UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Data Science Lab

2526-1-FDS02Q003

Learning objectives

To make students capable to set and lead data science projects on complex data systems.

Particularly, the course focuses on projects addressing soft and "open" questions, where a precise specification of analytical goals lacks, differently from a "classical statistics setting" and where data sources are the basis for new services, whose design and implementation requires creativity, managerial and communicative skills as well as methodological competencies.

In addition, some lessons/workshops will be devoted to algorithms and methodologies for treating some data structures of particular interest in the data science practice (e.g. temporal data or partially ordered data).

SPECIFIC LEARNING GOALS
Knowledge and understanding: addressing open problems, students will develop capabilities to understand and synthesize complex phenomena.

Applying knowledge and understanding: through the development of the project, students will learn to apply the knowledge acquired during their studies.

Making judgements: students must assess autonomously the value of their research question and similarly identify the most proper statistical path to answer it.

Communication skills: through the development of the project, students will learn to take the perspective of the final user of the research.

Learning skills: since the statistical methodologies required by the project depend upon the research question and are not identified in advance, .students will learn ongoing how to acquire new competencies.

Contents

The course illustrates and introduces to the kind of activities that a data scientist performs in socio-economic and business contexts and the topic of how to manage data science projects. In practice, these topics are addressed by means of concrete projects to be managed by groups of students, with the supervision of the teacher.

Detailed program

The course is divided into two parts. In the first some lessons, under the form of seminars, will be held to touch upon some specific basic topics, In the second part, students will be assigned projects to be managed and concluded within the course.

PART I

1. The current technological and economic context: complex socio-economic processes, the need for new knowledge and innovative services.

2. Data complexity and new data sources: web, e-commerce, Internet of Things, Smartphones...

3. Data-driven business processes: marketing, crm, operations...

4. Examples of data science projects.

5. Basics of Project Management: specificities and criticalities of data science projects, data quality and technological choices.

6. Computability and the limits of machine learning algorithms.

7. The statistical synthesis of complex phenomena and the construction of synthetic indicators.

8. Analysis of partially ordered data.

PART II

Definition and assignment of data science projects to student groups (each group is composed of 3-4 students; the content of the projects will be identified during the course) .

Prerequisites

There are no formal prerequisites but basic competencies in inferential statistics, data anaysis, data mining and R programming are necessary.

Teaching methods

Frontal lessons and project supervision (The didactic activity will be givrn as LECTURES, apart from project supervision, that will be perfomred in an INTERACTIVE way).

LECTURES: 35 hours
INTERACTIVE DIDACTIC: 13 hours

Assessment methods

The competence level will be assessed through a final PROJECT WORK, with no intermediate assessments.

The final mark is determined by assessing: (i) the relevance of the research question, (ii) the soundness of the statistical methodology employed, (iii) the quality of the argumentation/discourse and (iv) the editorial quality of the final deliverable.

This assessment method is motivated by the goal to put students into the setting of real business/institutional activities and to make their soft skills (e.g. organizational, communicative) and creativity emerge.

Textbooks and Reading Materials

Computers Ltd. What they really can't do, Harel D. Oxford University Press, 2000

Documents and slides provided by the teacher, available online

Semester

II semester

Teaching language

English

Sustainable Development Goals

NO POVERTY | GOOD HEALTH AND WELL-BEING | REDUCED INEQUALITIES | PEACE, JUSTICE AND STRONG INSTITUTIONS