Prerequisites

- Linear Algebra
- Probability Theory

Machine Learning (ML) Definition

"Machine learning is a subfield of artificial intelligence, which is broadly defined as the capability of a machine to imitate intelligent human behavior. Artificial intelligence systems are used to perform complex tasks in a way that is similar to how humans solve problems."

Pattern Recognition (PR) 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.

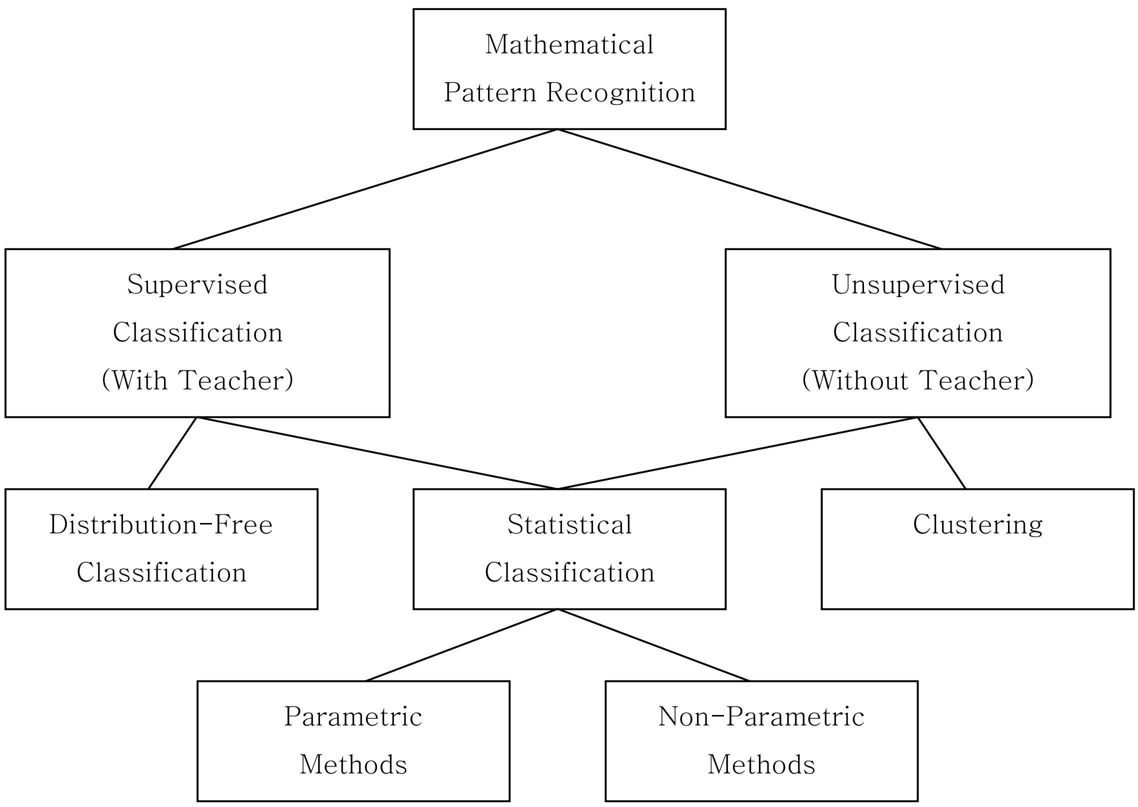
Sample Applications of Machine Learning/Pattern Recognition

- 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)

ML/PR Topics



Mathematical Machine Learning/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).

Machine Learning/Patter Recognition (Example)

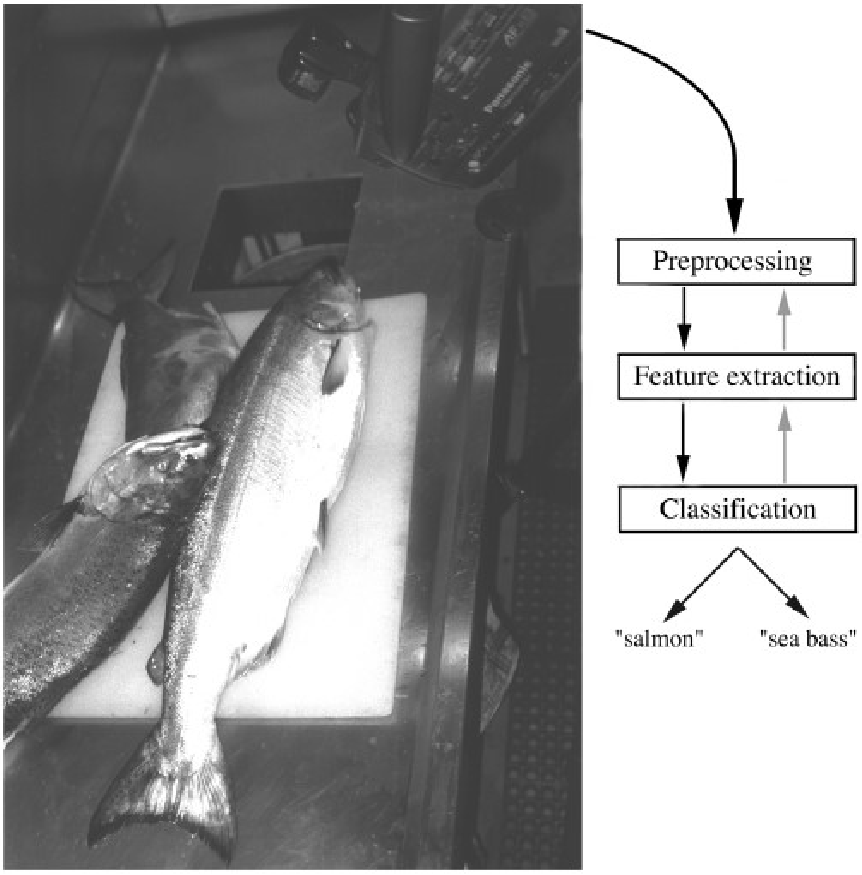


Figure 1.1: The objects to be classified are first sensed by a transducer (camera), whose signals are preprocessed, then the features extracted and finally the classification emitted (here either "salmon" or "sea bass"). Although the information flow is often chosen to be from the source to the classifier ("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 Machine Learning and 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

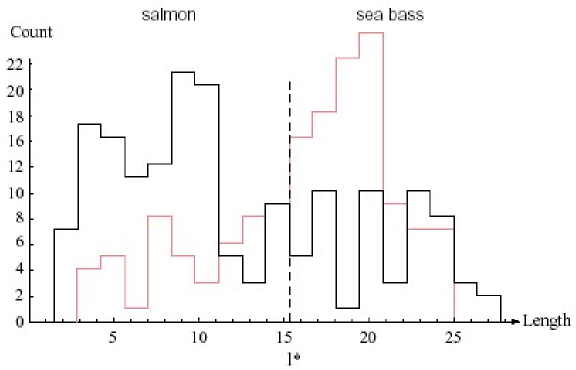


Figure 1.2: Histograms for the length feature for the two categories. No single threshold value l^* (decision boundary) will serve to unambiguously discriminate between the two categories; using length alone, we will have some errors. The value l^* marked will lead to the smallest number of errors, on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

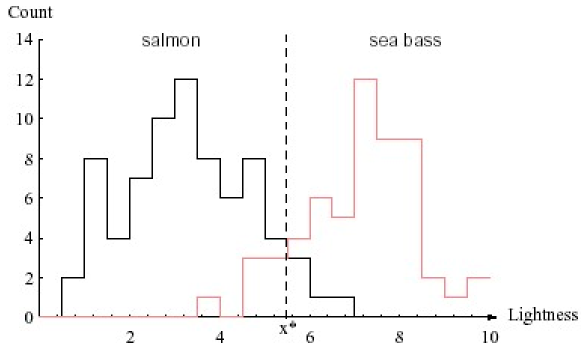


Figure 1.3: Histograms for the lightness feature for the two categories. No single threshold value x^* (decision boundary) will serve to unambiguously discriminate between the two categories; using lightness alone, we will have some errors. The value x^* marked will lead to the smallest number of errors,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

- Decision Theory

- Decision Boundary

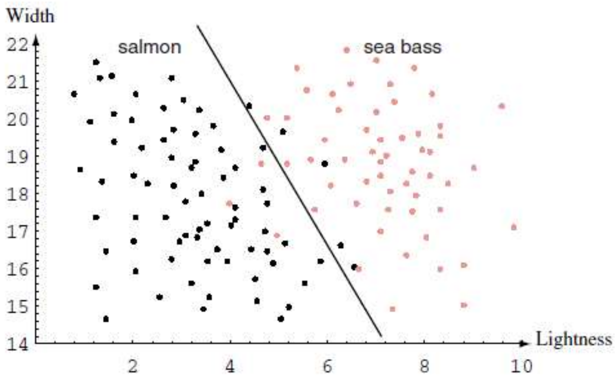


Figure 1.4: The two features of lightness and width for sea bass and salmon. The dark line might serve as a decision boundary of our classifier. Overall classification error on the data shown is lower than if we use only one feature as in Fig. 1.3, but there will still be some errors.

-

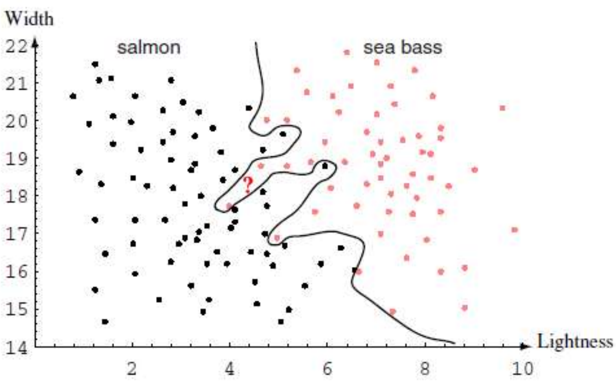


Figure 1.5: Overly complex models for the fish will lead to decision boundaries that are complicated. While such a decision may lead to perfect classification of our training samples, it would lead to poor performance on future patterns. The novel test point marked ? is evidently most likely a salmon, whereas the complex decision boundary shown leads it to be misclassified as a sea bass.

- Generalization

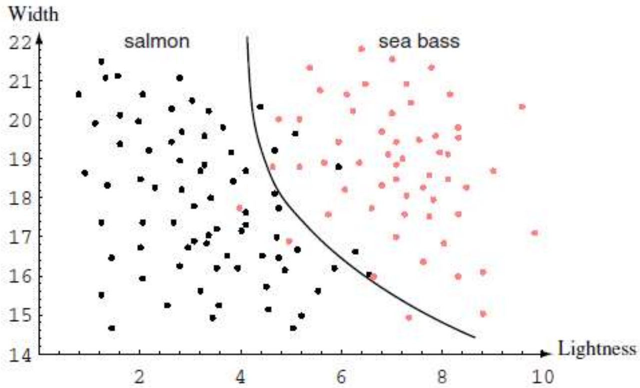


Figure 1.6: The decision boundary shown might represent the optimal tradeoff between performance on the training set and simplicity of classifier.

Components of Pattern Recognition Systems