Prerequisites

- Linear Algebra (Please review A.2 of Textbook)
- Probability Theory (Please review A.4 of Textbook)

Pattern Recognition Definition

"Pattern recognition is the categorization of input data into identifiable classes via the extraction of significant features or attributes of the data from a background of irrelevant detail." … J. T. Tou and R. Gonzales.

Sample Applications of Pattern Recognition

- Remote Sensing
  - Environment monitoring
  - Exploration of other planets
  - Water and crop resource management
- Fingerprint Identification
- Character Recognition
- Speech Recognition
- Image Analysis
  - Object recognition
  - Flexible and adaptive industrial automation
  - Robotics
  - Autonomous vehicle guidance
- Signal Analysis
  - Radar and sonar
  - Seismic
  - Communications
- Multimedia
  - Recognition of objects, actors, words, or voices in video clips or movies (for automated searches through stored databases)
- Human-Computer Interface
  - Face and expression recognition
  - Recognition of objects in a scene (e.g., hand against background)
- Biomedical
  - Gene analysis
  - DNA blood typing
  - EKG, EEG, X-ray, CT, MRI

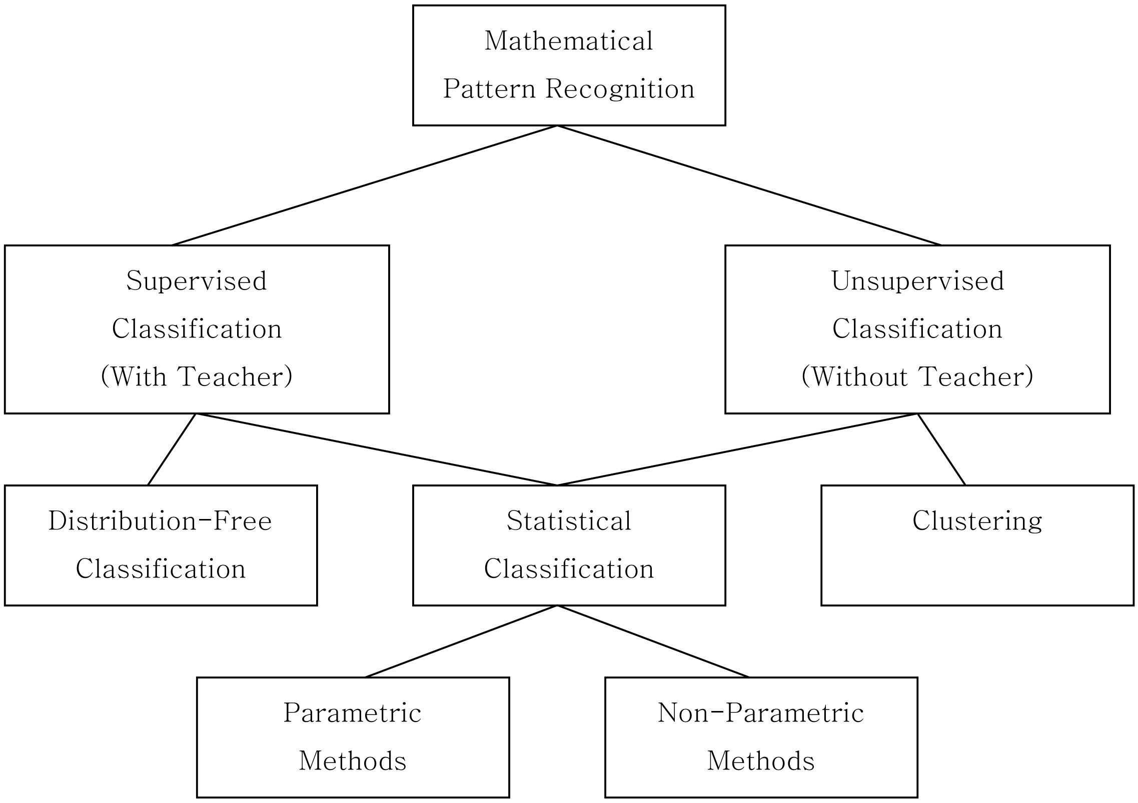
- Finance
  - Stock market
  - Economic analysis (economic indicators)
  - Banking (loan applications, signature verification)

Course Outline

1. Introduction (Ch. 1)
   A. Basic concepts in pattern recognition
   B. Approaches to pattern recognition
2. Distribution-Free Classification (Ch. 5)
   A. Classifier design – different techniques
      i. Discriminant functions
      ii. Linear vs. nonlinear
   B. Training algorithms
3. Statistical Classification (Ch. 3)
   A. Statistics are known
      i. Bayes decision theory
   B. Statistics are partially known or are assumed
      i. Parameter estimation
   C. Statistics are unknown
      i. Nonparametric techniques (Ch. 4)
   D. Supervised learning
4. Unsupervised Learning (Ch. 10)
   A. Statistical techniques (maximum likelihood)
   B. Criterion functions and iterative optimization
   C. Clustering techniques
   D. Component Analysis
5. Artificial Neural Network (Ch. 6)
   A. Single layer networks
   B. Multiple layer networks
   C. Supervised learning
   D. Unsupervised learning
   E. Capabilities and limitations
6. Deep Learning
   A. Convolutional Neural Networks
   B. Recurrent Neural Networks

C. Reinforcement Learning

Mathematical Pattern Recognition Topics



Mathematical Pattern Recognition: Fundamental Assumptions

1. There is a certain similarity of samples from a given class. A class is a collection of samples with something in common.
2. A transformation of the pattern space (feature extraction) will bunch or cluster elements within a given class and separate the different classes.
3. Pattern space contains sufficient information to permit classification
4. An appropriate distance measure is available
5. In the case of supervised pattern recognition, representative sets of samples (prototypes) are available (for training).

Patter Classification (Example)

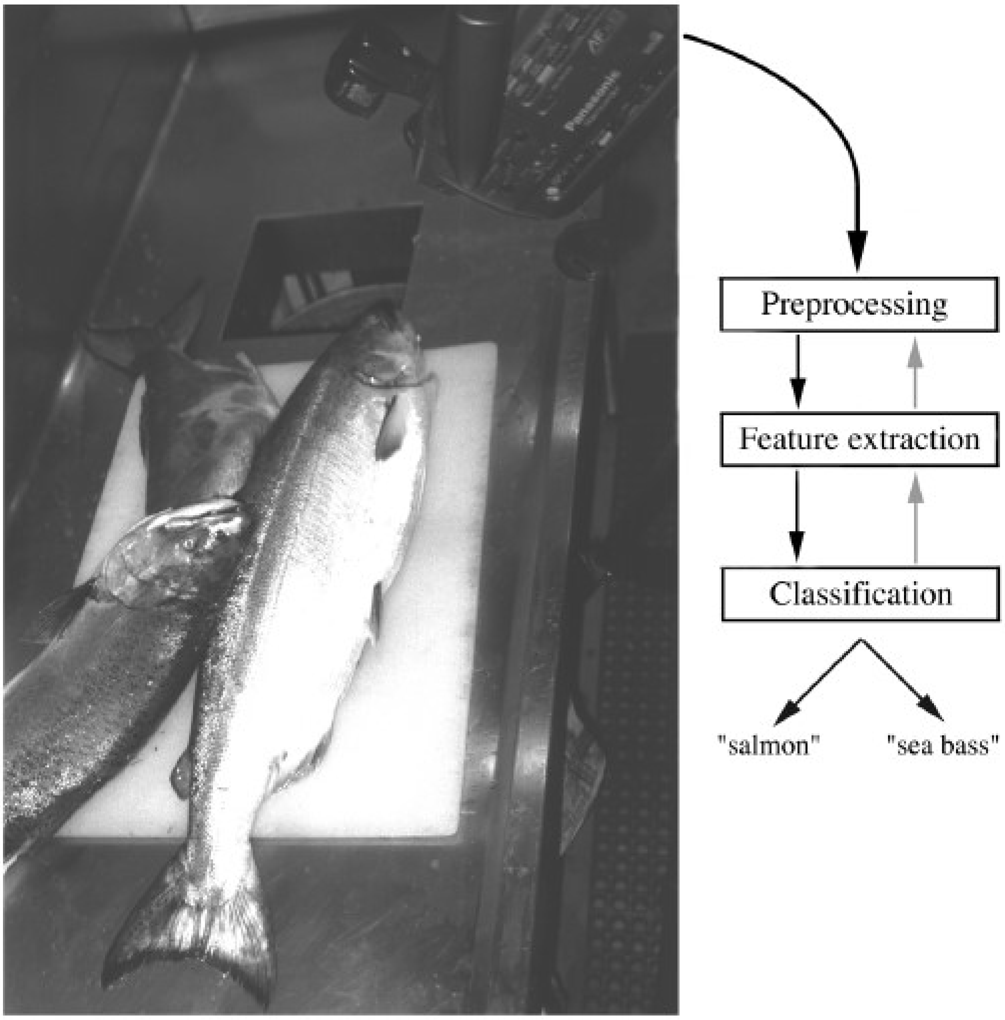


Figure 1.1: The objects to be classified are first sensed by a transducer (camera), whose signals are preprocessed, then the features extracted and finally the classification emitted (here either "salmon" or "sea bass"). Although the information flow is often chosen to be from the source to the classifier ("bottom-up"), some systems employ "top-down" flow as well, in which earlier levels of processing can be altered based on the tentative or preliminary response in later levels (gray arrows). Yet others combine two or more stages into a unified step, such as simultaneous segmentation and feature extraction.

Key Terminology in Pattern Classification

- Feature: lightness, width, number, shape, position, etc.
- Model: typically in meth forms
- Preprocessing
- Segmentation
- Feature Extraction
- Training Samples
- Cost: Fig. 1.2 and Fig. 1.3

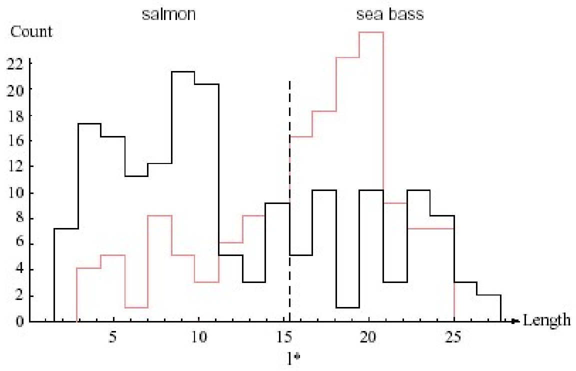


Figure 1.2: Histograms for the length feature for the two categories. No single threshold value l^* (decision boundary) will serve to unambiguously discriminate between the two categories; using length alone, we will have some errors. The value l^* marked will lead to the smallest number of errors, on average.

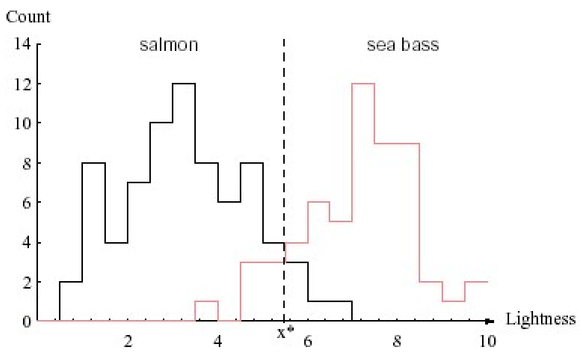
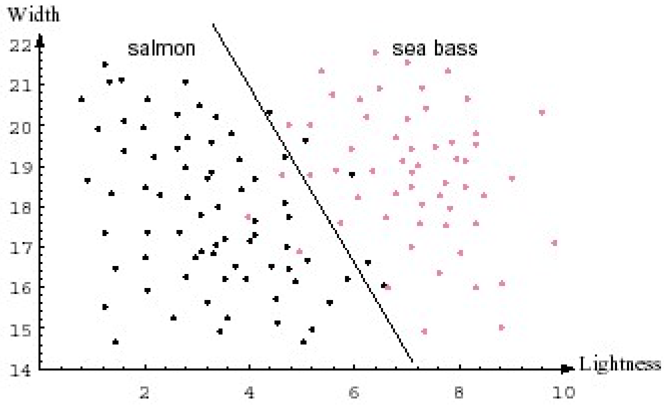
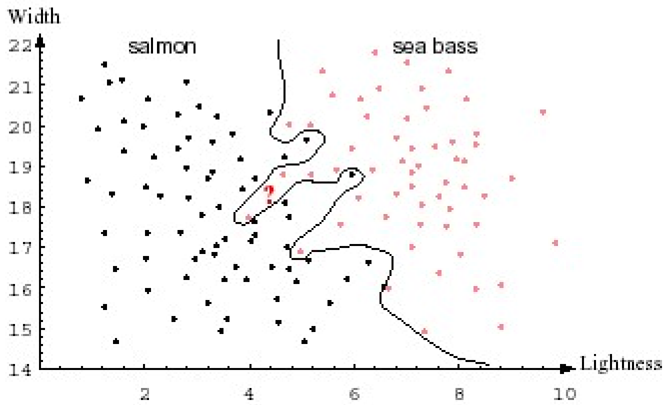
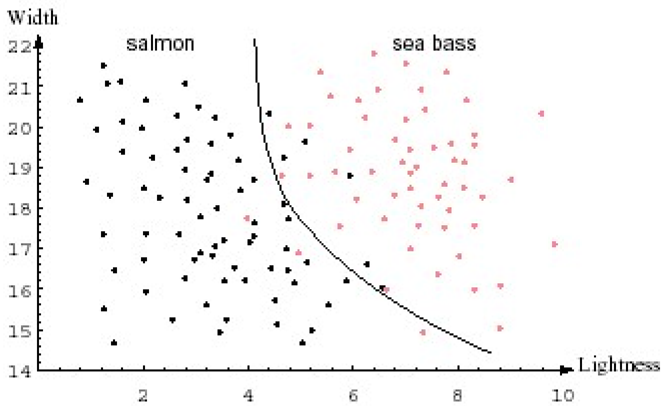


Figure 1.3: Histograms for the lightness feature for the two categories. No single threshold value x^* (decision boundary) will serve to unambiguously discriminate between the two categories; using lightness alone, we will have some errors. The value x^* marked will lead to the smallest number of errors, on average.

- Decision Theory
- Decision Boundary:







- Generalization

Components of Pattern Recognition Systems

- Fig. 1.7