Artificial Intelligence

AI

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

AI

Contents:

Tasks in Machine Learning

School of Electronic and Computer Engineering
Peking University

Wang Wenmin

Objectives 教学目的

■ This chapter will discuss in detail about the tasks that can be solved with machine learning.
本章将详细讨论可以通过机器学习解决的一些任务。

What are Learning Tasks 什么是学习任务

☐ The learning tasks are used to denote the general problems that can be solved by learning with desired output.

学习任务用于表示可以用机器学习解决的基本问题。

Why Study Learning Tasks 为什么要研究学习任务

☐ Various types of problems arising in applications:

应用中会产生各种类型的问题：

■ computer vision，计算机视觉，

■ pattern recognition，模式识别，

■ natural language processing，自然语言处理，

■ etc. 等等。

Typical Tasks in Machine Learning 机器学习中的典型任务

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

Contents:

Classification

School of Electronic and Computer Engineering
Peking University

Wang Wenmin

What is Classification 什么是分类

☐ A longer description 较长描述
Classification is the task of identifying to which of a set of categories a new observation belongs, on the basis of a training set of data containing observations whose category membership is known.

分类是基于包含已知类别成员观测值的训练数据集、来辨识新的观测值属于哪一组类别的任务。

☐ A shorter description 较短描述
To resolve such problems where the output is divided into two or more categories.

解决输出被分为两个或多个类别的问题。

☐ A very short description 极简描述
Assign a category to each item.

为每个项指定一个类别。

AI

Contents:

Classifier 分类器

☐ About classifier 关于分类器

An algorithm that implements classification, especially in a concrete implementation, is known as a classifier.
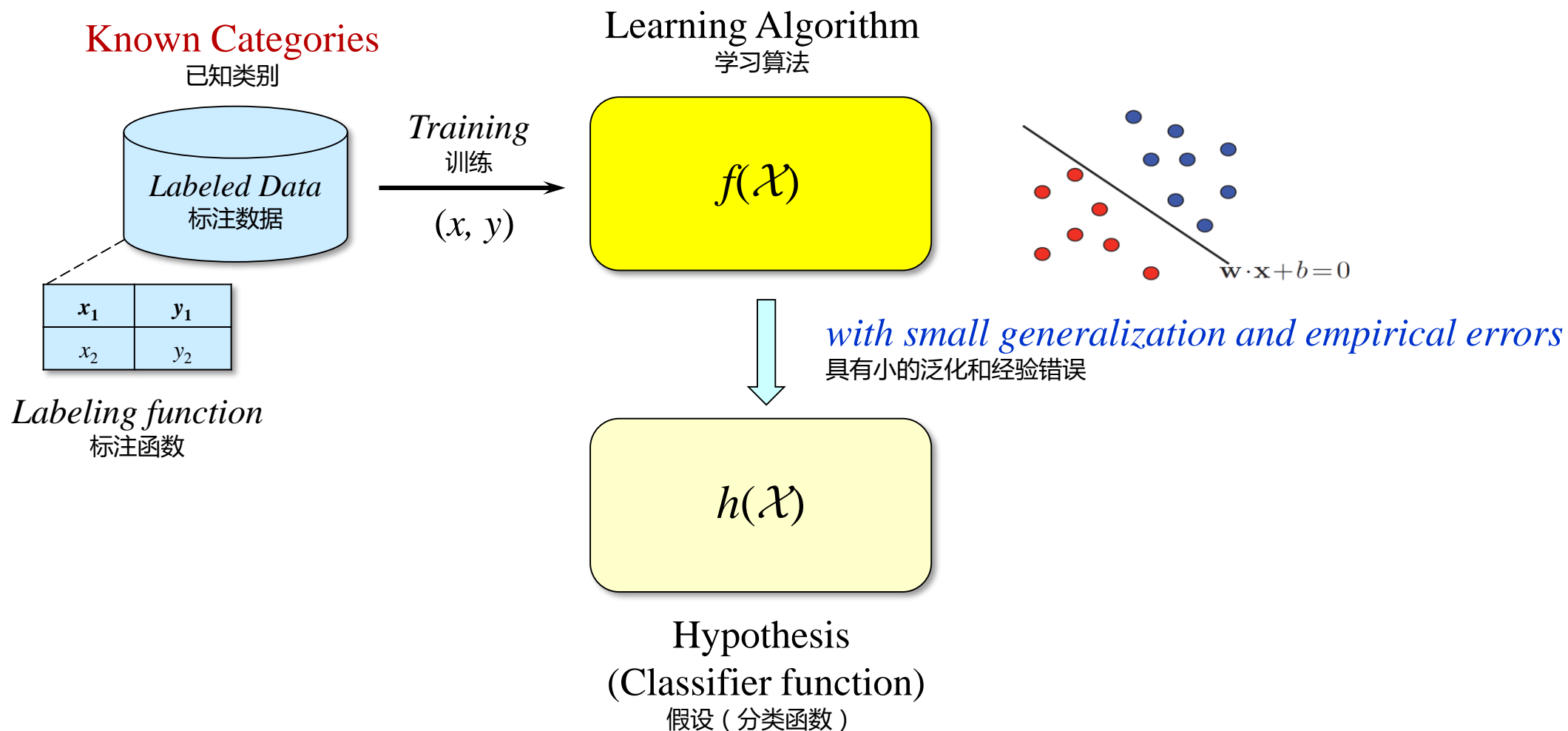
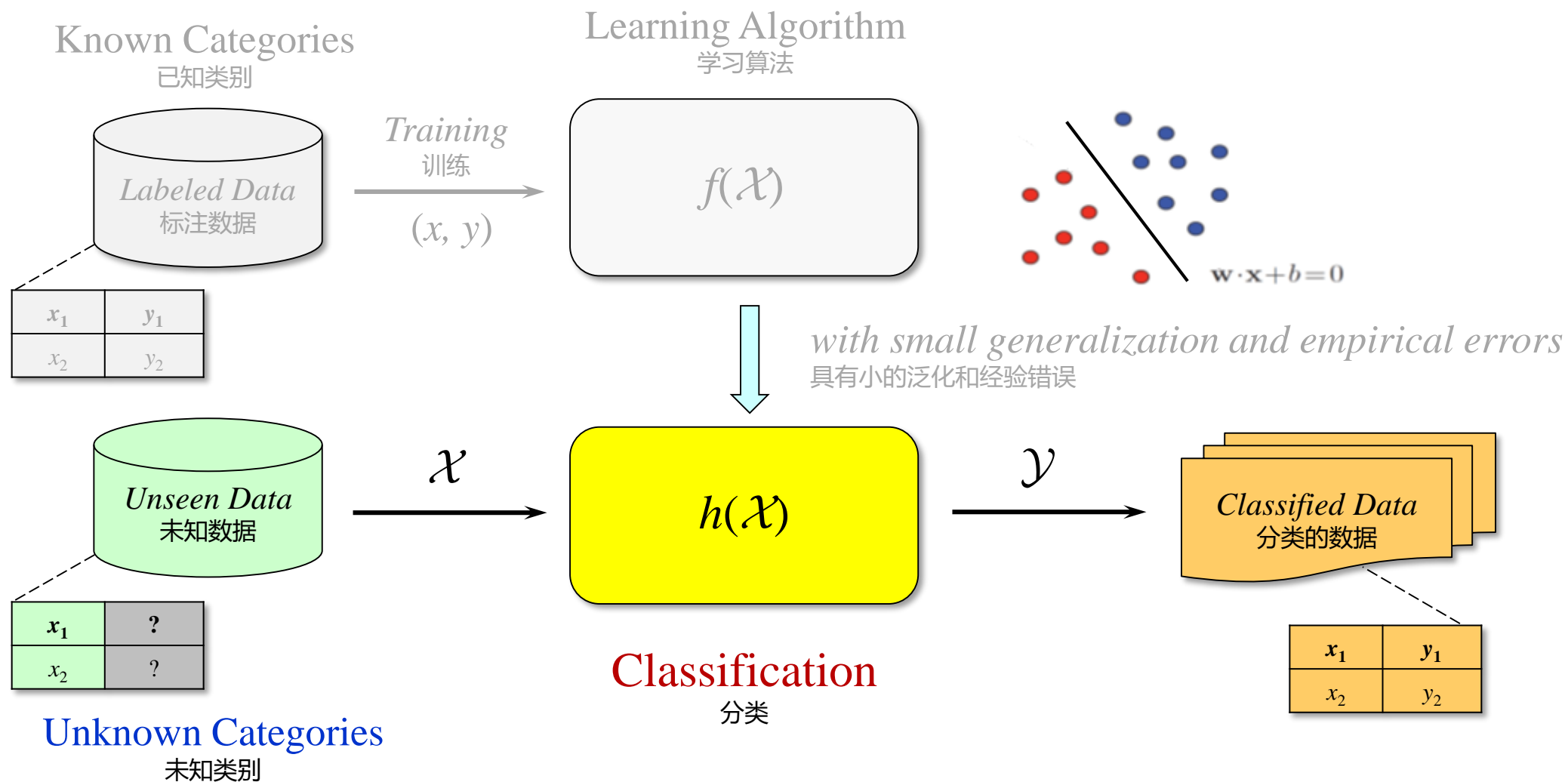一种实现分类、尤其是构成一种具体实现的算法，被称为一个分类器。

☐ About classifier function 关于分类器函数

The term "classifier" sometimes also refers to the mathematical function, implemented by a classification algorithm, that maps input data to a category.

"分类器"这个术语有时还指的是由分类算法所实现的数学函数，它将输入数据映射为一个类别。

Classification: Training 分类：训练

Known Categories
已知类别

Learning Algorithm
学习算法

Training
训练

(x, y)

Labeled Data
标注数据

$f(\mathcal{X})$

$$\mathbf{w} \cdot \mathbf{x} + b = 0$$

x_1	y_1
x_2	y_2

Labeling function
标注函数

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

$h(\mathcal{X})$

Hypothesis
(Classifier function)
假设（分类函数）

Classification: Testing 分类：实测

Known Categories
已知类别

Learning Algorithm
学习算法

Training
训练
(x, y)

Labeled Data
标注数据

$f(\mathcal{X})$

$\mathbf{w} \cdot \mathbf{x} + b = 0$

x_1	y_1
x_2	y_2

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

Unseen Data
未知数据

$\mathcal{X}$

$h(\mathcal{X})$

$\mathcal{Y}$

Classified Data
分类的数据

x_1	?
x_2	?

Classification
分类

x_1	y_1
x_2	y_2

Unknown Categories
未知类别

A Formal Description of Classification 一种分类的形式化描述

Let $\mathbb{R}^n$ ($n \cong 1$) denote a set of n-dimensional real-valued vectors, input space $\mathcal{X}$ is a subset of $\mathbb{R}^n$, output space $\mathcal{Y}$ is a set of categories, D is an unknown distribution over $\mathcal{X} \times \mathcal{Y}$, then:

设$\mathbb{R}^n$ ($n \cong 1$)表示一个n维实值向量集合，输入空间$\mathcal{X}$是$\mathbb{R}^n$的一个子集，输出空间$\mathcal{Y}$是一组类别，D是$\mathcal{X} \times \mathcal{Y}$的一个未知分布，则：

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

$$f : \mathcal{X} \to \mathcal{Y}$$

☐ Training set (Labeled training sample set): 训练集（标注的训练样本集）

$$\mathcal{S} = \{(x^{(i)}, y^{(j)}) \mid (x, y) \in \mathcal{X} \times \mathcal{Y}, i \in [1, m], j \in [1, n]\}$$

☐ Classification algorithm: 分类算法
Let a hypothesis set H are the mapping $\mathcal{X}$ to $\mathcal{Y}$, to determine a hypothesis (classifier function):

设一个假设函数集H是$\mathcal{X}$到$\mathcal{Y}$的映射，来决定一个假设（分类器函数）：

$$h : \mathcal{X} \to \mathcal{Y} \ \text{ and } \ h \in H$$

with small generalization error: 具有小的泛化错误

$$R(h) = \Pr_x[h(x) \neq f(x)]$$

A Formal Description of Classification 一种分类的形式化描述

☐ Classification: 分类

Given a testing data set of unknown categories:

给定一个未知类别的实测数据集：

$$\mathcal{X} = \{x^{(i)} \ / \ x \in \mathcal{X}, \ i \in [1, \ m]\}$$

Using the classifier function $h(\mathcal{X}) = \mathcal{Y}$ determined at above to predicate classifying results:

使用前面训练好的分类函数$h(\mathcal{X}) = \mathcal{Y}$来预测分类结果：

$$\mathcal{Y} = h(\mathcal{X}) = \{y^{(j)} \ / \ y \in \mathcal{Y}, \ j \in [1, \ n], \ h(x)=y\}$$

where 其中

$\mathcal{Y}$ is the set of known categories.

Y是该已知类别的集合。

Contents:

- ☐ 10.1.1. How Classification Works
- ☐ 10.1.2. Linear and Nonlinear
- ☐ 10.1.3. Dimensions and Classes
- ☐ 10.1.4. Applications and Algorithms

Linear Classification 线性分类

☐ Linear Classification is doing classification by a linear classifier.
  线性分类是通过线性分类器来进行分类。
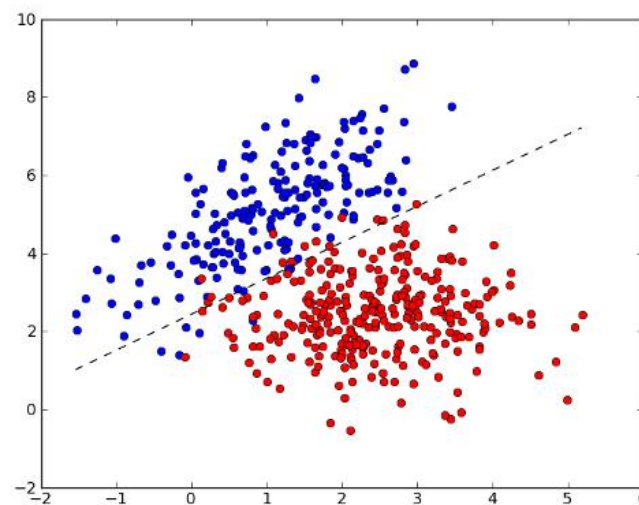
☐ A linear classifier is 一个线性分类器是

  ■ a linear discriminant function with a linear decision boundary.
    具有一个线性决策边界的线性判别函数。

Case Study: A Typical Linear Classifier 一个典型的线性分类器

$$H = \{ \mathbf{x} \mapsto y(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b \mid \mathbf{w} \in \mathbb{R}^n, b \in \mathbb{R} \}$$

where, 其中

$\mathbf{w}$ denotes a row vector, called a *weight vector,*

$\mathbf{w}$表示行向量、称为权向量，

$$\mathbf{w} = (w_1, ..., w_n)$$

$\mathbf{x}$ denotes a column vector,

$\mathbf{x}$表示列向量，

$$\mathbf{x} = (x_1, ..., x_n)^{\mathrm{T}}$$

b denotes a bias.

b表示偏差。

$y(\mathbf{x}) > 0$

$y(\mathbf{x}) = 0$

$y(\mathbf{x}) < 0$
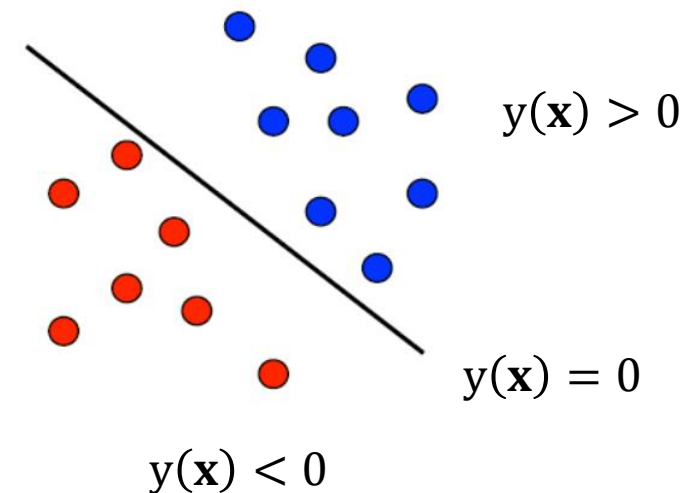
Nonlinear Classification 非线性分类

☐ Nonlinear Classification is doing classification by a nonlinear classifiers.

非线性分类是通过一个非线性分类器来进行分类。

☐ A nonlinear classifiers have 一个非线性分类器具有

nonlinear decision boundaries, and possibly discontinous decision boundaries.

若干非线性决定边界，并且可能是非连续决定边界。
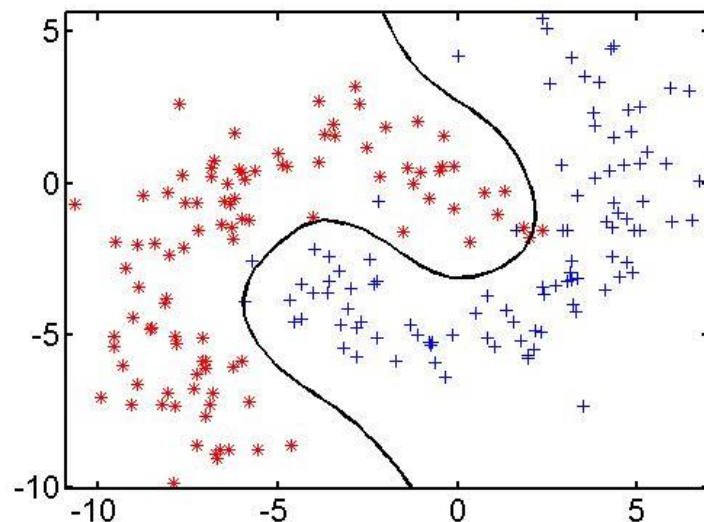
E.g., a nonlinear classifier in SVM is a nonlinear kernel function.

例如，在SVM中的非线性分类器是一个非线性核函数。

Contents:

☐ 10.1.1. How Classification Works

☐ 10.1.2. Linear and Nonlinear

☐ 10.1.3. Dimensions and Classes

☐ 10.1.4. Applications and Algorithms

Dimensions 维度

☐ If the problem space is n dimensional then its linear classifier is n-1 dimensional hyper-plane. E.g.,
如果问题空间的维度为n，则它的线性分类器的维度为n-1的超平面。例如：
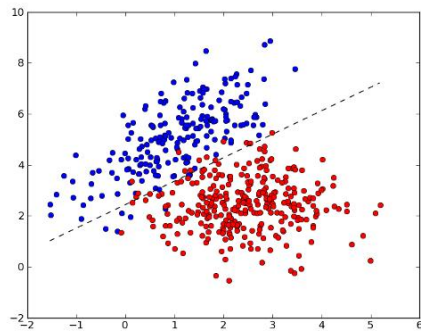
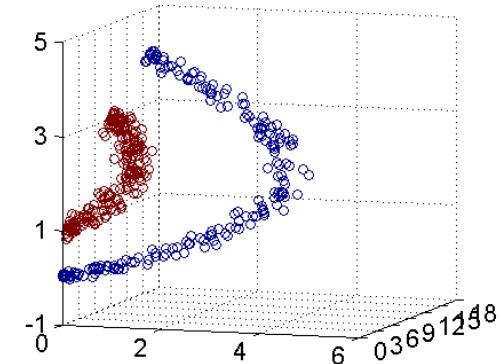2-dimensions
2维

in 2-dimensions, the hyper-plane is a line
2维空间中，该超平面为一条线
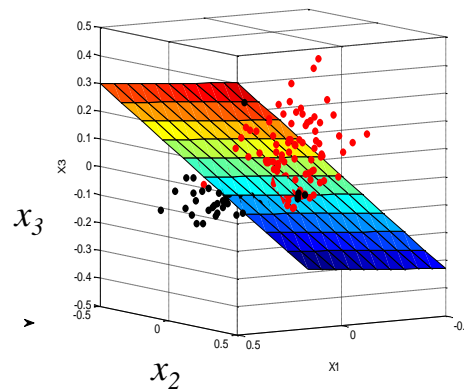
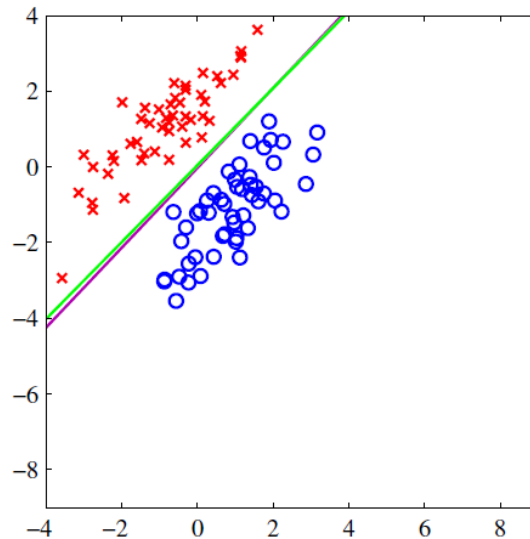3-dimensions
3维

in 3-dimensions, the hyper-plane is a plane
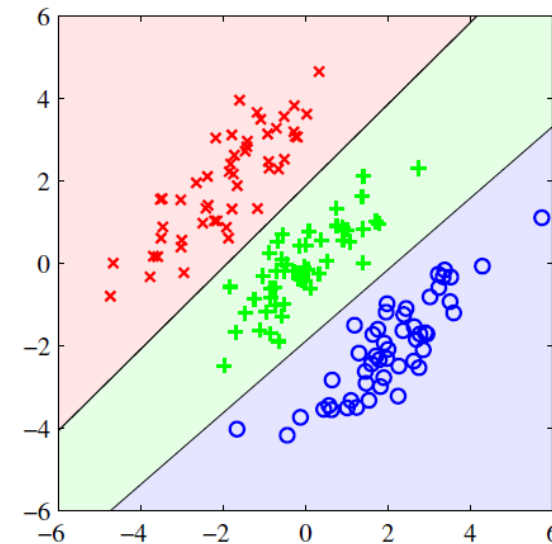3维空间中，该超平面为一个平面

Classes 类别

$$y_k(\mathbf{x}) = \mathbf{w}_k \cdot \mathbf{x} + b$$

■ **Two classes:**　　二元分类：　$k = 2$

■ **Multiple classes:** 多元分类：　$k > 2$

Two classes
二元分类

Three classes
三元分类

Case Study: Softmax Classifier Softmax分类器

☐ It is a multiclass classifier, implemented by a softmax function.
  这是一个多元分类器，由Softmax函数来实现。

☐ Softmax function maps a K-dimensional vector $\mathbf{x}$ of arbitrary real values to a K-dimensional vector $\sigma(\mathbf{x})$ of real values (range 0 to 1, add up to 1).
  Softmax函数将一个任意实数值的K维向量$\mathbf{x}$映射到一个实数值的K维向量$\sigma(\mathbf{x})$ (范围0到1，和为1)。

$$\sigma(\mathbf{x})_j = \frac{e^{x_j}}{\sum_{k-1}^{K} e^{x_k}} \qquad j = 1, \dots, K$$

☐ In probability theory, the output of the softmax function can be represented a categorical distribution, i.e., a probability distribution over K different outcomes.
 在概率论中，Softmax函数的输出可以被用来表示一个类别分布，即，一个涵盖K个不同结果的概率分布。

$$P(y = j \ / \ \mathbf{x}) = \frac{e^{\mathbf{x}^{\mathrm{T}}\mathbf{w}_j}}{\sum_{k-1}^{K} e^{\mathbf{x}^{\mathrm{T}}\mathbf{w}_k}}$$

Case Study: Softmax Classifier Softmax分类器

☐ Softmax function has been used in various multiclass classification methods, such as:

Softmax函数已经被用于各种多元分类方法，例如：

■ multinomial logistic regression,
多项式逻辑回归，

■ multiclass linear discriminant analysis,
多元线性判别分析，

■ naive Bayes classifiers,
朴素贝叶斯分类器，

■ artificial neural networks (ANN),
人工神经网络 (ANN)，

■ reinforcement learning.
强化学习。

Softmax function

$P(y=1 / \mathbf{x})$

$P(y=2 / \mathbf{x})$

$P(y=m / \mathbf{x})$

Output layer
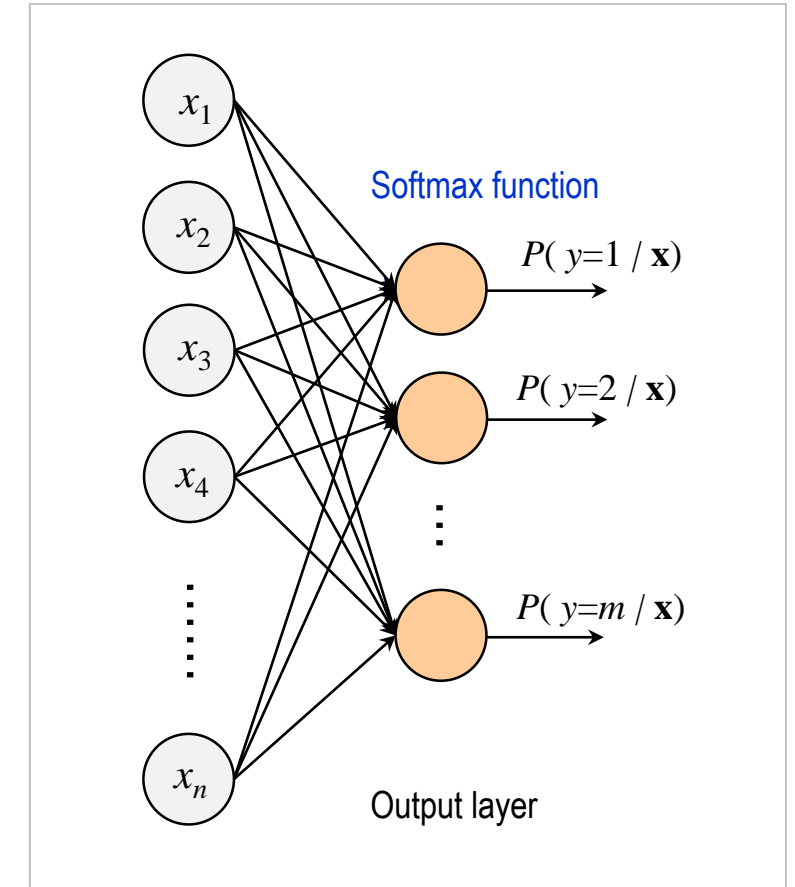
Softmax function used in ANN
as the final layer for multiclass classification
Softmax函数在ANN的最后一层用于多元分类

AI

Contents:

☐ 10.1.1. How Classification Works

☐ 10.1.2. Linear and Nonlinear

☐ 10.1.3. Dimensions and Classes

☐ 10.1.4. Applications and Algorithms

Typical Applications of Classification 分类的典型应用

- ☐ **Computer vision** 计算机视觉
  - ■ **Face, handwriting recognition** 人脸、手写体识别
  - ■ **Action recognition** 动作识别
  - ■ **Medical image analysis** 医学图像分析
  - ■ **Video tracking** 视频跟踪
- ☐ **Pattern recognition** 模式识别
- ☐ **Biometric identification** 生物特征识别
- ☐ **Statistical natural language processing** 统计自然语言处理
- ☐ **Document classification** 文档分类
- ☐ **Internet search engines** 互联网搜索引擎
- ☐ **Credit scoring** 信用评分

Typical Algorithms of Classification 分类的典型算法

☐ AdaBoost AdaBoost

☐ Decision tree 决策树

☐ Artificial neural networks 人工神经网络

☐ Bayesian networks 贝叶斯网络

☐ Hidden Markov models 隐马可夫模型

☐ K-nearest neighbors K-近邻

☐ Kernel method 核方法

☐ Linear discriminant analysis 线性判别分析

☐ Naive Bayes classifier 朴素贝叶斯分类器

☐ Softmax Softmax

☐ Support vector machine (SVM) 支撑向量机 (SVM)

Thank you for your attention!

AI