Pattern Recognition

(模式识别)

Instructor

Min-Ling Zhang (张敏灵)

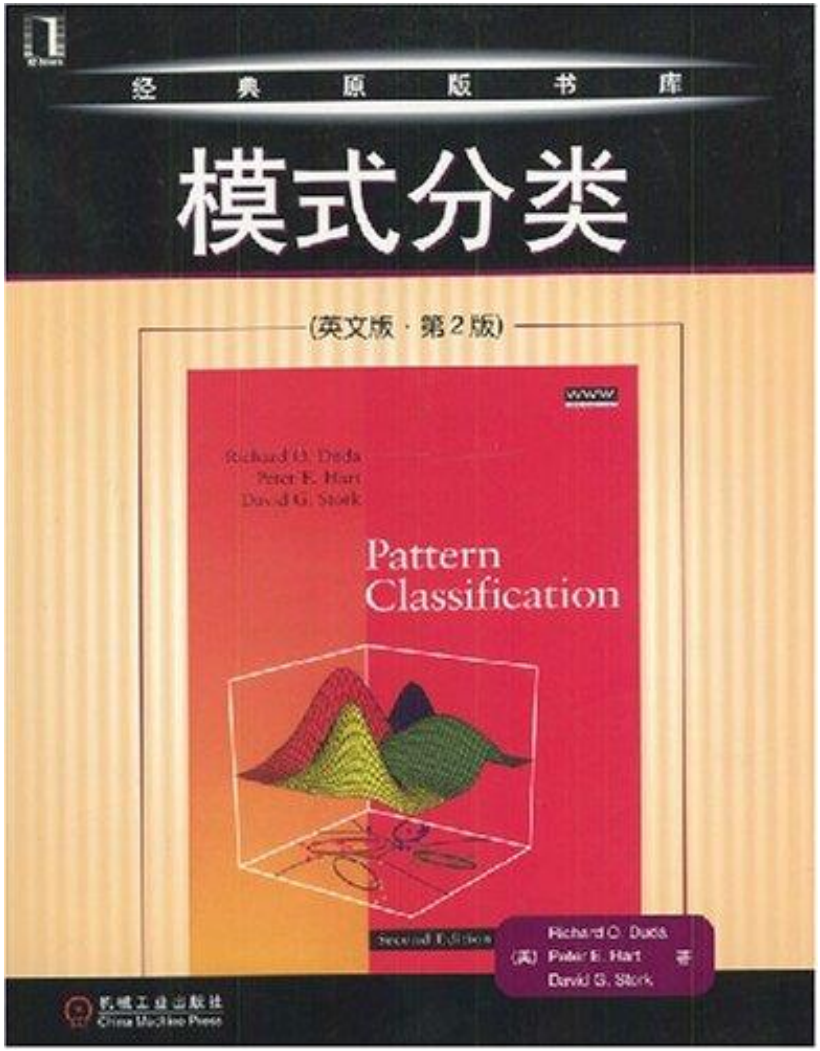
Email: zhangml@seu.edu.cn

URL: http://palm.seu.edu.cn/zhangml/

Southeast University

Spring Semester

Textbook



Richard O. Duda, Peter E. Hart, David G. Stork

Pattern Classification, 2nd edition

John Wiley & Sons, 2001

模式分类

(英文版·第2版)

机械工业出版社, 2004

Course Information

- Credits
  - 3 credits with 48 course hours and 16 project hours
  - Week 1 – Week 16, Wednesday & Thursday
- Contents
  - Chapters 1-6
- About scores
  - In class (attendance & assignments): 10%
  - Quiz: 10%
  - Project: 30%
  - Final Exam: 50%

References

- Books
  - S. Theodoridis, K. Koutroumbas. Pattern Recognition, 4th edition. Elsevier Publishers, 2009.
  - C. Bishop. Pattern Recognition and Machine Learning. Cambridge University Press, 2007.
  - 张学工. 模式识别(第三版). 清华大学出版社, 2010.
- Web Resources
  - International Association for Pattern Recognition (IAPR)
  - Pattern Recognition Journal (PRJ)
  - List of pattern recognition web sites

http://palm.seu.edu.cn/zhangml/

Home
Course
Publication
Activities
Resources

张敏灵

Min-Ling ZHANG

Professor

School of Computer Science and Engineering,
Key Laboratory of Computer Network and Information Integration of Ministry of Education (MOE),
Southeast University, China

Correspondence:

Mail: Min-Ling Zhang
School of Computer Science and Engineering
Southeast University
2 Sipailou
Nanjing 210096, China

Office: 528, Computer Building, Jiulonghu Campus of Southeast University
Tel: +86-25-5209-0869
Fax: +86-25-5209-0869
URL: http://cse.seu.edu.cn/PersonalPage/zhangml/
Email: zhangml@seu.edu.cn OR zhangml.seu@gmail.com

My research interests mainly include *machine learning* and *data mining*. Currently, I am a professor at the School of Computer Science and Engineering, Southeast University. Before joining So assistant professor (2007.10~2010.5) at the College of Computer and Information Engineering, Hohai University.

I received my B.Sc., M.Sc., and Ph.D. degrees in computer science all from Department of Computer Science & Technology, Nanjing University, China, in 2001, 2004 and 2007 respectively. I was led by my supervisor Prof. Zhi-Hua Zhou.

For related information and resources, please navigate via the links in the left bar. Contact me if you have any problems there.

http://palm.seu.edu.cn/zhangml/

Home

Course

Publication

Activities

Resources

NOTE: Refresh to find latest contents

Name: 模式识别 (全英文)

To: Undergraduate students

Semester: Spring 2021

Book: Richard O. Duda, Peter E. Hart, David G. Stork. **Pattern Classification**, John Wiley & Sons, 2001.
(Richard O. Duda, Peter E. Hart, David G. Stork. **模式分类[英文版·第2版]**, 机械工业出版社[影印版], 2004.)

Reference Books:
[1] S. Theodoridis, K. Koutroumbas. **Pattern Recognition, 4th edition**. Elsevier Publishers, 2009.
[2] C. Bishop. **Pattern Recognition and Machine Learning**. Cambridge University Press, 2007.
[3] 张学工. **模式识别(第三版)**. 清华大学出版社, 2010.

Credit: 3

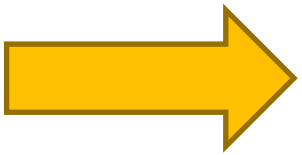
Slides (PDF): Ch1, Ch2, Ch3, Ch4, Ch5, Ch6

Homework I Homwork II Homework III Homework IV Homework V

Remarks

- Mathematical background
  - **Linear algebra**
  - **Probability theory**
  - Statistics
  - Information theory

Our course isn't a mathematical one

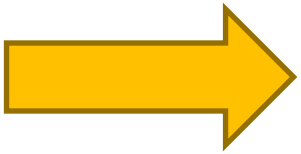


Carefully read and comprehend materials in Appendix "Mathematical Foundation"

Remarks (Cont.)

- **No pain, No gain**

Classroom lectures are important but not enough



Review what have been taught with at least 4~6 hours per week

- **Terminologies and Contents**

Important and difficult ones will be annotated and even revisited with Chinese

Only for reference purpose

Chapter 1

Introduction

The 3W of Pattern Recognition

- **W**hat is Pattern Recognition (PR)?

  What is **Pattern**?

  What is **Recognition**?

  What is **Pattern Recognition**?

- **W**hy do we need Pattern Recognition?

  The **necessity and importance** for pattern recognition

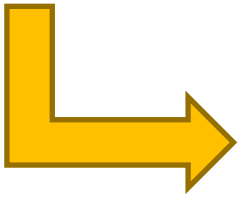
- Ho**W** to perform Pattern Recognition?

  The **building blocks** of a pattern recognition system

What is Pattern?

"To understand is to perceive patterns"

– Isaiah Berlin



Patterns are essential for human perception and understanding

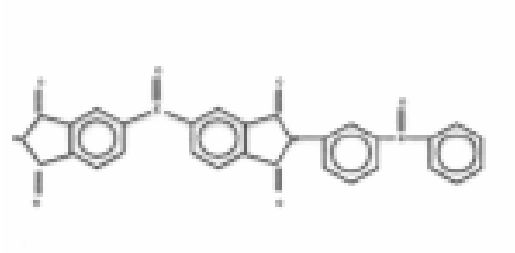
"A pattern is the opposite of a chaos; it is an entity vaguely defined, that could be given a name."

- Satoshi Watanabe

"模式(Pattern)是混沌(Chaos)的对立面，它是一个可赋予名字、无确切定义的实体"

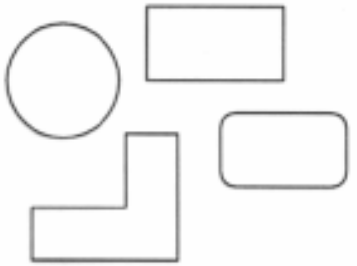
What is Pattern? (Cont.)

Some examples……



Molecules

Barcode



Geometry

Fingerprint

Footprint

Myself……

What is Pattern? (Cont.)

"A pattern is the opposite of a chaos; it is an entity vaguely defined, that could be given a name."

- There are various kinds of patterns
  - Visual patterns (视觉模式) such as eyes, nose, mouth, face, fingerprint, etc.
  - Temporal patterns (时序模式) such as speech, audios, videos, data streams, etc.
  - Logical patterns (逻辑模式) such as characters, strings, images, etc.
  - ……

What is Recognition?

"Identification of a pattern as a member of a category we already know, or we are familiar with"

"识别(Recognition)是将模式鉴定(Identification)为我们已知或者熟悉的类别(Category)的成员"

Two types of recognition

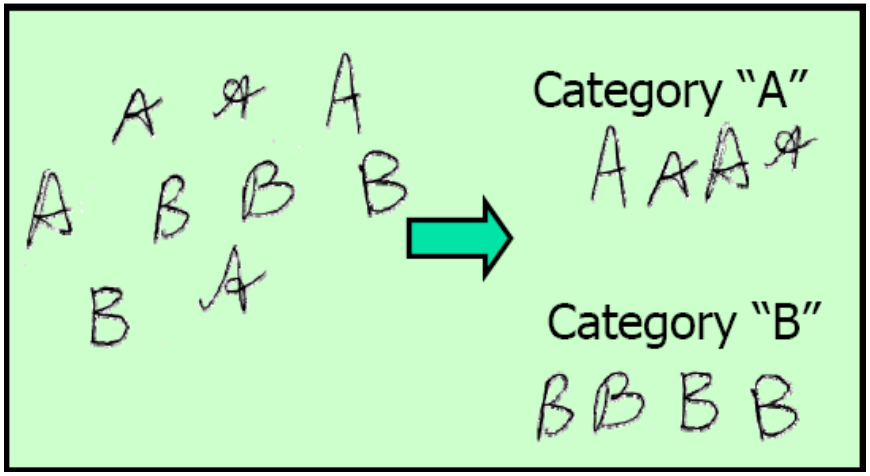
Classification (分类) → Categories are known and the task is to assign a proper class label for each pattern

Clustering (聚类) → Categories are unknown and the task is to learn categories and group the patterns accordingly

Classification vs. Clustering

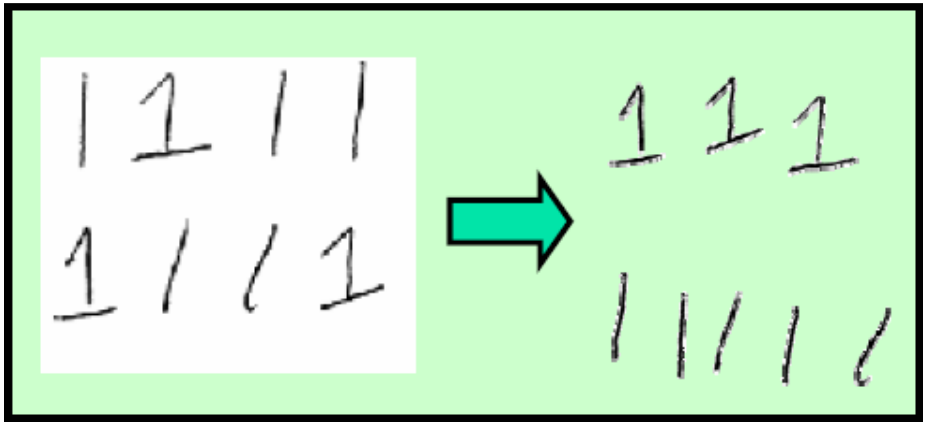
Classification: An example

We already know the categories of characters, and then classify the handwritten ones into category "A" and category "B"



Clustering: An example

We do not know the categories of symbols, and then learn the categories and group the symbols accordingly



What is Pattern Recognition?

Pattern recognition is the procedure of processing and analyzing diverse information (numerical, literal, logical) characterizing the objects or phenomenon, so as to provide descriptions, identifications, classifications and interpretations for them.

对表征事物或现象的各种形式的（数值的，文字的和逻辑关系的）信息进行处理和分析，从而对事物或现象进行描述、辨认、分类和解释的过程。(信息科学和人工智能的重要组成部分)

What is Pattern Recognition? (Cont.)

A "Perceive + Process + Prediction" View

It is the study of how machines can

- ✓ **Perceive**: Observe the environment (i.e. interact with the real-world)
- ✓ **Process**: Learn to distinguish patterns of interest from their background
- ✓ **Prediction**: make sound and reasonable decisions about the categories of the patterns

Why need pattern recognition?

"The real power of human thinking is based on recognizing patterns. The better computers get at pattern recognition, the more humanlike they will become."

– *Ray Kurzweil @ New York Times, 2003*

"The problem of searching for patterns in data is a fundamental one and has a long and successful history."

– *Christophe M. Bishop*

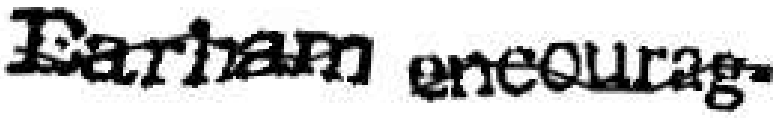
Pattern recognition is needed in designing almost all automated and intelligent systems!

Applications of Pattern Recognition

1) Character Recognition

[字符识别]

Input:
images with characters (normally contaminated with noise)



Output:
the identified character strings (Earham encourag)

Useful in scenarios such as automatic license plate recognition (ALPR), optical character recognition (OCR), etc.

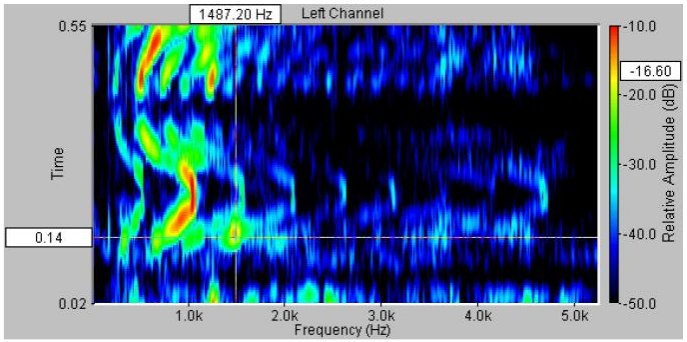
Applications of PR (Cont.)

2) Speech Recognition

[语音识别]

Input:

acoustic signal
(e.g. sound waves)



Output:

contents of the speech

Useful in scenarios such as speech-to-text (STT), voice command & control, etc.

Applications of PR (Cont.)

3) Fingerprint Recognition

[指纹识别]

Input:
fingerprints of some person

Output:

the person's identity

Useful in scenarios such as computerized access control, criminal pursuit, etc.

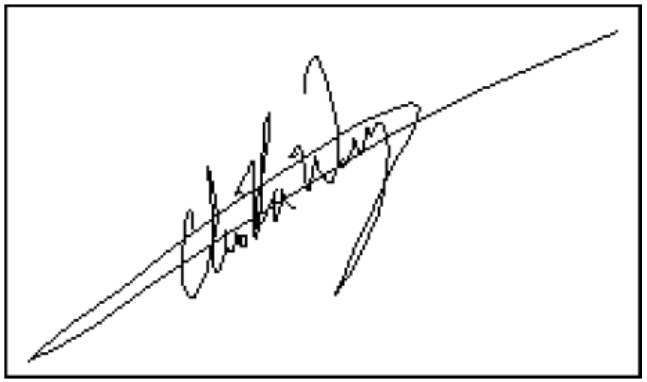
Applications of PR (Cont.)

4) Signature Identification

[签名验证]

Input:

signature of some person
(sequence of dots)



Output:

the signatory's identity

Useful in scenarios such as
digital signature verification,
credit card anti-fraud, etc.

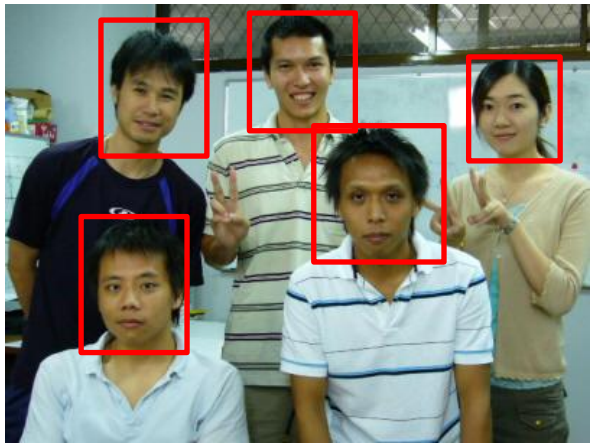
Applications of PR (Cont.)

5) Face Detection
[人脸检测]

Useful in scenarios such as digital camera capturing, video surveillance, etc.

Input:
images with several people

Output:
locations of the peoples' faces in the image



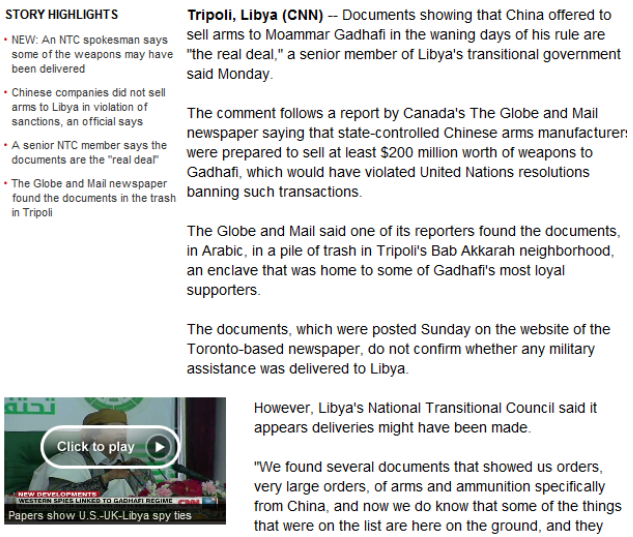
Applications of PR (Cont.)

6) Text Categorization

[文档分类]

Input:

document, web pages, etc.



STORY HIGHLIGHTS

- NEW: An NTC spokesman says some of the weapons may have been delivered
- Chinese companies did not sell arms to Libya in violation of sanctions, an official says
- A senior NTC member says the documents are the "real deal"
- The Globe and Mail newspaper found the documents in the trash in Tripoli

Tripoli, Libya (CNN) -- Documents showing that China offered to sell arms to Moammar Gadhafi in the waning days of his rule are "the real deal," a senior member of Libya's transitional government said Monday.

The comment follows a report by Canada's The Globe and Mail newspaper saying that state-controlled Chinese arms manufacturers were prepared to sell at least $200 million worth of weapons to Gadhafi, which would have violated United Nations resolutions banning such transactions.

The Globe and Mail said one of its reporters found the documents, in Arabic, in a pile of trash in Tripoli's Bab Akkarah neighborhood, an enclave that was home to some of Gadhafi's most loyal supporters.

The documents, which were posted Sunday on the website of the Toronto-based newspaper, do not confirm whether any military assistance was delivered to Libya.

However, Libya's National Transitional Council said it appears deliveries might have been made.

"We found several documents that showed us orders, very large orders, of arms and ammunition specifically from China, and now we do know that some of the things that were on the list are here on the ground, and they

Output:

category of the text, such as political, economic, military, sports, etc.

Useful in scenarios such as information retrieval, document organization, etc.

Applications of PR - More

Problem	Input	Output
Detection and diagnosis of disease	Electrocardiogram (ECG) waveforms, Electroencephalogram (EEG) waveforms	Types of cardiac conditions, classes of brain conditions
Natural resource identification	Multispectral images	Terrain forms, vegetation cover
Aerial reconnaissance	Visual, infrared, radar images	Tanks, airfields
Identification and counting of cells	Slides of blood samples, micro-sections of tissues	Type of cells
Inspection (PC boards, IC masks, textiles)	Scanned image (visible, infrared)	Acceptable/unacceptable
Manufacturing	3-D images (structured light, laser, stereo)	Identify objects, pose, assembly
Web search	Key words specified by a user	Text relevant to the user
……	……	……

Why need pattern recognition? (Cont.)

For humans, pattern recognition is natural & easy

recognize a face understand spoken words

read handwritten characters identify items by feel

decide whether an apple is ripe by its smell ……

For computers, pattern recognition is never easy

All in all, pattern recognition is important, useful, attractive, but rather challenging

Challenges ➔ Opportunities

Basic Concepts

Model (模型)

Descriptions which are typically mathematical in form

[以数学形式表达的性质]

e.g. image ➔ matrix; sound waves ➔ frequency vector

Sample (样本)

Representatives of the patterns we want to classify

[分类的基本对象，模式的实例]

e.g. fingerprint of a suspect; ECG of a patient

Training Set (训练集)

A set of samples used to train classifiers

[用于训练分类器的样本集合]

Basic Concepts (Cont.)

Test Set (测试集)

A set of samples to be classified, usually being mutually exclusive to training set

[用于测试分类器的样本集合,通常与训练集无交集]

"Training set" vs. "Test set" ⇔ *"Homeworks" vs. "Exams"*

Feature (特征)

Attributes which characterize properties of the samples

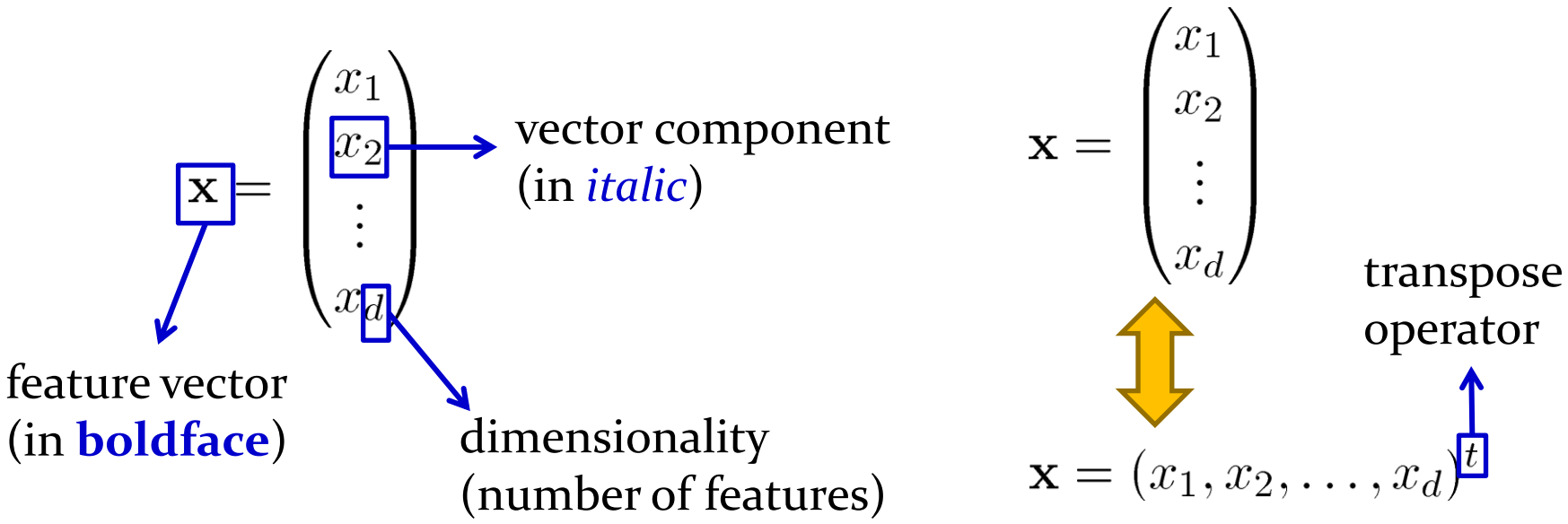
[用于刻画样本性质的属性]

e.g. to characterize a person, we may use features such as height, weight, age, salary, occupation, etc.

Basic Concepts (Cont.)

Feature Vector (特征向量)

Vector formed by a group of features, usually in column form
[由一组特征组成的向量，通常表示为列向量]



Basic Concepts (Cont.)

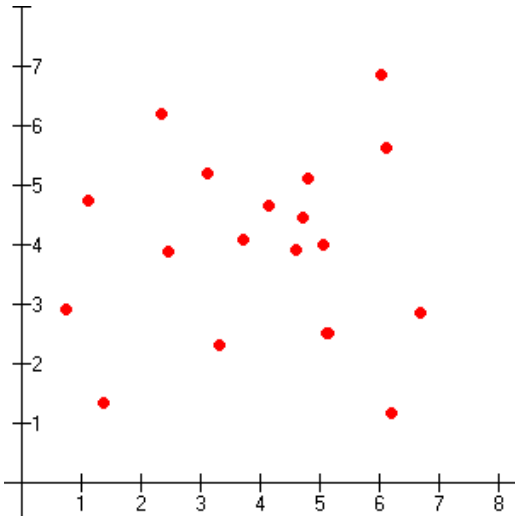
Feature Space (特征空间)

Space containing all the possible feature vectors
(由所有可能的特征向量组成的数据空间)

e.g. the d-dimensional Euclidean space $\mathbf{R}^d$

Scatter Plot (散布图)

Each sample is plotted as a point in the feature space
(将每个样本表示为特征空间中的一个点)

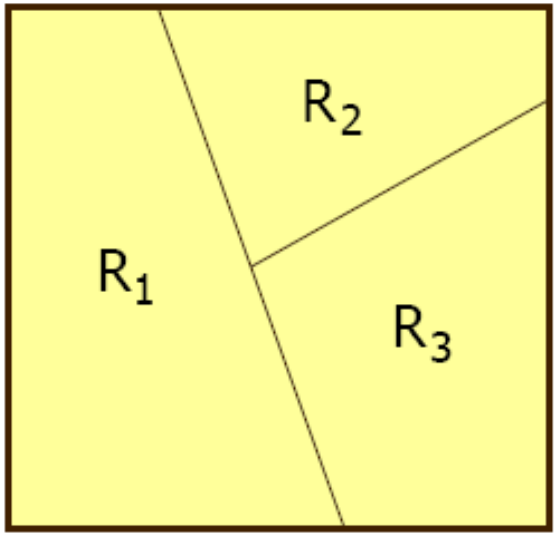


a 2D scatter plot

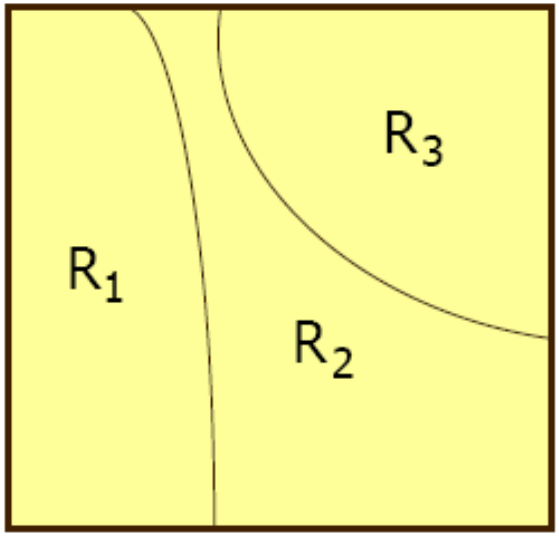
Basic Concepts (Cont.)

Decision Boundary (决策边界)

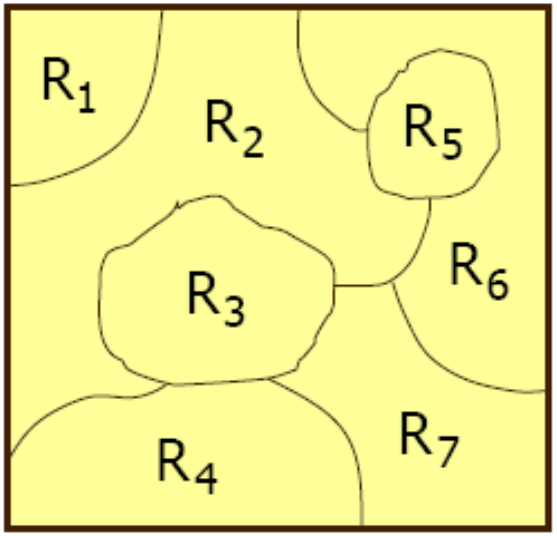
Boundaries in feature space which separate different categories (特征空间中区分各个类别的边界)



linear boundary



quadratic boundary



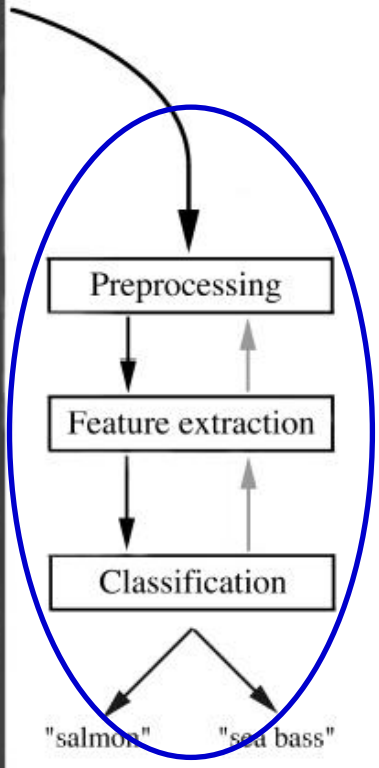
complex boundary

How to do pattern recognition?

An Example

The task: Automate the process of sorting incoming fish on a conveyor belt according to species

Separate *sea bass* from *salmon*
[鲈鱼 vs. 鲑鱼]



Three basic steps

Example: "Sea bass" vs. "Salmon" (Cont.)

Step I: Preprocessing (预处理)

Goal: Preprocess the image captured by the camera, such that subsequent operations could be simplified without losing relevant information

Routine image processing

- Adjust the level of illumination
- Denoising
- Enhance the level of contrast

segmentation

- Isolate different fishes from one another
- Isolate fishes from the background

.....

Example: "Sea bass" vs. "Salmon" (Cont.)

Step II: Feature Extraction (特征抽取)

Goal: Extract features (with good distinguishing ability) from the preprocessed image to be used for subsequent classification

Sea bass is usually longer than a salmon → "length" could be a good candidate for features

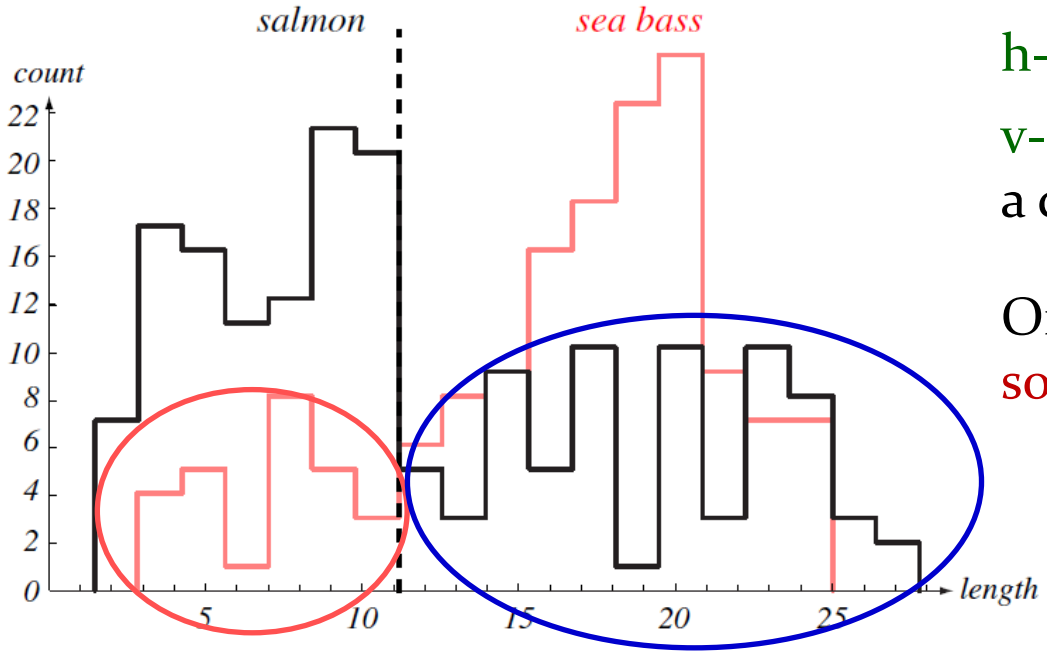
Sea bass is usually brighter than a salmon → "lightness" of *fish scales* could be another good candidate for features

……

Example: "Sea bass" vs. "Salmon" (Cont.)

Step III: Classification (分类)

Goal: To distinguish different types of objects (in this case, *sea bass* vs. *salmon*) based on the extracted features



histogram for *length*

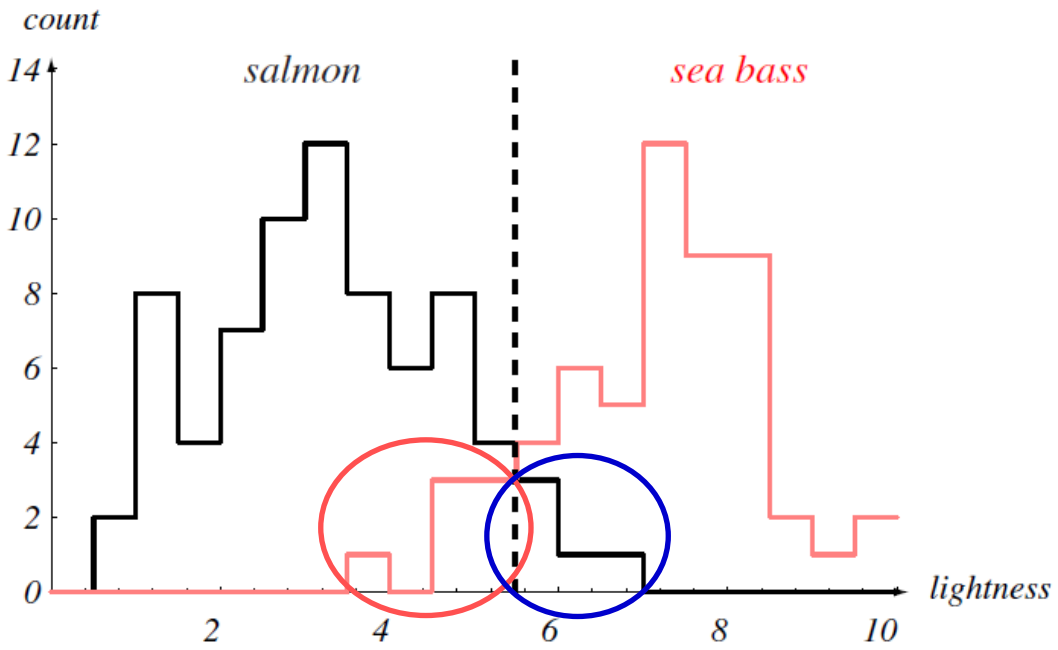
h-axis: length of fish

v-axis: number of fishes with a certain length

On average, sea bass is somewhat longer than salmon

Too much overlaps ➔ poor separation with the length feature

Example: "Sea bass" vs. "Salmon" (Cont.)



histogram for *lightness*

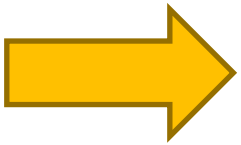
h-axis: lightness of fish scales

v-axis: number of fishes with a certain lightness

On average, sea bass is much brighter than salmon

Less overlaps ➔ better separation with the lightness feature, but still a bit unsatisfactory

What if no other single feature yields better performance?



Use more features at the same time!

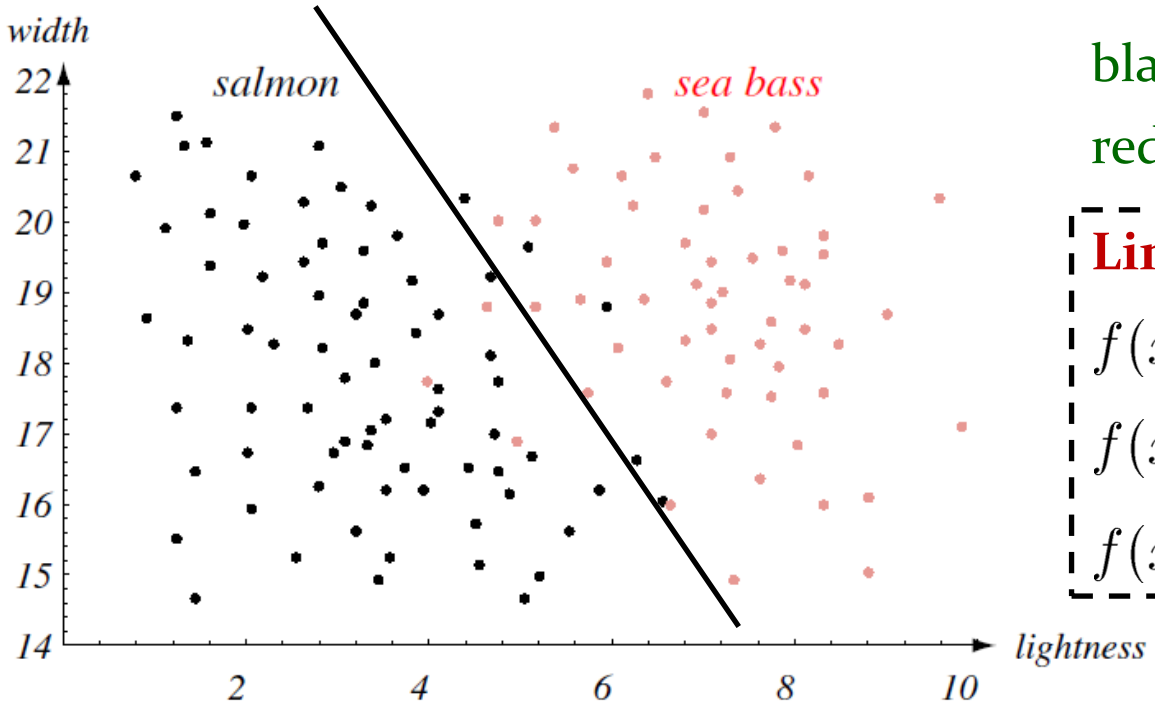
Example: "Sea bass" vs. "Salmon" (Cont.)

Using two features simultaneously

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

x_1 : fish width

x_2 : fish lightness



scatter plot for the feature vectors

black dots: salmon samples

red dots: sea bass samples

Linear decision boundary:

$$f(x_1, x_2) = a \cdot x_1 + b \cdot x_2 + c$$

$$f(x_1, x_2) > 0 \implies \text{sea bass}$$

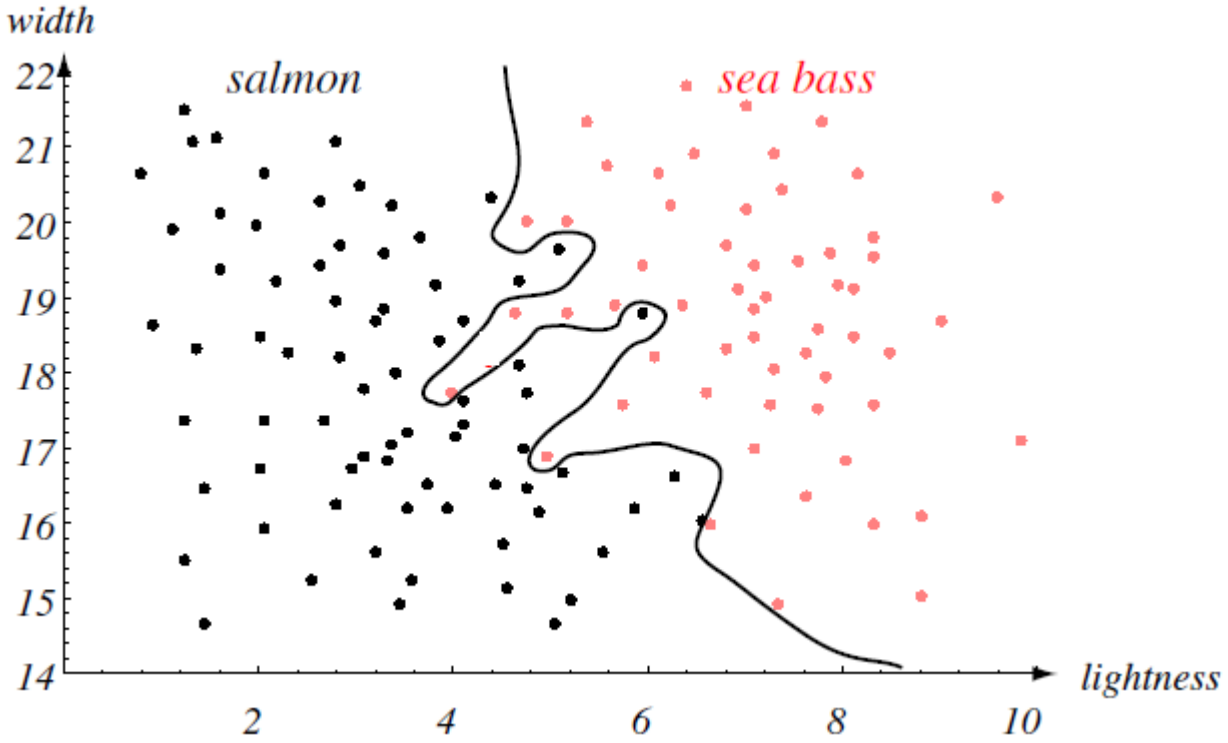
$$f(x_1, x_2) \leq 0 \implies \text{salmon}$$

Much better than single feature

Example: "Sea bass" vs. "Salmon" (Cont.)

Linear decision boundary:

Complex decision boundary



All the training samples (i.e. known patterns) have been separated perfectly

Can we truly feel satisfied?

Example: "Sea bass" vs. "Salmon" (Cont.)

Generalization

[泛化能力/推广能力]

The ultimate goal!

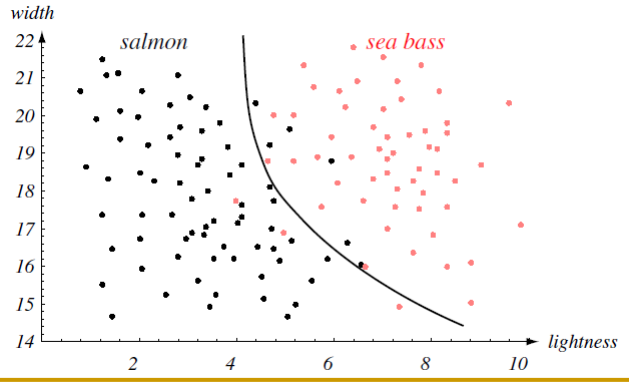
The central aim of designing a classifier is to make correct decisions when presented with *novel* (*unseen*/*test*) patterns, not on training patterns whose labels are already known

e.g. it's useless to get 100% accuracy when answering homework questions while get low accuracy when answering exam questions

Performance on the training set

Simplicity of the classifier

→ Tradeoff



Related Fields to PR

Pattern Recognition: **Pattern ➔ Category**

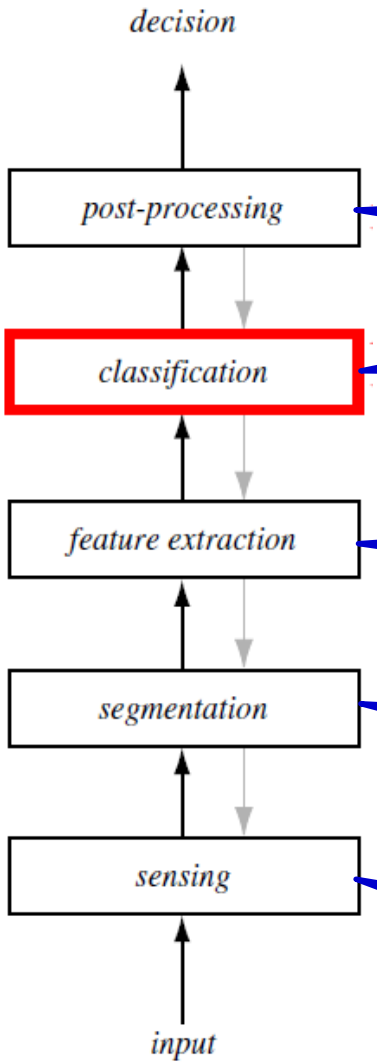
- Hypothesis Testing (假设检验)
  - Null hypothesis ➔ Rejection or Not [ref. pp.628]

  E.g.: To determine whether a drug is effective; Null hypothesis: it has not effect
- Image Processing (图像处理)
  - Image ➔ Image
- Associative Memory (联想存储器)
  - Pattern ➔ Pattern
- Regression (回归分析)
  - Pattern ➔ Real Value
- Interpolation (插值)
  - Pattern (unexplored input range) ➔ Interpolated Value
- Density Estimation (概率密度估计)
  - Patterns ➔ Probability density function (pdf) for different categories

Often employed as preliminary steps in pattern recognition

Pattern Recognition System

In addition to the usual "bottom-up" flow of data, some systems also employ feedback from higher levels back down to lower levels (gray arrows)



A post processor decide on the appropriate action based on the classification

A classifier uses extracted features to assign the sensed object to a category

A feature extractor measures object properties that are useful for classification

A segmentor isolates sensed objects from the background or from other object

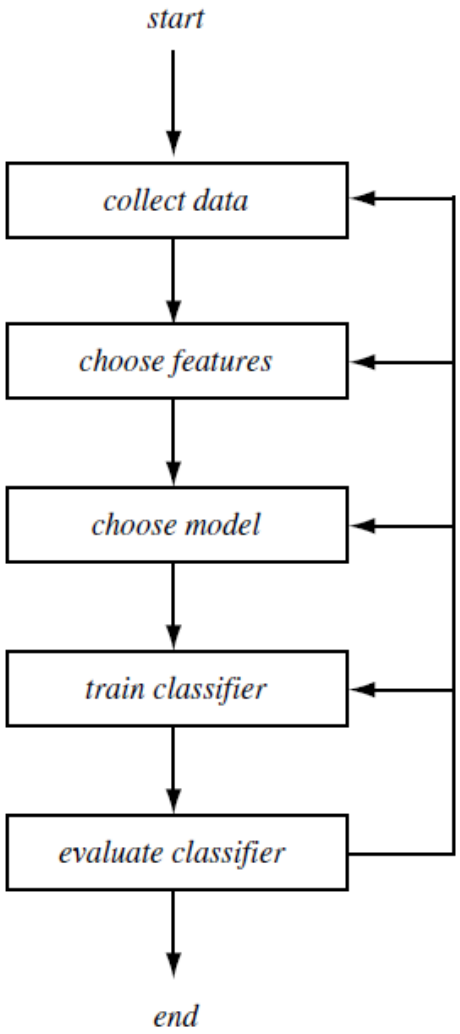
A sensor converts physical inputs (e.g. images, sounds) into digital signal data

Design Cycle of PR System

The design of a PR system usually entails a number of different activities, such as *data collection, feature choice, model choice, classifier training, classifier evaluation.*

- Data collection accounts for a large part of the cost of developing a PR system
- Feature choice and model choice are highly domain-dependent, where *prior knowledge* (先验知识) plays very important role

  e.g.: lightness might be a good feature for distinguishing sea bass and salmon; linear model might be preferred than nonlinear ones
- Various activities may be repeated in order to obtain satisfactory results



Important Issues in Pattern Recognition

- Noise
  (噪声)
- Segmentation
  (分割)
- Data Collection
  (数据采集)
- Domain Knowledge
  (领域知识)
- Feature Extraction
  (特征抽取)
- Pattern Representation
  (模式表示)
- Missing Features
  (特征缺失)
- Model Selection
  (模型选择)
- Overfitting
  (过配)
- Context
  (上下文)
- Classifier Ensemble
  (分类器集成)
- Costs and Risks
  (代价与风险)
- Computational Complexity
  (计算复杂度)
- ……

Noise

- General definition
  - Any property of the sensed pattern which is not due to the true underlying model but instead to intrinsic randomness of the world or the sensors
- Various types of noise exist
  - shadows, conveyor belt might shake, etc.
- Noise can reduce the reliability of the feature values measured
- Knowledge of the noise process can help improve performance

Segmentation

- Individual patterns have to be segmented for subsequent pattern recognition operations
- One of the deepest as well as hardest problems in pattern recognition
  - How can we segment the images without having categorized them firstly?
  - On the other hand, how can we categorize the images without having segmented them firstly?
- How do we "group" together the proper number of elements
  - **BEATS** ➔ BE, BEAT, EAT, AT, EATS?

Data Collection

- A small set of "typical" examples ➔ Preliminary study of system feasibility
- Much more data ➔ Assure good performance in the fielded system
- How do we know that we have collected:
  - Adequately large set of examples for training and testing the system?
  - Representative set of examples for training and testing the system?
- The efforts of data collection could be rather demanding

Domain Knowledge

- There is not sufficient data for training ➔ Incorporate domain knowledge (a.k.a. prior knowledge)
- **Type I:** Incorporate domain knowledge on the patterns themselves – Difficult!
  - To recognize all types of chairs
  - Astounding variety in *number of legs*, *material*, *shape*, and *so on* ➔ What is the commonness for chairs which could be regarded as domain knowledge?
- **Type II:** Incorporate domain knowledge on the pattern generation procedure
  - Optical character recognition ➔ Assume handwritten characters are written as a sequence of strokes
  - First try to recover stroke representations ➔ deduce the character from the identified strokes

Feature Extraction

- A domain-dependent problem which influences the classifier's performance
  - Good extracted features ➔ Make classification easier
- What kinds of features are promising?
  - **Distinguishing Capability:** Whose values are very similar for objects in the same category, while very different for objects in different categories
- What if a large set of candidate features available?
  - Choose those are simple to extract
  - Choose those are robust to noise
  - Choose those can lead to simpler decision boundaries
  - ……

Pattern Representation

- Various ways for pattern representation
  - **Statistical:** *feature vector* (the most popular)
  - **Template Matching:** *prototype templates*
  - **Syntactic:** *rules or grammars*
- Desired Properties
  - Patterns from the same classes should have similar representations
  - Patterns from different classes should have dissimilar representations
  - Pattern representations should be invariant to transformations such as *translations, rotations, resizes, reflections, non-rigid deformations*
  - Intra-class variation should be small
  - Inter-class variation should be large
  - ……

Missing Features

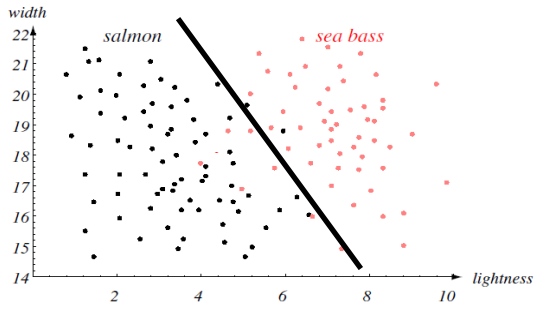
- In practical problems, values for certain features may be missing
  - Occlusion between fishes ➔ fish width can't be measured
- How could we train classifiers with missing features?
  - Naïve method could be used, but may not be optimal
    - Assuming the value of missing features is zero
    - Assigning the average value of patterns already seen for the missing feature
  - Sophisticated method might be better, but requires extra efforts in terms of storage and time
    - Fill in the missing values with regression techniques

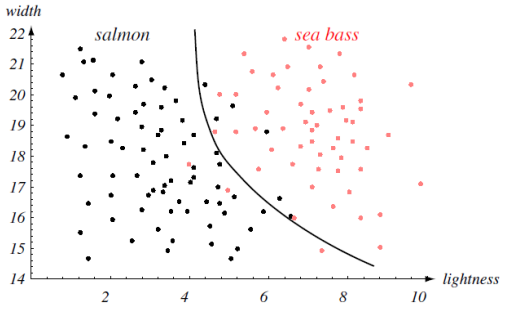
Model Selection

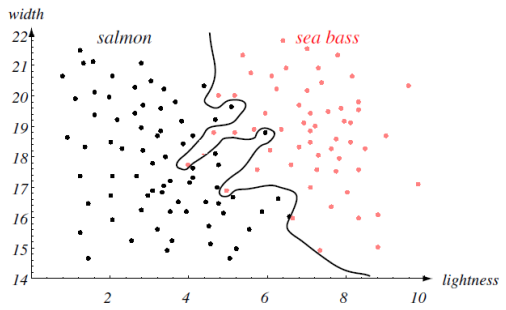
- Each pattern recognition method employs certain *model hypothesis*
- Every pattern recognition problem has its own *underlying true model*
- Fundamental questions on model selection
  - How do we know whether the hypothesized model is (relatively) consistent with the underlying true model?
  - How are we to know to reject a class of models and try another one?
  - Can we automate the process of model selection, instead of trial and error (试错) which is random and tedious?

Overfitting

- We can get perfect classification performance on the training data by choosing complex models
  - Complex models are tuned to the particular training samples, rather than the characteristics of the true model
- Models overly complex than necessary lead to overfitting
  - Good performance on the training data, but poor performance on novel data
- How can we find principled ways to obtain best complexity?







Context

- **Context:** Input-dependent information, other than from the pattern itself
  - context of language, context of videos, etc.
- The same pattern within different context might have different meanings
  - Use the context of a conversation to infer the meaning of the speaker
- Context is very helpful!

How m ch info mation
are y u mi sing

Classifier Ensemble

- Classifier ensemble aims to improve generalization performance by employing a number of classifiers for the same task
  - To improve the performance of speech recognizer: combine the results of *acoustic recognition* and *lip reading*
  - a.k.a. *Multi-classifier System*, *Mixture of Experts*, *Classifier Fusion*, etc.
  - Diverse ensemble techniques: *Bagging*, *Boosting*, *Random subspace*, etc. [ref. pp.475]
- How to combine different classifiers?
  - **Majority voting:** vote for the category where most classifiers agree
  - **Weighted voting:** weight each vote by classifier's confidence
  - **Stacking:** learn the rule of combination (more complicated)

Costs and Risks

- Cost is the loss after making incorrect decisions
  - **Equal cost:** In OCR, the cost of mistaking "6" as "9" might be equal to that of mistaking "9" as "6"
  - **Unequal cost:** In AIDS diagnosis, the cost of mistaking "positive (阳性)" as "negative (阴性)" would be much higher than that of mistaking "negative" as "positive"
- Risk is total expected cost which we want to optimize
  - Error rate (percentages of test patterns being wrongly classified)
  - Precision, Recall, Area under the ROC curve (AUC), etc.
- Questions on costs and risks
  - How do we incorporate knowledge of costs, e.g. unequal cost?
  - Can we estimate the *lowest* possible risk of any classifier?
  - ……

Computational Complexity

- How does an algorithm scale with
  - The number of features (dimensionality)
  - The number of training patterns
  - The number of possible categories
- Brute force (蛮力) approaches might lead to perfect classification, but with impractical time and storage requirements
  - In OCR, label all possible 20x20 binary pixel images with a category ➔ use simple table lookup (查表) to classify incoming patterns
  - Labeling each of the 2^{20x20} ($\approx 10^{120}$) patterns is prohibitive
- How can we find a good tradeoff between computational ease and classifier performance?

Summary

- What is Pattern Recognition?
  - Pattern
    - The opposite of chaos
    - Various kinds: visual patterns, temporal patterns, logical patterns, etc.
  - Recognition
    - Identification of a pattern as a member of a category
    - Classification: categories known ➔ assign proper class label for each pattern
    - Clustering: categories unknown ➔ learn categories and group patterns
  - Pattern Recognition
    - Perceive: observe the environment (i.e. interact with the real-world)
    - Process: learn to distinguish patterns of interest
    - Prediction: make sound and reasonable decisions about the categories

Summary (Cont.)

- Why Pattern Recognition?
  - Pattern recognition is needed in designing almost all automated and intelligent systems
  - Applications of pattern recognition are ubiquitous
    - Character recognition (images ➔ characters)
    - Speech recognition (speech ➔ text)
    - Fingerprint recognition (fingerprints ➔ person's identity)
    - Signature identification (signature ➔ signatory's identity)
    - Face detection (images ➔ face locations)
    - Text categorization (documents ➔ semantic categories)
    - ……

Summary (Cont.)

- How Pattern Recognition?
  - Basic concepts
    - model, sample, training set, test set, feature, feature vector, feature space, scatter plot, decision boundary
    - An illustrative example: “sea bass” vs. “salmon”
    - **Generalization**: Make correct decisions given novel patterns
  - Related fields
    - hypothesis testing, image processing, associative memory, regression, interpolation, density estimation
  - Components of Pattern Recognition System
    - sensing ➔ segmentation ➔ feature extraction ➔ classification ➔ post-processing ➔ .....

Summary (Cont.)

- How Pattern Recognition?
  - Design Cycle of Pattern Recognition System
    - collect data ➔ choose features ➔ choose model ➔ train classifier ➔ evaluate classifier ➔ ……
  - Important Issues
    - Noise
    - Segmentation
    - Data Collection
    - Domain Knowledge
    - Feature Extraction
    - Pattern Representation
    - Missing Features
    - Model Selection
    - Overfitting
    - Context
    - Classifier Ensemble
    - Costs and Risks
    - Computational Complexity
    - ……