

Machine learning for structured domains by Kernel methods

Prof. Alessandro Sperduti¹

¹ *University of Padua*
Department of Mathematics
Email: sperduti@math.unipd.it

Timetable: 14 hrs. First lecture on June 7, 2013, 9:00 (dates already fixed, see the calendar), Torre Archimede, Room 2BC/30.

Course requirements: Basic knowledge of linear algebra, algorithms, calculus, probability.

Examination and grading: The students will be evaluated on an written essay on one of the topics covered in the course.

SSD: INF/01 - Computer Science

Aim: The aim of the course is to give a comprehensive description of the field of machine learning using kernel methods for tree and graph structured data, outlining the specific challenges posed by these domains.

Course contents: The amount of data available in electronic format is increasing at such a rapid pace that intelligent automatic techniques for extraction of relevant information are gaining more and more importance. In many application domains, such as biology, chemistry, natural language processing, data is naturally represented in structured form: proteins and phylogenetic trees, molecular graphs, hypertextual and XML documents, parse trees.

Since traditional methods in machine learning deal with vectorial information, they require an a-priori form of preprocessing. However, various algorithms have been proposed in literature which are able to directly deal with structured data. Among the most popular, there are neural based algorithms and kernel methods. Especially the latter are recognized to have a strong theoretical background and state of the art results in many applications.

Syllabus:

1. Introduction to Machine Learning
2. Kernel Methods
3. Expressiveness versus Efficiency in kernel definition
4. General Frameworks
5. Kernels for Trees: Algorithmic, Neural and Probabilistic approaches
6. Kernels for Graphs: based on walks, paths and subtree patterns
7. Software resources and application examples
8. Future Directions