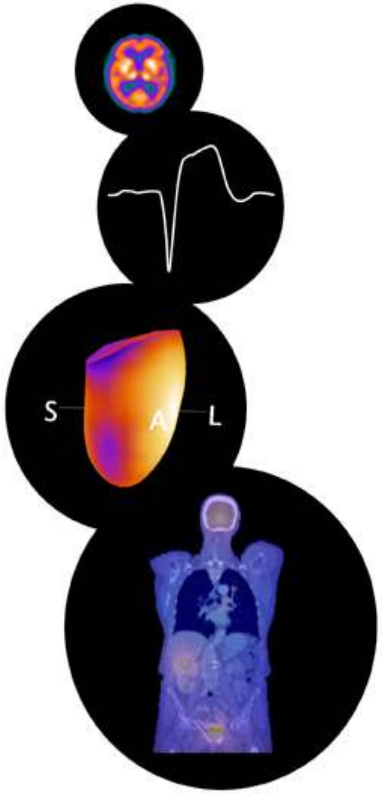


METU Informatics Institute

Min720

Pattern Classification Bio-Medical Applications

Lecture Notes

by

Neşe Yalabık

Spring 2011

Part 1:Introduction

Pattern Recognition and Classification: An Introduction

We human beings do pattern recognition everyday.

- We "**recognize**" and **classify** many things, even if it is corrupted by **noise, distorted** and **variable**.
  - Faces of people
  - Spoken information
  - Written information
  - Medical data
  - Classification is the result of recognition: categorization, generalization

How do we do it?

Automatic pattern recognition has 50 years of history

- Many different approaches tried
- Limited success in many problems
- Successful only with restricted environments and limited categories.

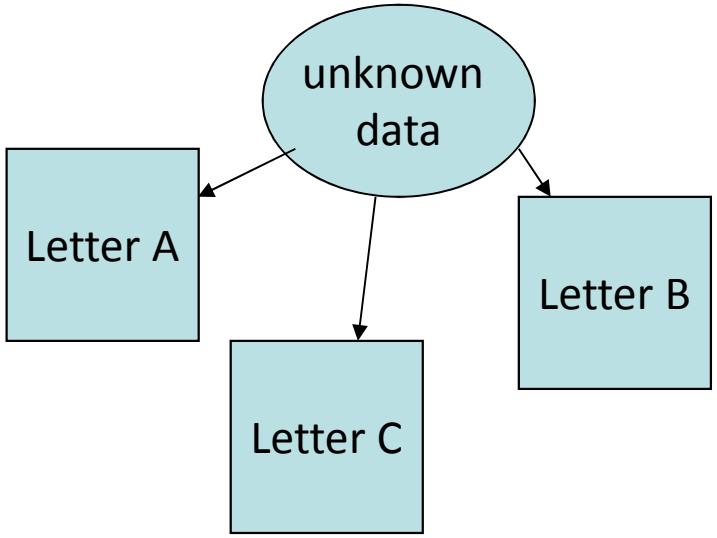
Turns out that **unlimited recognition is still a dream,**

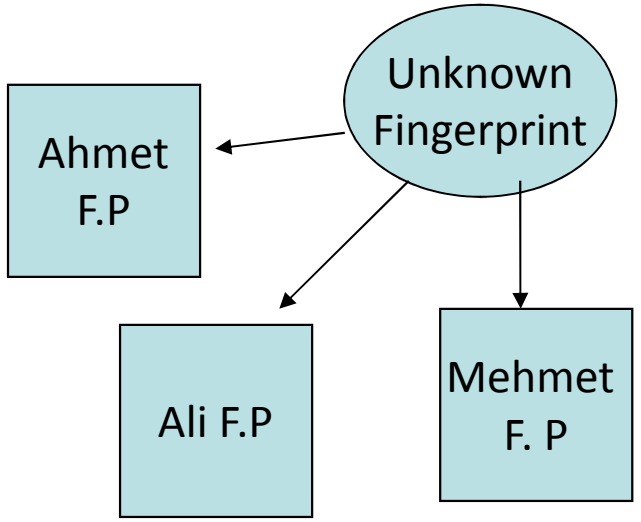
such as:

- Continuous speech recognition
- Cursive script
- Unlimited medical diagnosis
- Unlimited fingerprint recognition

Today **applications aim at limiting these to simpler problems.**

Definition P.R.: The process of machine perception for an automatic labeling of an object or an event into one of the predefined categories.





Objective

Minimize the average error: (at least as good as a human being)

Minimize the risk: wrong decision could be more risky in some cases such as medical diagnosis

Why automize? Obvious reason:save from time and effort

(Ex: consensus forms: enter 100 million records into electronic medium).

How do we solve it? Many different approaches in history

- Template matching
- Use statistics, decision theory "statistical pattern recognition"
- Use " neural networks" Self learning systems
- Structural descriptive approaches: non-numeric information processing – makes use of Formal Language Theory

  ex: Letter A: can be "described" as "two lines intersecting at the top and a third line intersecting the two in the middle

Bio-medical Aplications

- Here, we will quickly look at the applications in medicine, biology and genetics.
- Medicine: many attempts
- Biology

Definition and Terminology

- **Medical Informatics** : Is an interdisciplinary scientific field of research that deals with the use of Information and Communication Technologies and Systems for clinical health care, **for more accurate and faster service to people.**

- **Medical Pattern Recognition**: All PR techniques used in diagnosis, decision support and treatment of illnesses.

Pattern Recognition in Medical Diagnosis Decision Support

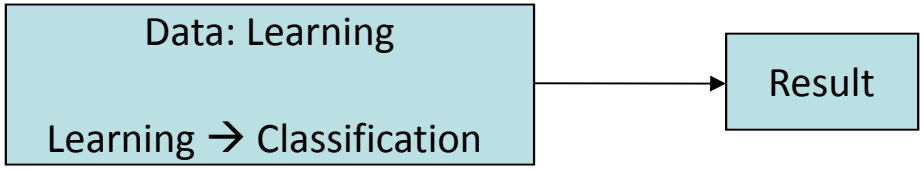
- Earlier in 70's, it was thought that it was easy
- Enter the symptoms, diagnose the desease
- Unfortunately it did not work!
- Most successful application: **Mycin**
- **was designed to diagnose infectious blood diseases and recommend antibiotics.**
- Used 'Expert Systems' approach: 500 rules(if-then statements)
- a correct diagnosis rate of about 65%(better than most physicians),
- Legal issues : Who is responsible for the wrong diagnosis?

Pattern Recognition in Medical Decision Support

- Today, we make systems that we call '**decision support**' that only gives opinion to physician
- Interpreting all kinds of test data
- EKG waveform interpretations
- Locating tumors in x-rays and in other imaging devices
- You will be studying and presenting an application

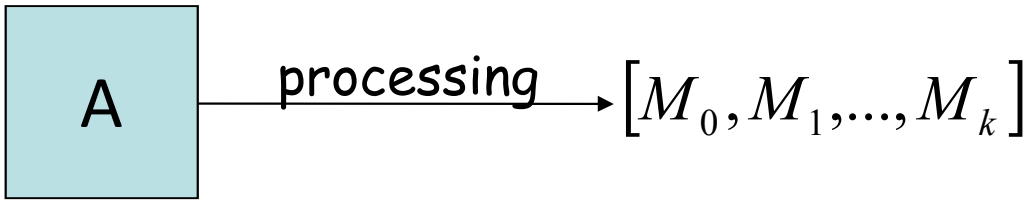
Whichever approach is used, there's a classification process

P.R system:

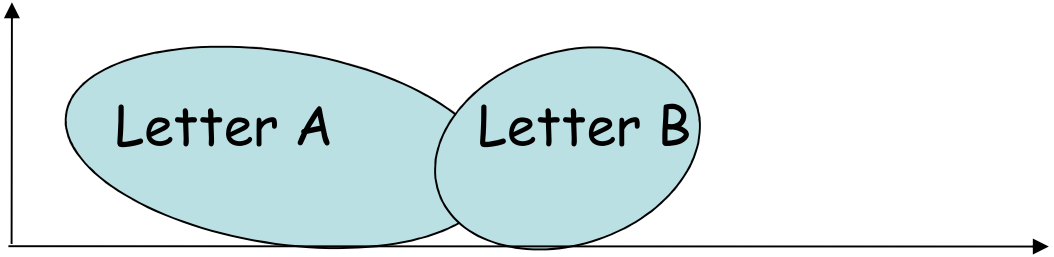


- "Learning samples" : Large data sets to be used in training, or estimating parameters, etc.
- "Result" a decision on the category sample belongs.
- "Test Samples" used in testing the classifier performance.
  L.S and T.S may have an overlap.
- "Data" a raw data pre-processing feature set.
- "Feature" a discriminating, easily measurable characteristics of our data.
- In all approaches, samples from different categories should give distant numerical values for features.

Ex. For letter A, a feature

2-d array 

- M: moments invariants (center of growing obtained from the A feature vector! A model of the underlying system that generated it.



- There is always an error probability in decision!

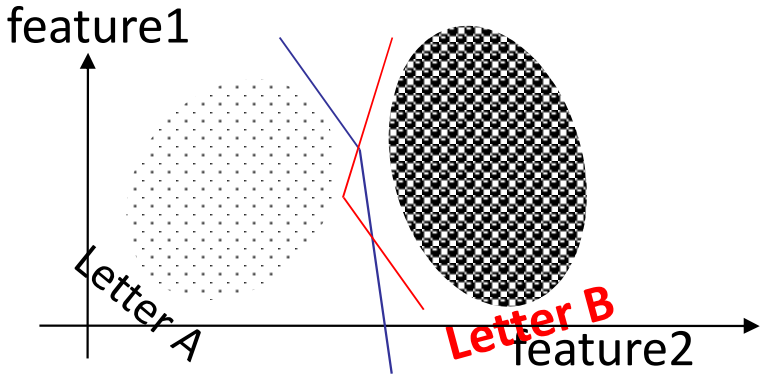
- How many features should we use?

Not small, but not too large either.

(curse of dimensionality)

- Which features should be selected?

Classification



How do we separate A 's from B 's?

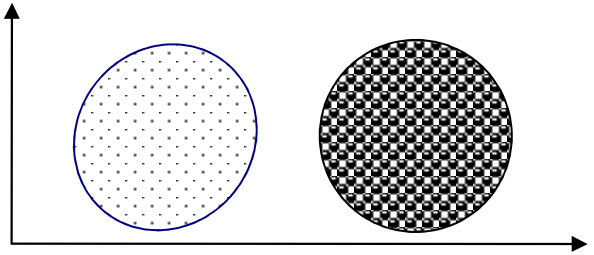
- From a decision boundary
- Classify the sample to the side it falls

Many classification methods exist:

- Parametric: Bayes Decision Theory, Parameterize as belonging to a probabilistic variable.
- Non-parametric: discriminant functions, nearest neighbor rule use only learning samples
- Tree classifiers

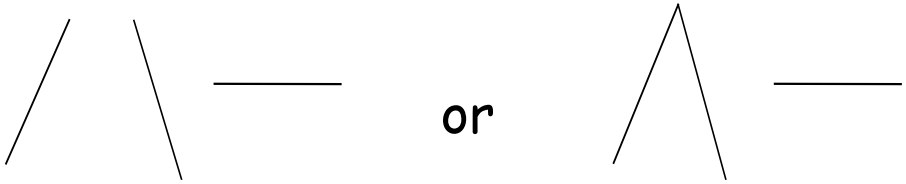
Learning

- Given the learning data set, supervised learning, learn parameters of P.R.



clustering

- If we do not have enough data, we incorporate "domain knowledge" for example, we already know that letter A is written by hand in form of 2 or 3 strokes.



- So maybe recognizing strokes rather than the complete letters first is a better idea. Also consider the text.

- Once the characters are recognized, they might be checked if they are correct by a spell-checker-like system. This is called post-processing "context-dependence".
- To complete system put together looks like: