UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Data Analytics

2122-2-F1801Q104

Aims

The main objectives of the course concern the acquisition of knowledge related to the main techniques of data analysis, considering both structured and unstructured data, developing specific skills regarding data, networks and social media analytics. Students will also be able to design and deploy applications for the development of specific analytics functionalities (for example, python and R scripts).

Contents

Introduction to data analytics, with particular reference to different types of data and different analysis techniques. Focuses on specific data analytics contexts such as network analytics and social media analytics.

Detailed program

1 The data analytics process. Data types: structured, semi-structured and unstructured data. Types of analysis: predictive, prescriptive, descriptive

2 Introduction to Network Analysis: basic definitions and application domains

3 Static and dynamic networks: statistical properties for structural analysis

4 Centrality measures

5 Graph clustering: algorithms and applications

6 Models of influence and contagion in social networks

7 Introduction to Social Media Analytics

8 Subjectivity Detection, Sentiment Mining, Irony Detection: lexicons, probabilistic language models, ensemble learning

9 Named-Entity Recognition and Linking: Conditional Random Fields, Probabilistic Topic Models, Word Sense Disambiguation

10 Social Media Tagging and Summarization: Tag Generation, Tag analysis, Summarization

11 Open issues and visualization techniques

Prerequisites

No essential prerequisite. Basic knowledge of linear algebra and graph theory is helpful.

Teaching form

Classroom lectures, exercises and laboratory activities.

During the Covid-19 emergency period, the lessons will take place completely remotely synchronously and asynchronously with some physical presence events. All the lessons will be recorded and made available to the students.

Textbook and teaching resource

Albert-László BARABÁSI. **Network science**. Cambridge university press (2016).

D. Easley, J. Kleinberg. **Networks, Crowds, and Markets: Reasoning About a Highly Connected World**. Cambridge University Press (2010).

Charu C. Aggarwal . **Social Network Data Analytics**, Springer (2012).

N. Danneman, R. Heimann. Socia Media Mining with R. (2014).

Semester

Second semester.

Assessment method

Team project (with oral presentation) and oral exam. No intermediate tests.

The project will consist of the development of an analytics tool based on methods and models presented in class. The project evaluation consists of a numerical evaluation expressed in a range of 0-24. For the evaluation of each project, the following points will be evaluated:

- Adopted methodology (models and methods): 7 points

- Demonstrator and visualization techniques adopted: 5 points

- Analysis of experimental results: 7 points

- Presentation: 5 points

The oral exam includes 4 theoretical questions among the course topics listed in the detailed program. For each question, a score equal to -2 will be given to a wrong response or a missing answer, and a score equal to +2 for a correct answer.

During the Covid-19 emergency period, exams will take place completely remotely.

Office hours

On appointment.

During the Covid-19 emergency period, the meetings will take place completely remotely.