ECEN 649 Pattern Recognition

Introduction

Ulisses Braga-Neto

ECE Department

Texas A&M University

What is Pattern Recognition?

Possible answer:

"Mathematical, statistical, and computational methods that attempt to automatize the way humans routinely recognize familiar patterns."

Alternate Names

- Machine Learning

- Decision Theory

- Pattern Classification

- Pattern Analysis

- Data Mining

- Artificial Intelligence
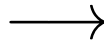
The Prototypical PR System

Vision → HUMAN (mainly right brain hemisphere) ← Hearing ← Smell ← Touch ← Taste
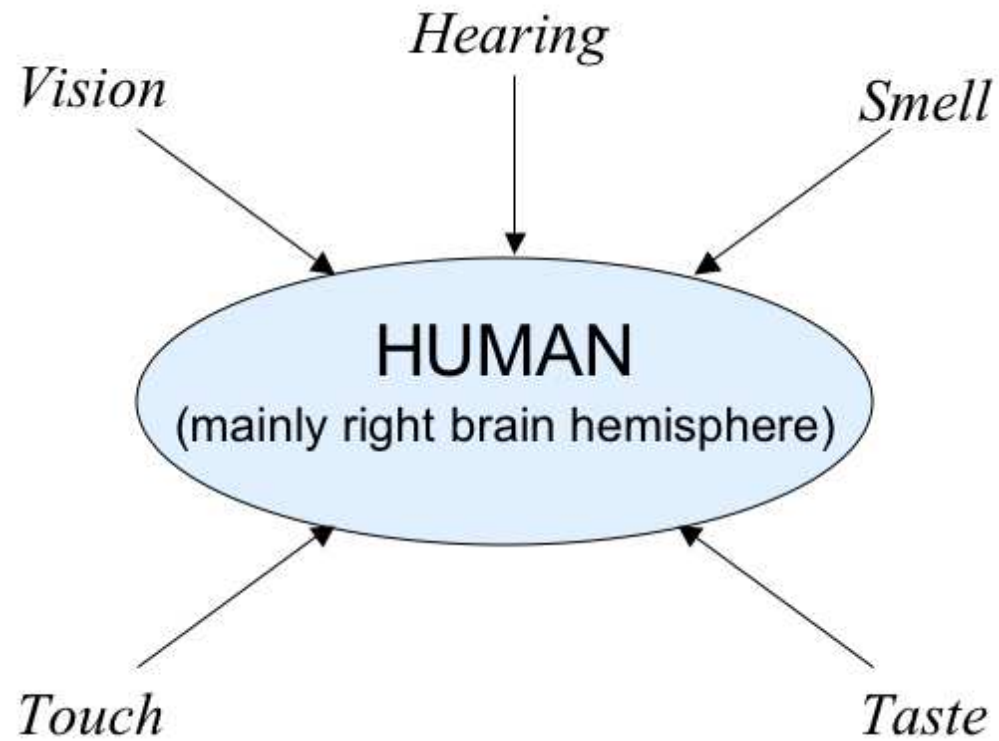
Human PR is Sometimes "Too Good"
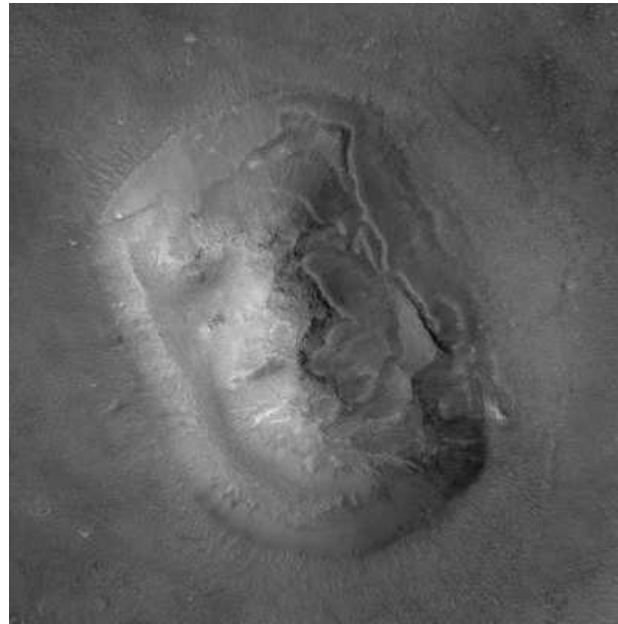
The Infamous "Face" on Mars

→

Picture taken by Viking spacecraft in 1976

This is what the human mind "sees"

Human PR is Sometimes "Too Good"
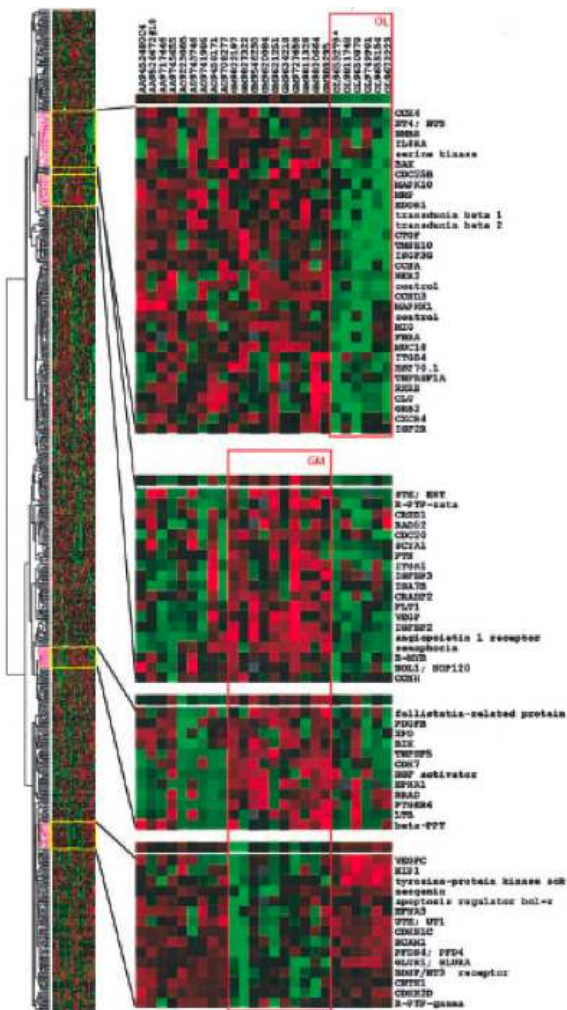
What is really there

Picture taken by Mars Global
Surveyor spacecraft in 2001

Applications of Pattern Recognition

- Image Analysis

- Remote Sensing

- Medical Imaging Diagnostics

- Speech Recognition

- Artificial Noses/Taste Buds

- Robotics

- Genomic Signal Processing
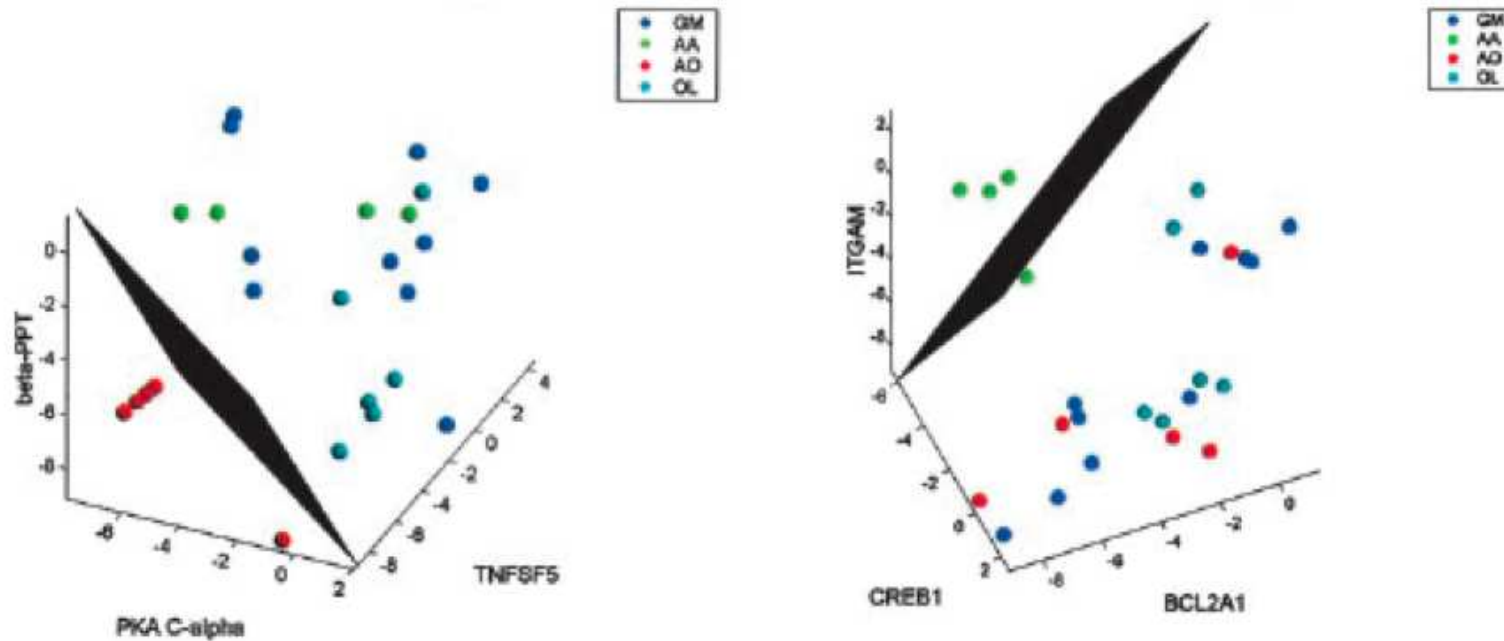
Functional Genomics

Here, the patterns correspond to gene expression values corresponding to 4 types of gliomas: OL, GM, AA, AO.

Single genes can distinguish OL and GM types of Glioma.

From: Kim et al., "Identification of Combination Gene Sets for Glioma Classification," *Molecular Cancer Therapeutics*, 1:1229-1236, 2002

Functional Genomics

Combination of three genes, or *features*,
accomplishes discrimination of AO and AA

Basic Mathematical Setting of PR

- In Pattern Recognition, we have:
  - A *feature vector* X, which contains relevant attributes of the observed entity (the process of obtaining X is called feature extraction/selection).
  - A *label* discrete variable Y (the "state of nature"). E.g., for binary classification, $Y = \{0, 1\}$

- In a complete-information scenario, there is a function f such that $Y = f(X)$.

- Such is rarely the case, however, due to *noise* (sensor imprecision, latent variables, etc.)
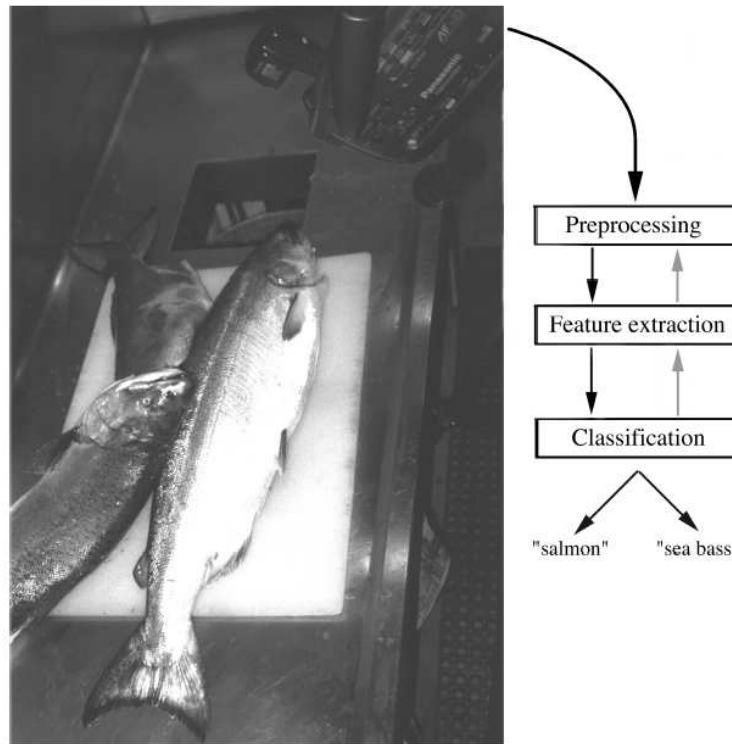
Stochastic Setting

SIGMA CLASSIFIER (*GAUSSIAN*-spread) w/ σ = 1.000
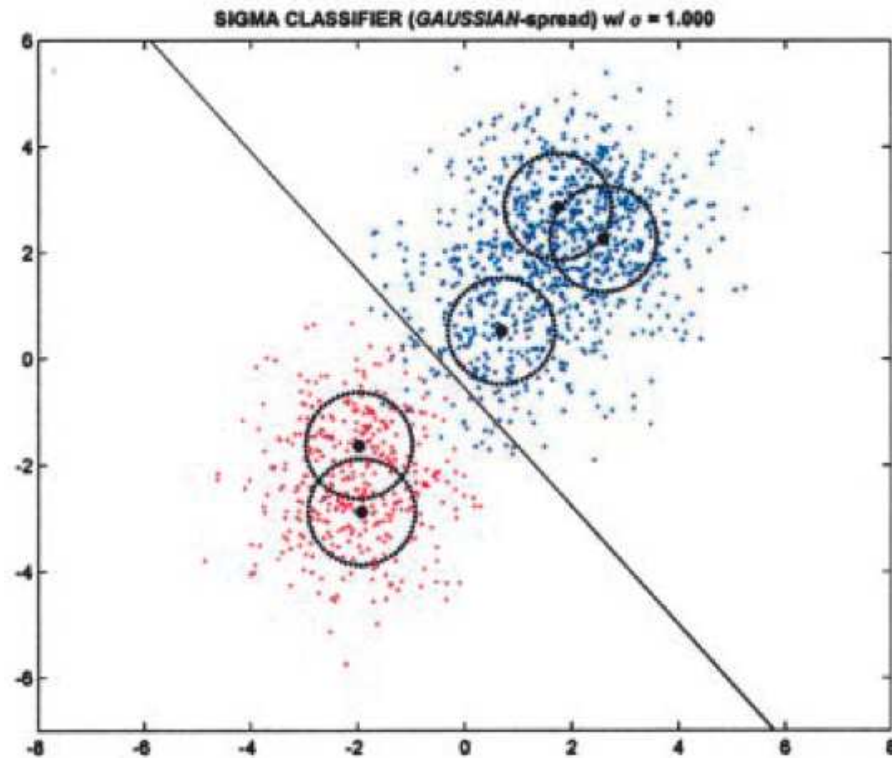
Due to noise, the relationship between Y and X is given by a *joint probability distribution* F_{XY}.

Pattern Recognition Example

Automatic System for Fish-Processing Plant.
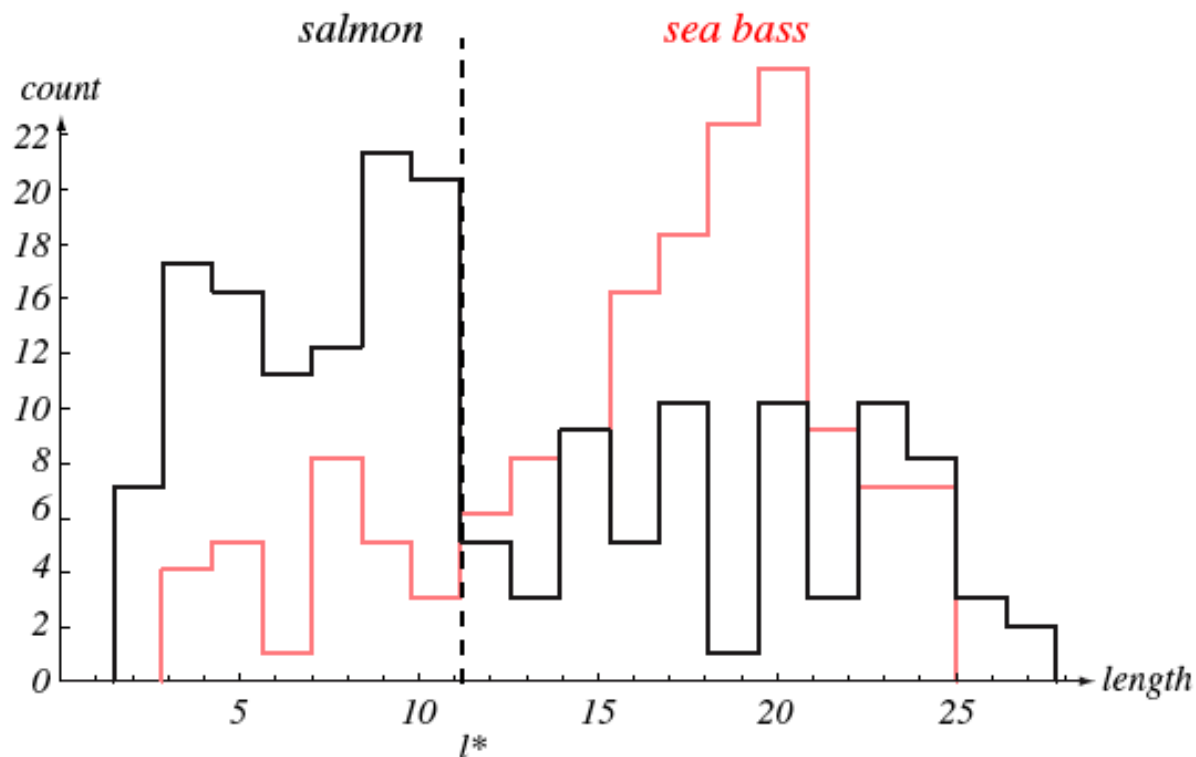Classes: salmon or sea-bass.

$Y = 0$: salmon
$Y = 1$: sea-bass

Pattern Recognition Example - II

Someone may have observed that sea-bass is generally longer than salmon. One may examine the *histogram* of past-observed lengths for each kind of fish.
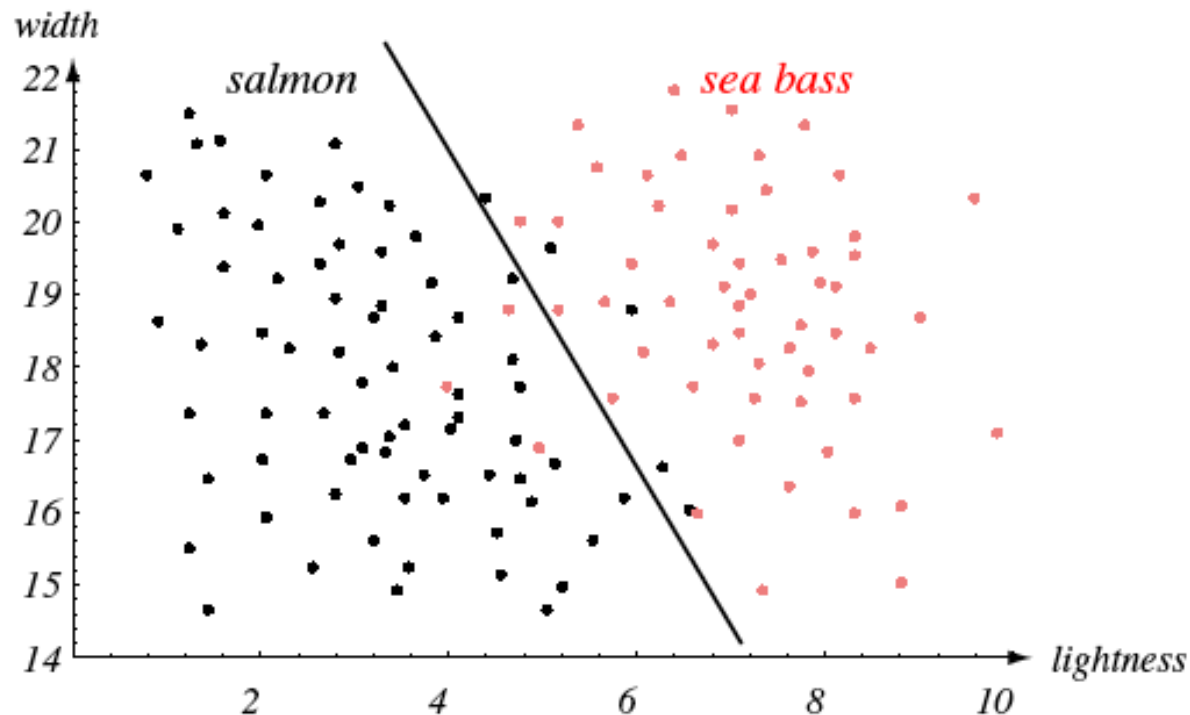
Pattern Recognition Example - III

Length does not seem to work so well, so someone says that sea-bass is generally *lighter* than salmon. One may then also examine the histogram of past-observed lightness for each kind of fish.
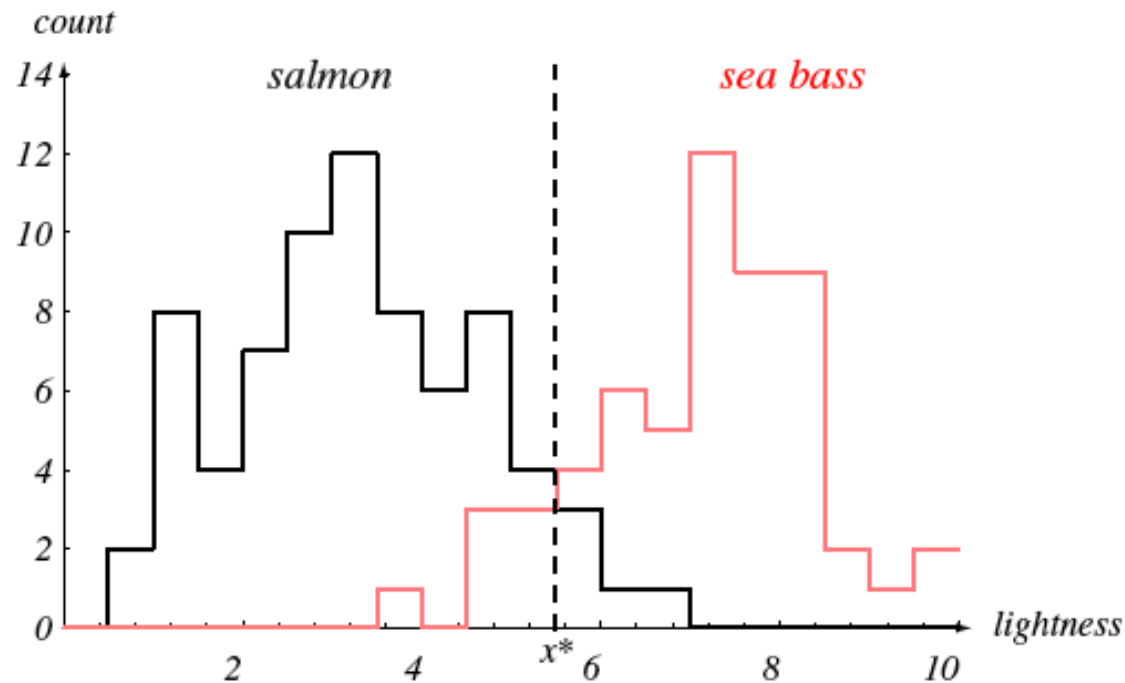
Pattern Recognition Example - IV

One could also use more than one features and obtain a 2-D feature vector. This leads to a 2-D *feature space* and a 2-D *decision boundary* for our classifier.

Error Estimation

- Notice from the previous example that there is an inevitable element of *error* in Pattern Recognition.

- There are three basic kinds of classification error:

  - *Optimal* classification error (minimum possible error).
  - *True* error of *designed* classifier.
  - *Estimated* error of designed classifier.

- Assessment of classification error (called *error estimation*) is a key component of the PR design cycle. This can be based on an independent test set, or on the training data itself.

Feature Selection

- In addition, we can see from the example that the correct choice of features is fundamental to obtain a small classification error. This problem is known as *feature selection*.

- There are two types of feature selection approaches:
  - Filter Feature Selection
  - Wrapper Feature Selection

- More generally, feature selection is an example of *dimensionality reduction*.

Pattern Recognition Design Cycle