UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Machine Learning

2526-1-F1802Q101

Aims

The course aims to introduce the basic elements and provide the basic concepts and tools based on computational methods, to represent the learning, knowledge and reasoning under uncertainty.

The student will be able to design and develop software systems based on machine learning techniques to solve simple problems of data analysis, knowledge discovery and decision support in the presence of uncertain or incomplete information. Acquired skills are the basis for the development of advanced data mining and knowledge discovery and application specific software systems in decision support.

The specific objectives with respect to the Dublin Descriptors (DdD) are the following:
1. Knowledge and understanding

- the theoretical background of machine learning;
- understanding of the main models (decision trees, neural networks, SVM, Bayesian models, k-means, etc.);

2. Applied knowledge and understanding

- design and develop software tools based on machine learning models;
- apply supervised and unsupervised classification methods to real problems;

3. Autonomy of judgment

- select the most appropriate methods in relation to the nature of the problem and the data;
- evaluate the quality of the proposed solutions based on the results obtained by the models with respect to performance metrics;
- critically analyze the assumptions of the models (e.g. linear separability, independence of variables) and the results obtained with respect to these assumptions.

4. Communication skills

- communicate design and methodological choices in a reasoned manner
- present and discuss clearly the results of a project during the oral exam;
- work in a group, sharing knowledge.

5. Learning ability

- develop autonomous study skills through the completion of optional assignments resulting from laboratory activities focused on the use of machine learning models

Contents

The course will provide an in-depth discussion of the large collection of methods and tools that have become available for developing autonomous learning systems and for aiding in the analysis of complex multivariate data. These tools include decision trees, neural networks, belief networks, as well unsupervised clustering algorithms and increasingly sophisticated combinations of these architectures. Examples of recent applications will be introduced.

Detailed program

Introduction and terminology

Concept learning
· Decision trees
· ID3 Algorithm

Neural networks
· perceptron; linear separability
· perceptron learning algorithm
· delta rule and gradient descent
· multilayer perceptron
· error backpropagation

Support vector machines
· Optimal separation
· Hard and Soft SVM
· Kernels

Bayesian Learning
· Bayes theorem
· Naive Bayes classifier

Unsupervised learning techniques
· k-means algorithm

Performance evaluation
· evaluation for supervised techniques
· evaluation for clustering

Deep neural network (Deep Learning)

· Introduction
· Main reference models

Prerequisites

Basic knowledge of algorithms and data structures.

Elements of probability and statistics

Teaching form

- 8 frontal lessons of 2 hours each held by the teacher in presence;
- 10 sessions for exercises of 2 hours each held by the teacher in presence, 50% frontal 50% interactive;
- 12 interactive laboratory lessons of 2 hours each held by the teacher in presence;

Textbook and teaching resource

References on the elearning page of this course

Semester

First semester

Assessment method

The exam consists of a written test and an oral test, each valid for 50% of the grade, and is carried out according to the following methods:

1. **Written test** with theoretical questions and exercises. To access the oral test, it is necessary to obtain a score >= 6 in the written test. The written test will have a score from 0 to 30.

2. **Oral test** focused on the discussion of a project typically carried out in a group. The evaluation of the project is based on the overall quality of the report and on questions that are specifically related to the teaching program. The oral test will have a score from 0 to 30.

The **final grade** will be determined as average between the written test and the oral test.

Students have the opportunity to take two intermediate tests as a substitute for the written test. Each intermediate test will consist of theoretical questions and exercises and will have a score from 0 to 30.

Office hours

by requiring an appointment

Sustainable Development Goals