UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Web and Social Networks Search and Analysis

2324-3-E311PV030

Aims

The course aims to provide students with the main concepts behind the management of data originating from the Web and social media, from collection, to modeling, to subsequent analysis. Students will be able, in particular, to retrieve and store data from the Social Web, either through the use of APIs or scraping, to use advanced representations (both topological and semantic), and to analyze and visualize complex network structures and related analyses. Part of the course will focus especially on the concepts of "Web search" and "social search" and the study of the most appropriate models and dimensions of relevance in the context of the Social Web.

Contents

- Introduction to the Social Web;
- Data in the Web and social media;
- Representing complex network structures: graph and network theory;
- Social Network Analysis (SNA): metrics and algorithms for community detection;
- Social Content Analysis (SCA): subjectivity and sentiment analysis, irony detection, Named-Entity Recognition (NER) and linking;
- Web and Social search: IR models in the Social Web, relevance dimensions, evaluation of search results;
- Visualization of data retrieved from the Web and social media and analysis of such data.

Detailed program

1. Introduction

- Introduction to the Web (1.0, 2.0), Semantic Web and Social Web and the terminology used;

- The "information objects" in the Social Web;
- Web and Social Media Analytics: definition and objectives, the concepts of self-presentation and self-revelation, implicit and explicit incentives.

2. Data in the Social Web.

- The main platforms, data types, programming interfaces, the process of crawling and scraping.
- Pre-processing and storage of Social Web data.
- Hints of data collection issues, both from a legal (the GDPR) and technological perspective.

3. Representing complex online data structures: graph and network theory

- Elementary and complex data structures;
- Representation of network structures using graphs (graph theory, types of networks).

4. Social Network Analysis (SNA)

- Link analysis, Web link analysis, main metrics;
- Network clustering: community detection algorithms;
- Influence and contagion models in social networks.

5. Social Content Analysis (SCA)

- Introduction to Natural Language Processing concepts in the context of social networks;
- Objectivity/subjectivity, polarity, emotion and irony in social networks;
- Lexical and semantic approaches;
- Named-Entity Recognition and Linking.

6. Web and Social search

- The main IR models in the Social Web;
- Dimensions of relevance;
- The evaluation of search results.

7. Visualization of data retrieved from the Social Web and analysis of such data

- Graphical interfaces;
- Usability and user studies.

Prerequisites

Basic skills in linear algebra and programming are required.

Teaching form

- Frontal lectures;
- Classroom exercises;
- Laboratory exercises.

Teaching will be given in English.

Textbook and teaching resource

Slides, insights and suggested readings during the course will be provided.

Suggested books:

- Greenlaw, R., & Hepp, E. (2001). Inline/online: fundamentals of the internet and the world wide web. McGraw-Hill, Inc.
- Rahman, M. S. (2017). Basic graph theory (Vol. 9). Cham: Springer.
- Knoke, D., & Yang, S. (2019). Social network analysis. SAGE publications.
- Liu, B. (2020). Sentiment analysis: Mining opinions, sentiments, and emotions. Cambridge university press.
- Ledford, J. L. (2015). Search engine optimization bible (Vol. 584). John Wiley & Sons.

Semester

Second semester

Assessment method

Written exam with exercises and open questions
The written exam aims at extensive and intensive assessment of theoretical and theoretical-practical skills acquired during the course.

Group project (with oral presentation)
The project aims to assess the students' ability to translate the skills acquired during the course into real application areas through the development and use of technological solutions for data analysis and retrieval in the Social Web.

Global evaluation

- Written exam is graded on a scale of 0 to 24
- Students must score greater than or equal to 12 in the written test;
- The project, with associated oral discussion, is graded on a scale of 0 to 8;
- The final grade will be given by the sum of the grade obtained in the written exam and the grade related to the project.

Office hours

By appointment: Marco Viviani (marco.viviani@unimib.it)

Sustainable Development Goals

GOOD HEALTH AND WELL-BEING | QUALITY EDUCATION | GENDER EQUALITY