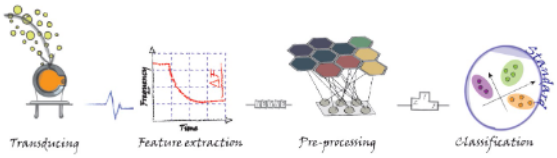


Pattern Recognition and Machine Learning

Prof. Eugenio Martinelli

Contacts : martinelli@ing.uniroma2.it
eugenio.martinelli72@gmail.com

1

Goal

- The main goal of the course of Pattern Recognition and Machine Learning is to give the necessary tools analysing multivariate data (temporal sequences, classification of data, pattern analysis). In particular we will study:
  - Explorative algorithms
  - Regression models
  - Classification models: linear , non linear, bio-inspired
  - And also ….***Deep Learning***
- During the course, a series of practical session will be done.
  - In these sessions the algorithms showed during the lessons will be applied to experimental datasets and the results will be discussed.

2

2

The program

- What is the ***Pattern Recognition and the multivariate data analysis***
- The experimental data and the characteristic of the measurement.
- Fundamentals of statistical data analysis.
- Vectors, vectorial spaces, matrix, statistic of the matrices (covariance matrix eigenvectors, eigenvalue)
- Explorative data analysis:
  - Feature extraction.
  - Data preprocessing.
  - Outliers identification.
  - Principal Component Analysis (PCA)
- Statistical Regression
  - Multiple linear regression(MLR) , Principal component Regression (PCR)
  - Partial Least Square (PLS).
- Theory and algorithm of Pattern Recognition.
  - Linear Classification models supervised and unsupervised (Fisher discriminant analysis, KNN, Mahalanobis classifier, PLS-DA, K-means,…).
  - Validation technique (Leave one out cross validation, Venetian blind, …).
  - Artificial Neural Network (ANN) and Support Vector Machines (SVM).
- Bio-inspired algorithm and their applications.
- Deep Learning: fundamentals

3

Modalità Esame: why in italian…

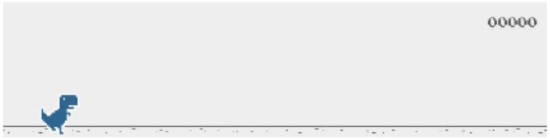
- ***L' esame consiste in***
  - Tesina di Analisi Dati e sua discussione.
    - Analisi di un set di dati sperimentali con tecniche di analisi scelte dal docente.
  - ***Commento di un articolo scientifico***
    - Analisi dettagliata dell'articolo dalla struttura del lavoro , alla presentazione dei risultati, etc…
  - Colloquio Orale sul programma
- ***Durante il corso***
  - Esercitazioni in ambiente matlab.
  - Con i toolbox che verranno usati per la tesina di analisi dati
- ***3 CFU addizionali***

4

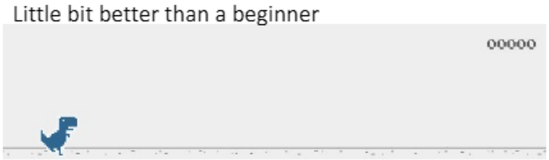
4

From this year: Reinforcement Learning

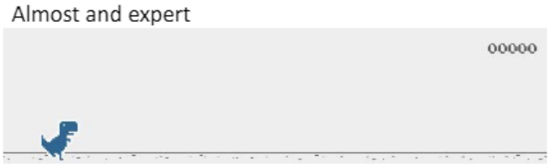
Beginners

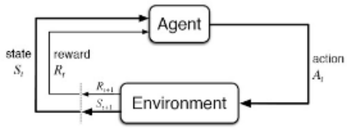


Little bit better than a beginner



Almost and expert

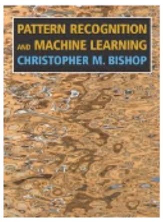


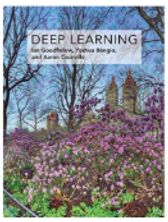


Additional 3 CFUs :
Deep Learning project
(also Reinforcement Learning)

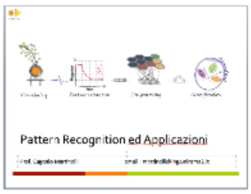
5

Books..





Lesson Slides that will be given by the professor.



Title: Introduzione all'analisi dei dati sperimentali
Authors: D'Amico, Di Natale, Martinelli
Editor: Aracne
2006
ISBN: 8854805939 **Pages:** 180

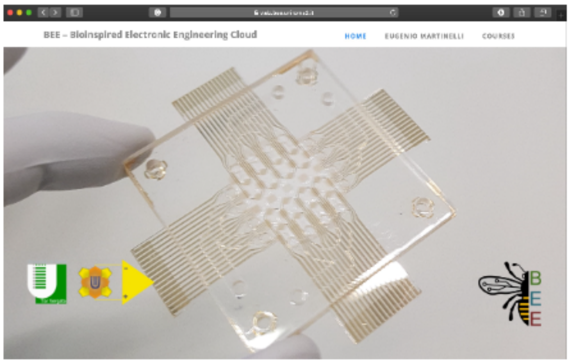
6

6

Where you can find the slides: Forum

https://web.bee.uniroma2.it

https://web.bee.uniroma2.it/category/mlpr/

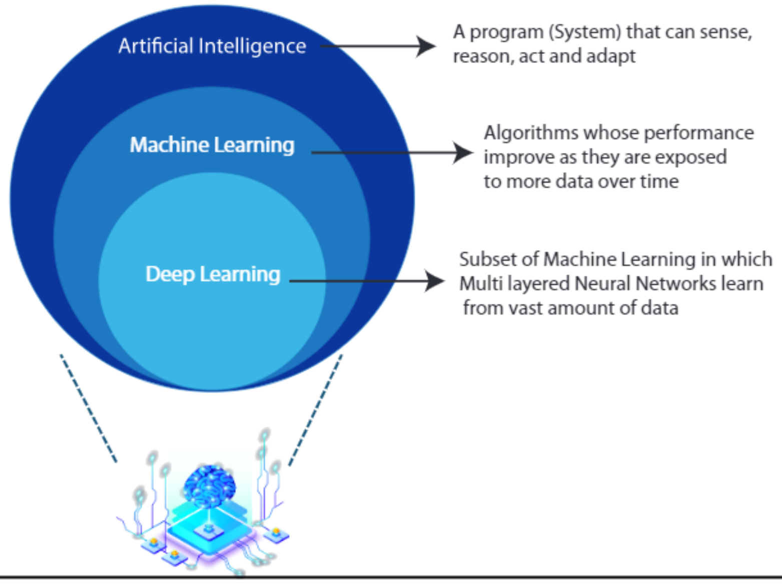


7

Before to start… few definitions

8

AI, Machine Learning and Deep Learning



9

Machine Learning: a Few Quotes…in 2015

- "Machine learning is the next Internet" (Tony Tether, Director, DARPA)
- Machine learning is the hot new thing" (John Hennessy, President, Stanford)
- "Web rankings today are mostly a matter of machine learning" (Prabhakar Raghavan, Dir. Research, Yahoo)
- "Machine learning is going to result in a real revolution" (Greg Papadopoulos, CTO, Sun)
- "Machine learning is today's discontinuity" (Jerry Yang, CEO, Yahoo)

10

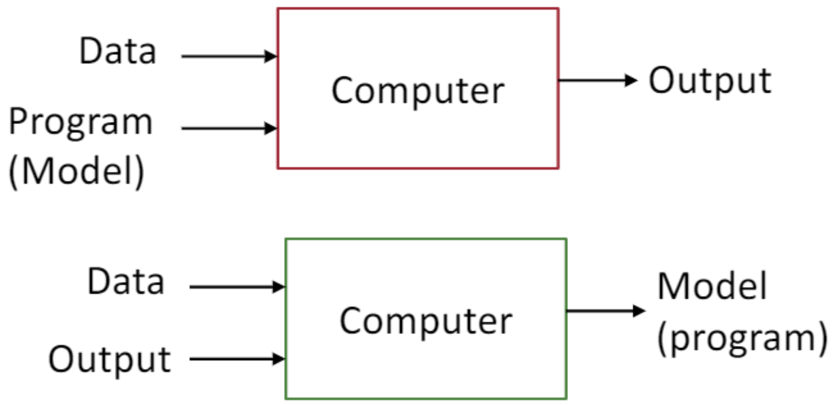
So What Is Machine Learning?

- Automating automation
- Getting computers to program themselves
- Writing software is the bottleneck
- Let the data do the work instead!

11

So What Is Machine Learning?

Traditional Programming



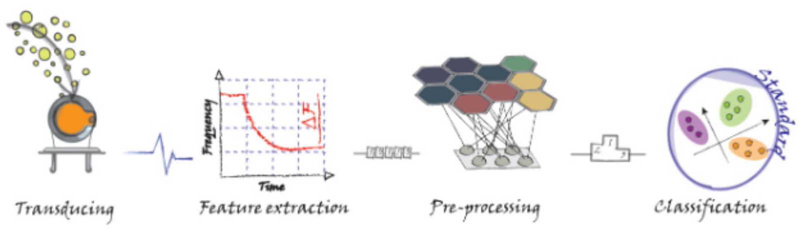
Machine Learning

12

Applications??

- Web search
- Computational biology
- Finance
- E-commerce
- Space exploration
- Robotics
- Information extraction
- Social networks
- Debugging
- [Your favorite area]

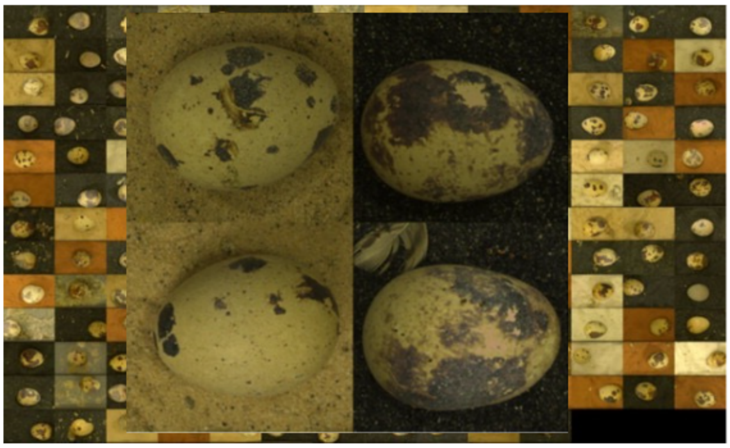
13



Pattern Recognition: a possible definition

14

Egg quail recognition task



15

Pattern Recognition

- Pattern recognition stems from the need for automated machine recognition of objects, signals or images, or the need for automated decision-making based on a given set of parameters.
- A fundamental challenge in automated recognition and decision-making is the fact that pattern recognition problems that appear to be simple for even a 5-year old may in fact be quite difficult when it transferred to machine domain.

ROBI POLIKAR, Pattern Recognition, Rowan University

16

Simple for even a 5-year old but difficult for an engineer

- Consider the problem of identifying the gender of a person by looking at a pictorial representation.

ROBI POLIKAR, *Pattern Recognition*, Rowan University

17

Simple for even a 5-year old but difficult for an engineer

- What are the most meaningful features to distinguish between these two classes—males and females—that the machine should consider to make an intelligent decision?

- It is not difficult to realize that many of the features that initially come to mind, such as
  - hair length,
  - height-to-weight ratio,
  - body curvature, facial expressions, or facial bone structure
- even when used in combination—may fail to provide correct male/female classification of these images!

18

What applications?

- The list of applications can be infinitely extended, but all of these applications share a common factor: *automated classification* or decision making based on observed parameters, such as a signal, image, or in general a pattern, obtained by combining several observations or measurements.

Industrial Application

Biomedical Application

Space Application

Safety Application

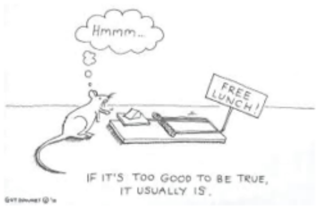
Security Application

ROBI POLIKAR, Pattern Recognition, Rowan University

19

An important remark

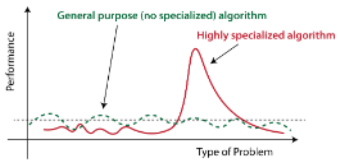
- Literally hundreds of pattern recognition approaches and algorithms exist, and *it is often asked whether any one of them is consistently better than the others.*



20

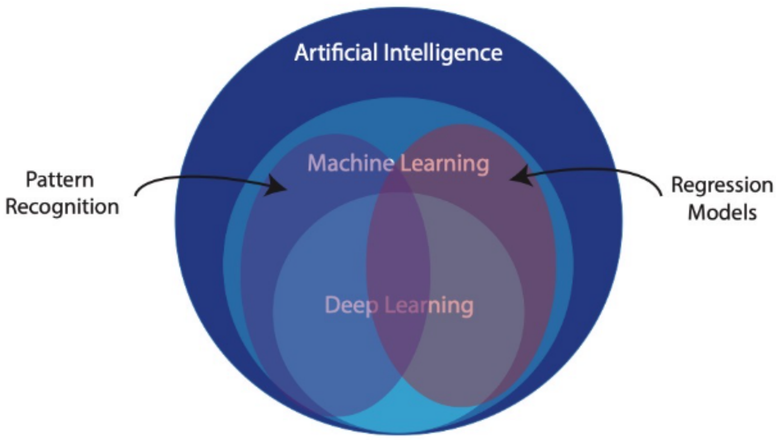
"No free lunch" theorem

- *Then, no-one algorithm is universally superior to all others in the absence of any additional information.*
  - *In fact, it can be proven that problems exist for which random guessing will outperform any other algorithm.*
- ***No Free Lunch Theorem*:**
  - The choice of the most appropriate algorithm almost invariably depends
    - **on the nature of the problem**
    - **the distribution that provides the data for that problem**
    - **the prior knowledge available to the practitioners.**

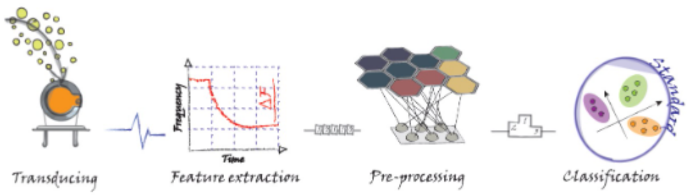


21

Back to AI, Machine Learning and Deep Learning



22



Lesson 1

Pattern Recognition problem

A simple example to explain the problem of patter recognition

23

SeaBass or Salmon

- Suppose that:
  - A fish packing plant wants to automate the process of sorting incoming fish on a conveyor belt according to species.
  - There are two species:
    - Sea bass,
    - Salmon.

24

Sea Bass and Samon

- How to ***distinguish*** one specie from the other ?
  - length,
  - width,
  - weight,
  - number and shape of fins,
  - tail shape,
  - etc.

25

Sea Bass and Salmon

- Suppose somebody at the fish plant say us that:
  - Sea bass is generally longer than a salmon
- Then our ***models*** for the fish:
  - ***Sea bass*** have some typical length, and this is greater than that for ***salmon***.

26

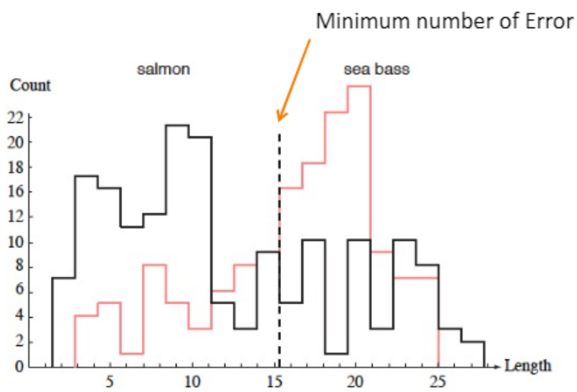
Sea Bass and Salmon (cont..)

- Then length becomes a ***feature***,
- We might attempt to classify the fish by seeing whether or not the ***length*** of a fish exceeds some ***critical value (threshold value)***
- How can we define the ***critical value (threshold value) ?***
  - We could obtain some training samples of different types of fish,
  - make length measurements,
  - Inspect the results.

27

Sea Bass and Salmon (cont..)

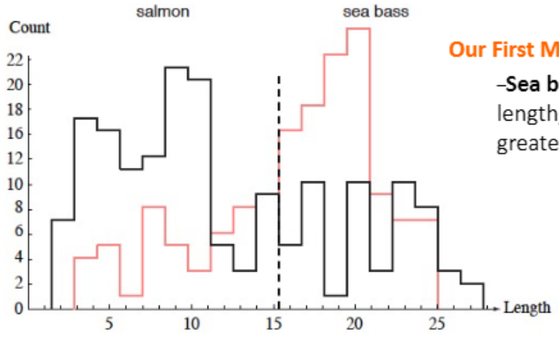
- Measurement results on the training sample related to two species.



28

Sea Bass and Salmon (cont..)

- Can we ***reliably separate*** sea bass from salmon by using length as a feature ?



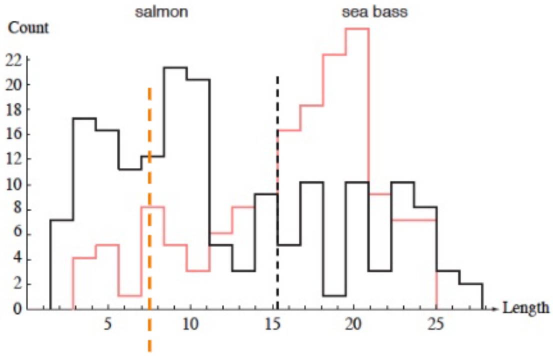
Our First Model:

–**Sea bass** have some typical length, and this should be greater than that for **salmon**.

29

Sea Bass and Salmon (cont..)

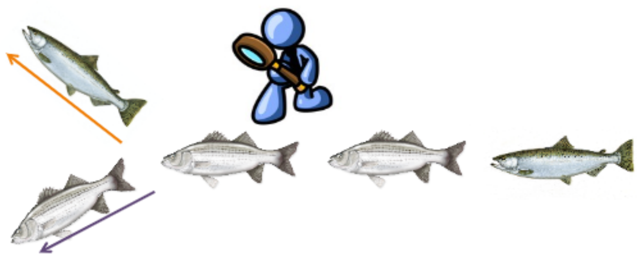
- From histogram we can see that single criteria is quite poor.
- But what is the real goal of our model?



30

A second feature

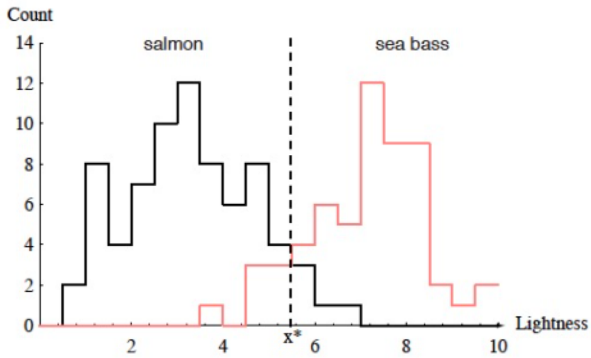
- First Conclusion: it clear that the feature "Length" is not a good feature.
- ***What we can do to separate sea bass from salmon?***
- ***Try another feature:***
  - ***average lightness of the fish scales.***



31

A second feature

- Can we ***reliably separate*** sea bass from salmon by using lightness as a feature ?

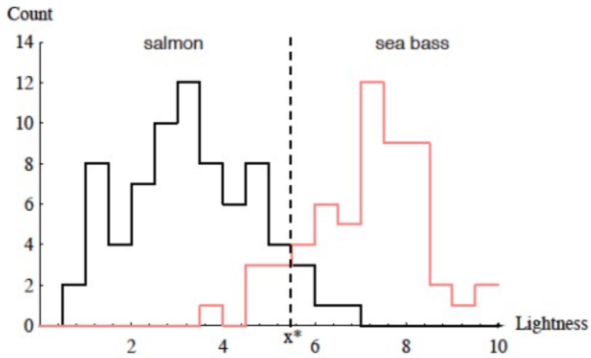


32

A second feature…the same problem

- Lightness is better than length as a feature but…

  ….. again there are some problems.



33

A simple solution

- Resuming:
  - Sea bass are typically wider than salmon.
  - Sea bass are typically ligthness than salmon.

THEN…..

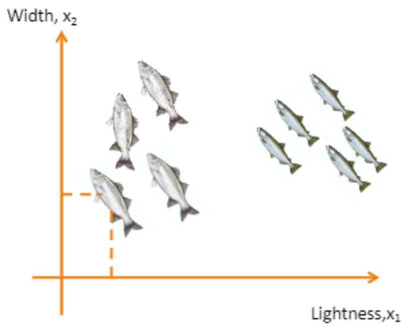
- We can use more than one feature for our decision:
  - Lightness (x_1) and width (x_2)

34

An Example

- Each fish is now a point in two dimension.
  - Lightness (x_1) and width (x_2)

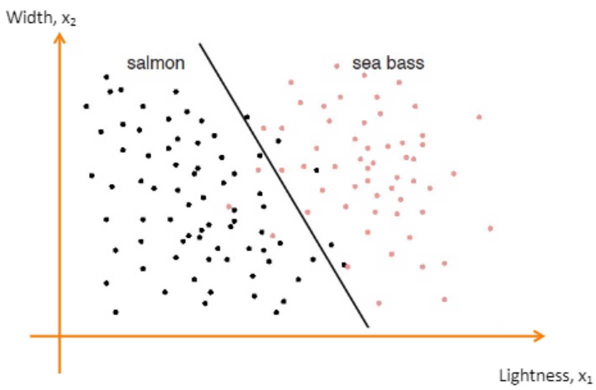
$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$



35

An Example

- Each fish is now a point in two dimensions.
  - Lightness (x_1) and width (x_2)



36

Cost of error

- ***Cost of different errors*** must be considered when making decisions.
- We try to make a decision rule so as to ***minimize*** such a cost.

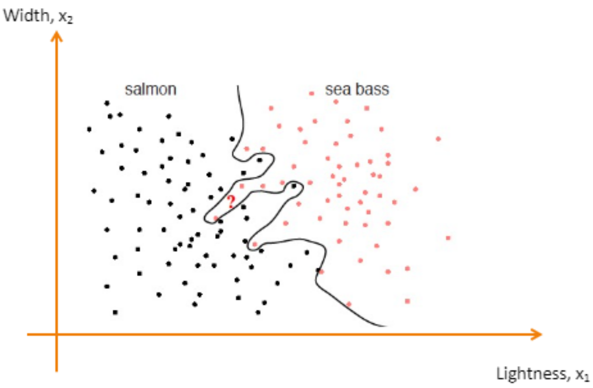
This is the central task of decision theory.

- For example, if the fish packing company knows that:
  - Customers who buy salmon will ***object*** if they see sea bass in their cans.
  - Customers who buy sea bass will ***not be unhappy*** if they occasionally see some expensive salmon in their cans.

37

Decision boundaries

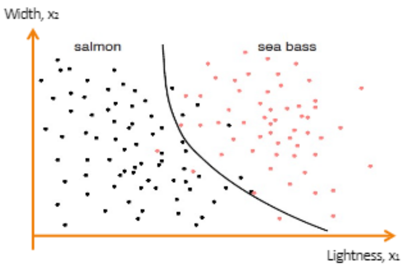
- We can perform better if we use more complex decision boundaries.



38

Decision boundaries

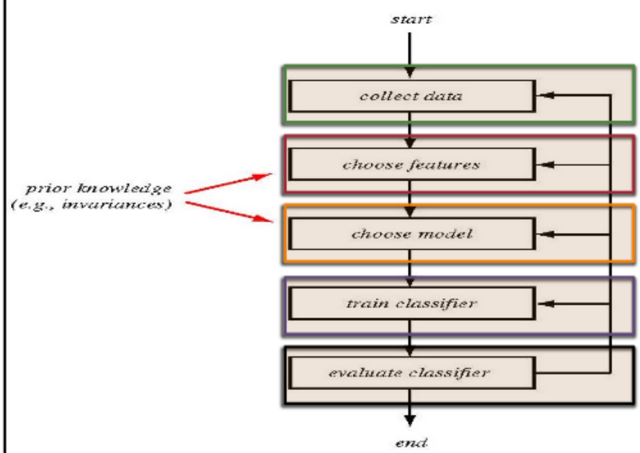
- There is a trade off between complexity of the decision rules and their performances to unknown samples.
- ***Generalization:*** The ability of the classifier to produce correct results on *novel* patterns.
- ***Simplify the decision boundary!***



39

The design cycle

....advices to the practitioners



- Experimental Design.
- Sources of disturbances

Garbage input...Garbage out

Experimental Design.
Sources of disturbances..

- Representative Dataset

- External dataset
- Problem of problems of overfitting and generalization.

40

Summary…

- Pattern recognition techniques find applications in many areas: machine learning, statistics, mathematics, computer science, biology, etc.
- There are many sub-problems in the design process.
- Many of these problems can indeed be solved.
- More complex learning, searching and optimization algorithms are developed with advances in computer technology.
- There remain many fascinating unsolved problems.

41

References

- Srihari, S.N., Covindaraju, Pattern recognition, Chapman &Hall, London, 1034-1041, 1993,
- Sergios Theodoridis, Konstantinos Koutroumbas , pattern recognition , Pattern Recognition ,Elsevier(USA)) ,1982
- R.O. Duda, P.E. Hart, and D.G. Stork, Pattern Classification, New York: John Wiley, 2001,

42