UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Big Data in Economics

2526-2-FDS01Q021-FDS01Q021M

Learning objectives

By the end of the course, students will have acquired:

1. Knowledge and understanding
   Foundational and advanced knowledge in data science, computer science, and applied statistics, with a focus on economic and business-related problems.
   Understanding of core concepts related to causal inference, prediction, and unsupervised classification.
2. Applying knowledge and understanding
   Ability to apply data analysis techniques to real-world economic and business problems using case studies and real datasets.
   Use of computational tools and quantitative models in complex business scenarios.
3. Making judgements
   Development of critical thinking in selecting appropriate methodological tools for each type of problem (causality, prediction, clustering).
   Ability to interpret statistical results for decision-making and strategic purposes.
4. Communication skills
   Ability to clearly and effectively communicate analytical results, including visual and technical reporting, to both specialist and non-specialist audiences in the economic and business domain.
5. Learning skills
   Ability to independently deepen knowledge of data analysis and machine learning tools, supported by a strong methodological foundation, and apply them in evolving professional contexts.

Contents

The course is divided into 4 parts.

The first part discusses the role of big data within within the firm and the new challenges. The remaining three parts of the course separately discuss the three main areas of application (causality, prediction and unsupervised classification) with specific examples mainly concerning risk management and consumer choices.

Finally, in hands-on lab students learn to develop R algorithm for data analysis.

Detailed program

1. Introduction and definition of the problem: the Big Data Challenge

2. The role of uncertainty: Cause, prediction and unsupervised classification.

3. Causal mechanisms: fundamental elements and a case study.

4. Prediction: the challenge of assessing uncertainty in predictive models.

5. Unsupervised learning: Self-Organizing-Map and marketing

6. Bonus track: the analysis of drift in business models.

7. Reporting of company results: creating a narrative around the model

Prerequisites

Principles of programming in R or Python. Basic statistics.

Teaching methods

The 21-hour course consists of:

11 hours of remote synchronous laboratory sessions.
10 hours of in-person lecturing sessions.

Assessment methods

Project and written exam about the course content (50%-50%) . The assessment depends on the correctness and the clarity of the answers.

Textbooks and Reading Materials

The reading material is based on journal articles and selected book chapters. The material will be available on the e-learning platform.

Book

Data Science for Business
What You Need to Know about Data Mining and Data-Analytic Thinking
By Foster Provost, Tom Fawcett

Semester

II semester

Teaching language

English

Sustainable Development Goals

INDUSTRY, INNOVATION AND INFRASTRUCTURE