



UNIVERSITÀ  
DEGLI STUDI DI MILANO-BICOCCA

## SYLLABUS DEL CORSO

### Data Science Lab in Public Policies and Services

2425-2-FDS01Q028

---

#### Aims

This module aims at teaching students how to analyze medical data (especially, data of electronic health records) through computational statistics and machine learning techniques to infer new knowledge about the conditions of patients.

#### Contents

- Dataset search and retrieval
- Data preparation and data cleaning
- Exploratory data analysis
- Unsupervised machine learning
- Supervised machine learning
- Feature ranking
- Result understanding and validation
- R and Python programming languages

#### Detailed program

- Dataset search and retrieval
- Data preparation and data cleaning
- Exploratory data analysis
- Unsupervised machine learning
- Supervised machine learning
- Feature ranking

Result understanding and validation  
R and Python programming languages

## **Prerequisites**

Basic statistics and basic machine learning  
Basic knowledge of R or Python

## **Teaching form**

In-person theory classes and practice exercise classes

## **Textbook and teaching resource**

Classes slides and scientific papers mentioned during classes

## **Semester**

Second semester

## **Assessment method**

Personal work on a scientific project, delivery of a report on the work done, and oral presentation of the work done

## **Office hours**

To define via email by writing to [davide.chicco\(AT\)unimib.it](mailto:davide.chicco(AT)unimib.it)

## **Sustainable Development Goals**

GOOD HEALTH AND WELL-BEING

---