UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Statistical Learning

2425-2-F8204B018-F8204B033M

Learning objectives

The course aims to acquire the main techniques of statistical learning (SL) and their implementation in the R programming environment. During the course, emphasis will be placed on the algorithmic modelling culture, while also paying attention to the estimation of uncertainty in predictions.

The course contributes to the achievement of the learning objectives in the subject area of the MSc: "Statistics".

Contents

Main subjects are:

- Tree-based methods and computational aspects.
- Deep Learning for unstructured data.
- Uncertainty estimation.

Detailed program

Tree-based methods.

- Decision trees: classification and regression.
- Bagging.
- Random forests.
- Boosting and additive trees.
- Ensemble learning

Focus: gradient boosting algorithm.

**Deep learning

- Neural networks.
- Convolutional neural networks.
- Recurrent neural networks.

Uncertainty estimation.

- Conformal prediction.

Prerequisites

Knowledge of topics covered in the courses Probability and Statistics M, Advanced Statistics M and Data Mining (module of Data Science M) is highly recommended.

Teaching methods

Lessons are held both in classroom and in lab, integrating theoretical principles with practicals aspects of data analysis and programming in R.

The 47 hours of teaching are organized as follows:

- 35 hours of lectures, in person;
- 12 hours of laboratory activities conducted interactively and remotely.

Assessment methods

The examination consists of a written/practical examination to be conducted in the laboratory.

(32 points out of 32) Written examination with open and closed questions, in which the theoretical aspects of the course are assessed. There will also be questions to be completed in R.

Textbooks and Reading Materials

T. Hastie, R. Tibshirani, J. Friedman (2017) The Elements of Statistical Learning. Springer.
D. Efron, T. Hastie (2016) Computer-Age Statistical Inference: Algorithms, Evidence, and Data Science. Cambridge University Press.
I. Goodfellow, Y. Bengio, A. Courville (2016) Deep Learning. MIT Press.
F. Chollet, J.J. Allaire (2018) Deep Learning with R. Manning.

Further readings will be suggested during the course.

Semester

First semester, second cycle.

Teaching language

The lessons are held in Italian, but the materials and textbooks are in English.
Lessons can be conducted in English if necessary.

Sustainable Development Goals

QUALITY EDUCATION