Computational analysis of biological structures and networks

Manuele Bicego

Master's degree in Medical Bioinformatics

Computer Science Dept.
University of Verona

The course

Timetable:

Monday   8.30 - 11.30 (Lecture Hall G)

Monday   15.30 - 18.30 (Lab Alfa)

Lecturer

Manuele Bicego

Office	Ca' Vignal 2, Floor 1, Room 48b
Telephone	045 8027072
Email	manuele.bicego@univr.it
Office Hours	During the semester: Wednesday 8.30-10.30 In alternative: send an e-mail

Requirements

- Notions from the course "Riconoscimento e Recupero dell'informazione per Bioinformatica" (BS degree in Bioinformatics)

- Basic notions of Algorithms, Probability, Statistics, Algebra

- Programming skills (for lab part)
  - Programming language used: Matlab

Overview

- Title: "Computational analysis of biological structures and networks"

- Many possible viewpoints: here we adopt the Pattern Recognition perspective

- The course is about **Pattern Recognition tools** and **techniques** to model biological *complex* objects

  - Objects with a *complex structure* (strings, 3D structures, sets, graphs, networks...)

Contents

The course is divided in two parts:

PART 1: **Theory**

PR tools and techniques to model structured data

PART 2: **Laboratory**

Implementation of algorithms studied during the theory part (*matlab*)

Program (Theory)

- **Chapter 1.** Introduction
  - Basic Pattern Recognition concepts (recap from Bs. Course)
  - Introduction to structured data (data with complex structure)

- **Chapter 2.** Representation of structured data
  - The Bag of words representation
  - The dissimilarity-based representation
  - Advanced dimensionality reduction techniques

Program (Theory)

- **Chapter 3.** Models for structured data
  - Generative models (Bayes Networks)
  - Learning and inference

- **Chapter 4.** Advanced concepts
  - Kernels for structured data
  - Advanced learning paradigms for structured data (Multiple instance learning, semi supervised learning, transfer learning)

Material

- Slides, notes, suggested readings...

- Course slides will be posted on line *before* classes

- All info can be found a the course homepage

Reference books

- R. Duda, P. Hart, D. Stork *Pattern Classification.* Wiley, 2001 (2nd edition).

- S. Theodoridis, K. Koutroumbas: *Pattern Recognition*, Second edition, Academic press, 2003

- C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

- B. Frey: *Graphical Models for Machine Learning and Digital Communication*, MIT Press, 1998

- E. Pekalska, B. Duin, *The Dissimilarity Representation for Pattern Recognition*, World Scientific Press, 2005

Some specific readings will be suggested for every chapter

Assessment methods

Two parts:

- First part: few questions on course topics (**written exam**, during exam sessions)

- Second part: **talk** within a thematic workshop (as in a conference)

Assessment methods

Talk within a thematic workshop: details

- The topic of the thematic workshop will be decided in advance (before middle of November)

- Each student has to choose a scientific paper to be presented in 10 minutes

- One thematic workshop will be held at the end of the course (registration needed by early December)

- Other sessions in June and September