11.1.4. Algorithms of Supervised Learning
- ☐ 11.1.5. Applications of Supervised Learning
- ☐ 11.1.6. Variants of Supervised Learning

# Six Steps by Supervised Learning 有监督学习的6个步骤

4) *Determine the algorithm to the task* 设计该任务的算法

3) *Determine the feature extraction approach*

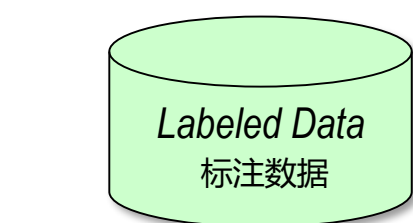
确定特征提取方法

2) *Gather a training set*

收集训练集

1) *Determine the training type*

确定训练类型



$x_1$	$y_1$
$x_2$	$y_2$

$$f(x) = y$$

Training  
训练  
( $x, y$ )

Learning Algorithm  
学习算法

$f(X)$

5) *Training the algorithm*

训练该算法

*with small generalization and empirical errors*  
具有小的泛化和经验错误

$X$

$h(X)$

Hypothesis  
假设

6) *Evaluate the accuracy*

评估其精确性

$Y$

## Six Steps by Supervised Learning 有监督学习的6个步骤

### □ 1) Determine the training type 确定训练类型

You should decide firstly what kind of data is to be used as a training set.

应该首先确定使用何种数据作为训练集。

- E.g., for handwriting recognition, that may be a single handwritten character, an entire handwritten word, or an entire line of handwriting.

例如，对于手写体识别，可以是一个手写字符、一个完整的手写单词、或是手写的一行。

### □ 2) Gather a training set 收集训练集

The training set needs to be representative of the real-world use of the function.

训练集需要代表实际使用的功能。

- Thus, a set of input objects is gathered and corresponding outputs are also gathered, either from human experts or from measurements.

因此，由人类专家或者通过测量，筛选出一组输入对象以及对应的输出。



## Six Steps by Supervised Learning 有监督学习的6个步骤

### □ 3) Determine the feature extraction approach 确定特征提取方法

Typically, there are two kind of approaches to extract the feature from input data:  
通常，有两种从输入数据提取特征的方法：

- *handcraft* feature extraction: by some feature descriptor.

手工特征提取：通过某种特征描述子。

- *automated* feature extraction: by some deep neural network.

自动特征提取：通过某种深度神经网络。

### □ 4) Design the algorithm to the task 设计该任务的算法

This depends on what your task is.

这取决于你的任务是什么。

- E.g., for classification, you may choose to use SVM, decision tree, Softmax, etc.

例如，对于分类来说，你可以选择使用SVM、决策树、Softmax、等等。

## Six Steps by Supervised Learning 有监督学习的6个步骤

### □ 5) Training the algorithm 训练该算法

Run the learning algorithm on the gathered training set.

在收集的训练数据集上运行该学习算法。

- Some algorithms require the user to determine certain control parameters.  
某些算法需要用户来确定某些控制参数。
- These parameters may be adjusted by optimizing performance on a subset of the training set.  
这些参数可以通过在训练子集上优化性能来调整。

### □ 6) Evaluate the accuracy 评估其精确性

After parameter adjustment and learning, the performance of the resulting function should be measured on a validation set that is separate from the training set.

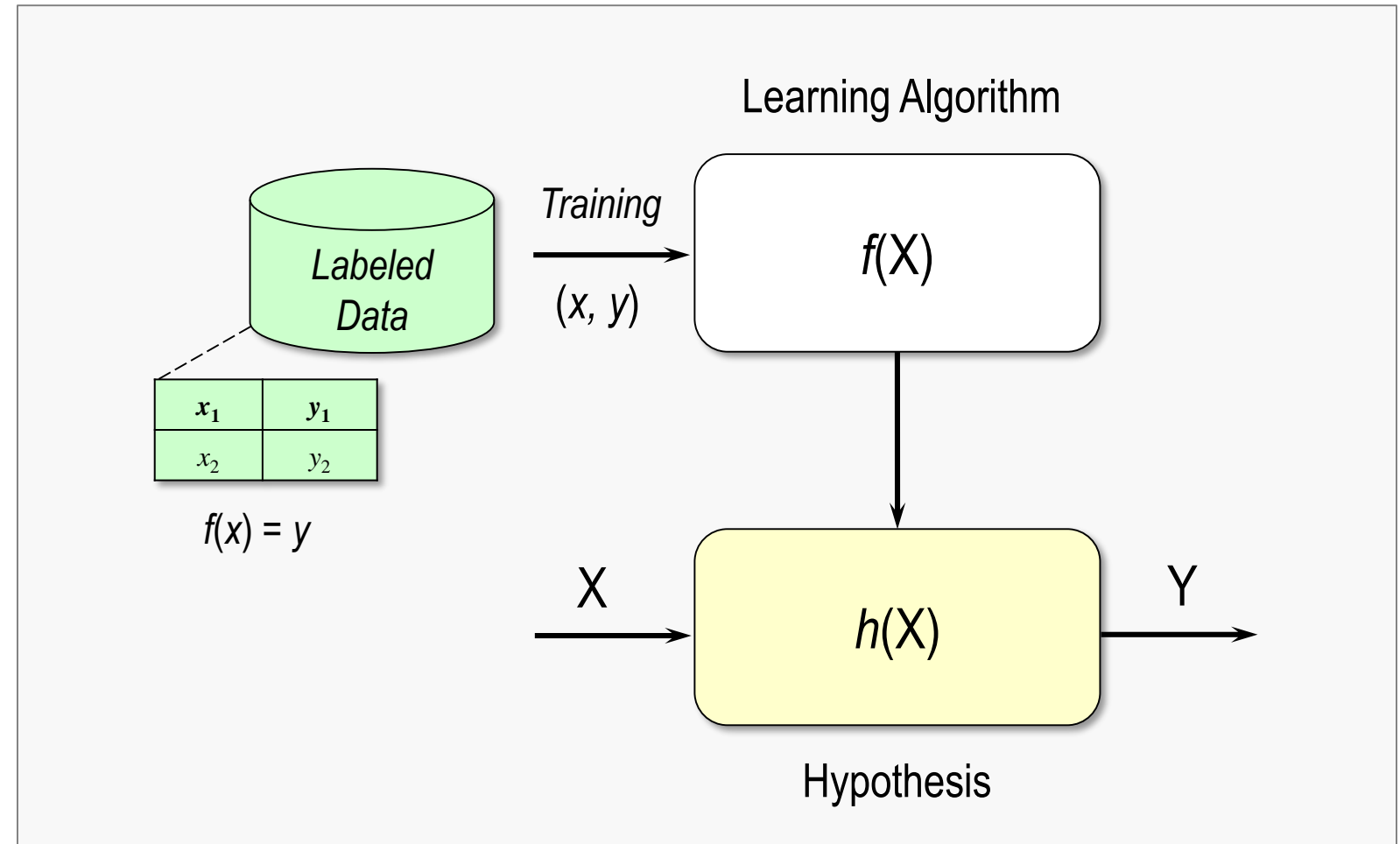
在参数调整和学习之后，应当在（独立于训练集的）验证数据集上对结果函数的性能进行度量。

## Contents:

- ☐ 11.1.1. Overview of Supervised Learning
- ☐ 11.1.2. Suitable Learning Tasks
- ☐ 11.1.3. Formal Description
- ☐ 11.1.4. Algorithms of Supervised Learning
- ☐ 11.1.5. Applications of Supervised Learning
- ☐ 11.1.6. Variants of Supervised Learning

# Tasks Associated with Supervised Learning 与有监督学习相关的任务

- ☐ Classification,  
分类
- ☐ Regression,  
回归
- ☐ Ranking.  
排名



### Contents:

- ☐ 11.1.1. Overview of Supervised Learning
- ☐ 11.1.2. Suitable Learning Tasks
- ☐ 11.1.3. Formal Description
- ☐ 11.1.4. Algorithms of Supervised Learning
- ☐ 11.1.5. Applications of Supervised Learning
- ☐ 11.1.6. Variants of Supervised Learning

## A Formal Description for Supervised Learning 一种有监督学习的形式化描述

Let  $X$  denote input space,  $Y$  denote output space, and  $D$  an unknown distribution over  $X \times Y$ .

设 $X$ 表示输入空间， $Y$ 表示输出空间，并且 $D$ 表示 $X \times Y$ 上的一个未知分布。

□ Let target labeling function: 设目标标注函数

$$f: \square \rightarrow \square$$

□ Training set (a labeled sample set): 训练集（标注的训练样本集）

$$\square = \{(x^{(i)}, y^{(i)}) \mid (x, y) \in \square \times \square, i \in [1, m]\}$$

□ Given a hypothesis set  $H$ , to find a hypothesis  $h \in H$  that is the mapping:

给定假设集 $H$ ，来发现一个假设 $h \in H$ ，满足如下映射：

$$h: \square \rightarrow \square$$

# A Formal Description for Supervised Learning 一种有监督学习的形式化描述

## □ Classification 分类

output space  $Y$  is a set of **categories**.

输出空间 $Y$ 是一组类别。

## □ Regression 回归

output space  $Y$  is a set of **real continues numbers**.

输出空间 $Y$ 是一组连续的实数值。

## □ Ranking 排名

output space  $Y$  is a set **with relative order**.

输出空间 $Y$ 是一组相对的顺序。

## Contents:

- ☐ 11.1.1. Overview of Supervised Learning
- ☐ 11.1.2. Suitable Learning Tasks
- ☐ 11.1.3. Formal Description
- ☐ 11.1.4. Algorithms of Supervised Learning
- ☐ 11.1.5. Applications of Supervised Learning
- ☐ 11.1.6. Variants of Supervised Learning

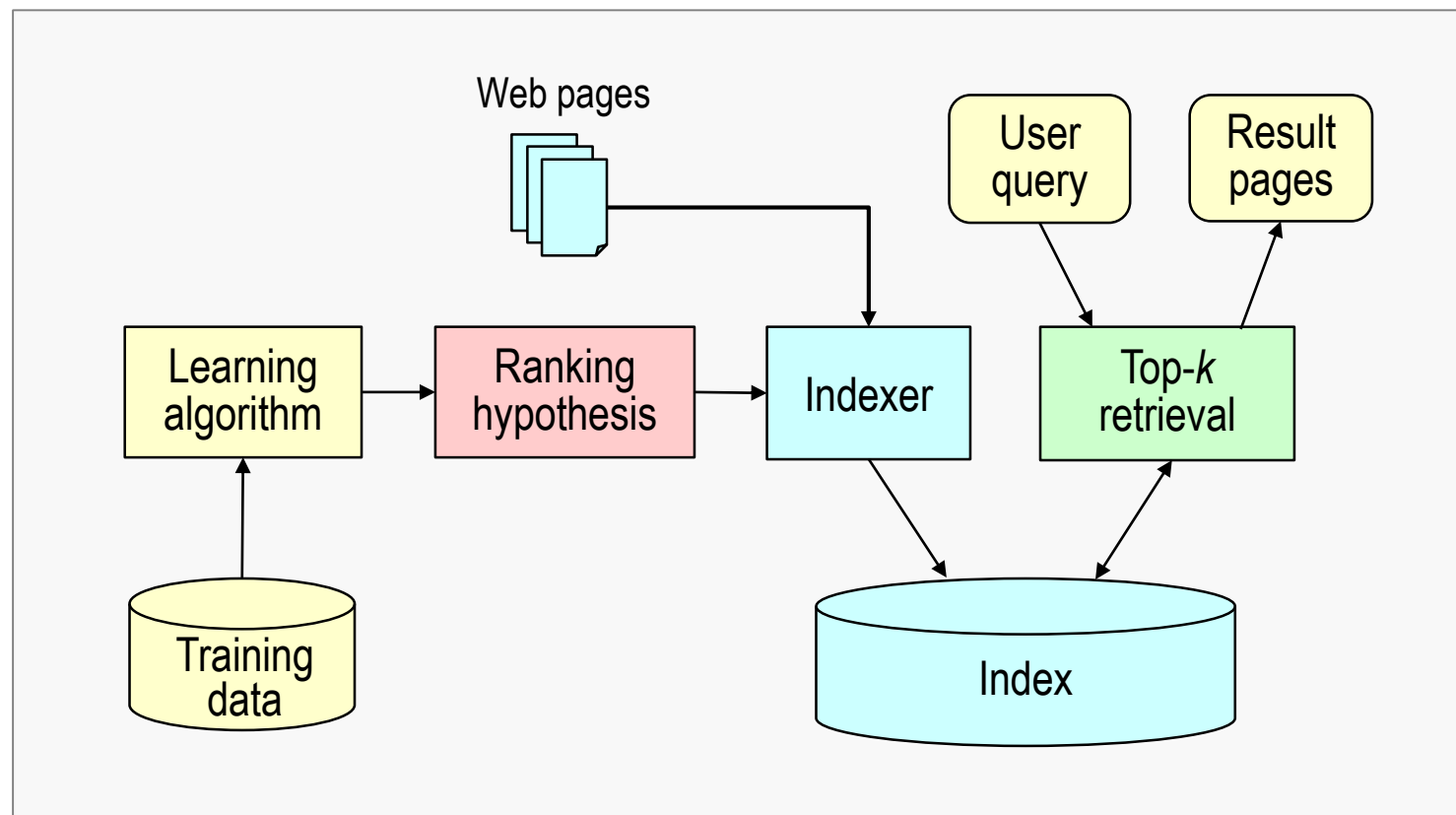


# Typical Classification and Regression Algorithms 典型的分类与回归算法

Algorithm 算法	Task Types 任务类型	Predictive accuracy 预测精度	Training speed 训练速度
AdaBoost 自适应增强	Either 两者	Higher 高	Slow 慢
Artificial neural network 人工神经网络	Either 两者	Higher 高	Slow 慢
$k$ -Nearest neighbor $k$ 近邻	Either 两者	Lower 低	Fast 快
<u>Linear regression</u> 线性回归	<u>Regression</u> 回归	Lower 低	Fast 快
<u>Logistic regression</u> 逻辑回归	<u>Classification</u> 分类	Lower 低	Fast 快
Naive Bayes 朴素贝叶斯	Classification 分类	Lower 低	Fast 快
Decision tree 决策树	Either 两者	Lower 低	Fast 快
Random Forests 随机森林	Either 两者	Higher 高	Slow 慢
Support vector machines 支撑向量机	Either 两者	Higher 高	Slow 慢

## Typical Ranking Algorithms 典型的排名算法

- ☐ AdaRank
- ☐ BayesRank
- ☐ BoltzRank
- ☐ LambdaRank
- ☐ RankBoost
- ☐ Ranking Refinement
- ☐ RankSVM
- ☐ PageRank



A concept architecture of a machine-learned search engine.

一种机器学习搜索引擎的概念架构

## Contents:

- ☐ 11.1.1. Overview of Supervised Learning
- ☐ 11.1.2. Suitable Learning Tasks
- ☐ 11.1.3. Formal Description
- ☐ 11.1.4. Algorithms of Supervised Learning
- ☐ 11.1.5. Applications of Supervised Learning
- ☐ 11.1.6. Variants of Supervised Learning

## Some Applications of Supervised Learning 有监督学习的一些应用

Object recognition in computer vision	<input type="checkbox"/> 计算机视觉中的物体识别
Optical character recognition (OCR)	<input type="checkbox"/> 光学字符识别 (OCR)
Handwriting recognition	<input type="checkbox"/> 手写体识别
Information retrieval	<input type="checkbox"/> 信息检索
Learning to rank	<input type="checkbox"/> 学会排名
Spam detection	<input type="checkbox"/> 垃圾邮件检测
Speech recognition	<input type="checkbox"/> 语音识别
Bioinformatics	<input type="checkbox"/> 生物信息学
Cheminformatics	<input type="checkbox"/> 化学信息学

## Some Examples of Supervised Learning 几个有监督学习的例子

### □ Spam Detection 垃圾邮件检测

Mapping email to {Spam, Not Spam}

将电子邮件分为 {Spam, Not Spam}

### □ Digit Recognition 数字识别

Mapping handwriting digit to {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}

将手写体数字映射为 {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}

### □ Price Prediction for Used Cars 二手车价格预测

Mapping a used car to a real price, based on the historical data collected from used car market.

根据二手车市场收集到的历史数据，估算一台二手车的实际价格。

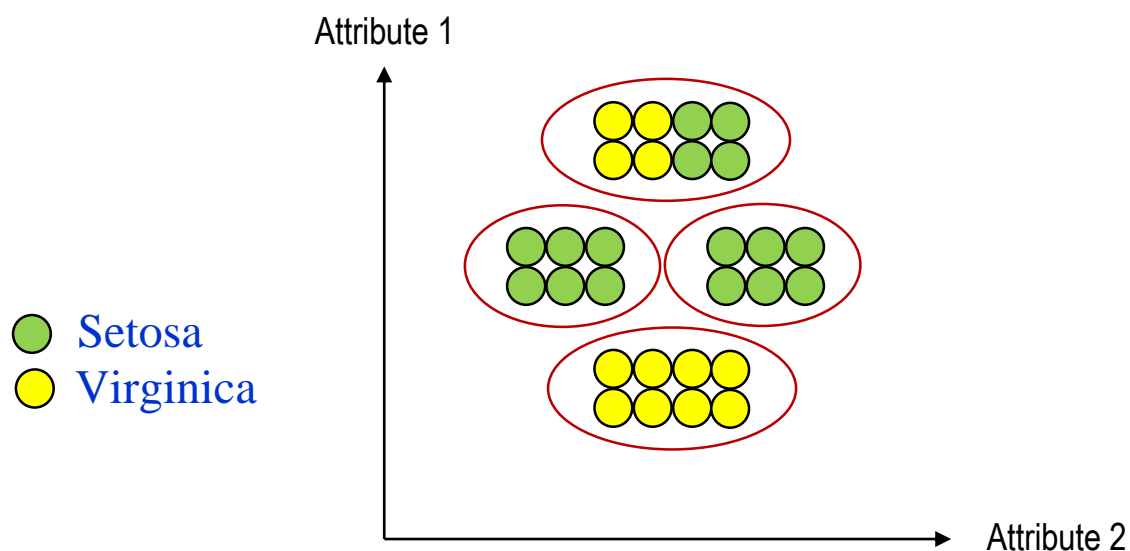
## Special Case 特例

### □ Supervised Clustering 有监督聚类

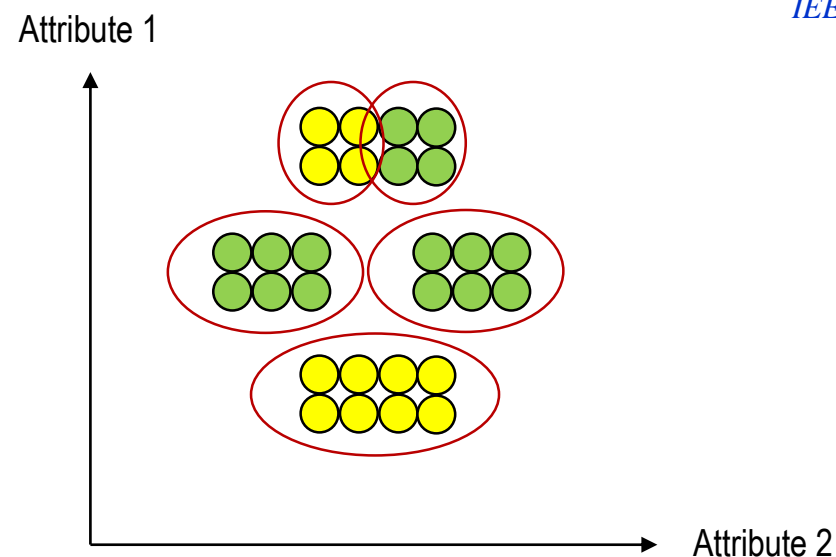
to identify class-uniform clusters that have high probability densities.

识别具有高概率密度的类别统一类聚。

Source: "Supervised clustering - algorithms and benefits",  
IEEE ICTAI, 2004



(a) Traditional Clustering 传统的聚类



(b) Supervised Clustering 有监督聚类

Differences between Traditional Clustering and Supervised Clustering  
传统聚类和有监督聚类的差异

### Contents:

- ☐ 11.1.1. Overview of Supervised Learning
- ☐ 11.1.2. Suitable Learning Tasks
- ☐ 11.1.3. Formal Description
- ☐ 11.1.4. Algorithms of Supervised Learning
- ☐ 11.1.5. Applications of Supervised Learning
- ☐ 11.1.6. Variants of Supervised Learning

# Variants of Supervised Learning 有监督学习的变体

Paradigms 范式	Brief Statements 简介
Semi-supervised learning 半监督学习	A class of supervised learning techniques that also make use of unlabeled data for training. 属于有监督学习算法一类，此外还利用未标记数据进行训练。
Weakly supervised learning 弱监督学习	It aims to learn some information using a limited amount of training examples. 旨在采用有限数量的训练样本来学习一些信息。
One-shot learning 一次性学习	It aims to learn some information from one, or only a few, training examples. 旨在从一个、或仅有的几个训练样本中学习一些信息。
Zero-shot learning 零次性学习	It is able to solve a task despite not having received any training examples of that task. 即使没有得到某个任务的任何训练样本，也能够求解该任务。



## Contents:

- ☐ 11.1. Supervised Learning Paradigm
- ☐ 11.2. Unsupervised Learning Paradigm
- ☐ 11.3. Reinforcement Learning Paradigm
- ☐ 11.4. Relations and Other Paradigms

Thank you for your attention!

**AI**