UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Statistical Learning M

2324-2-F8204B015

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.

By the end of the course, the student will be able to deal with complex applications.

Contents

Algorithms/models: tree-based methods.
Forecasting: uncertainty estimation.
A new paradigm: reinforcement learning.
An SL problem: p>>n

Detailed program

Algorithms/models: tree-based methods.

- Decision trees
- Bagging
- Random forests
- Boosting and additive trees
- Ensemble learning

Forecasting: uncertainty estimation

- Conformal prediction: an introduction

A new paradigm: reinforcement learning.

- Markov decision process
- Policy search
- Q-learning

An SL problem: p>>n

- When p is much greater than n
- Sure Independent Screaning (SIS)

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 taught in classroom and lab.

Assessment methods

Written exam.

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.

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