UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Social Media Analytics

2122-2-F9101Q021

Aims

At the end of the course, the student will have learned the main concepts underlying the management of data originating in social media (access, pre-processing, modeling) and their subsequent analysis. The student will be able, in particular, to collect, process and analyze data from the main social media, using the most suitable technologies for the purpose. S/he will also be able to provide a representation of complex social structures in order to extract useful information.

Contents

1. The Social Web: introduction and related terminology;
2. Retrieving data in social media.
3. Representing complex social data structures (graph theory and networks);
4. Description of some open issues (information diffusion and credibility);
5. Social Network Analysis;
6. Subjectivity and sentiment analysis, irony detection;
7. Named Entity Recognition and Linking;
8. Visualization of social media data: open issues and techniques.

Detailed program

PART I

1. Introduction

- The Social Web: introduction and related terminology: Social Web, Social Networks, Social Media, User-Generated Content (UGC).
- Different classes of social “information objects”: 1) texts: posts, blogs, microblogs, 2) images, 3) audios, 4)

videos.
- Social Media Analytics: definition and objectives. The concepts of auto-presentation and auto-disclosure.

2. Retrieving data in social media

- Main social platforms, Advanced Programming Interfaces (API)s, the “crawling” process.
- Pre-processing and storing of social data.
- Outline of legal issues related to data collection (the GDPR).

3. The social data representation issue

- Elementary and complex data structures.
- Representing social structures via a graph-based representation (graph theory, networks, network analysis: link analysis, web link analysis, centrality measures).

4. Description of some open issues

- Information tracking over time.
- Information credibility of online information.

PART II

5. Social Network Analysis

- Network clustering: community-detection algorithms.
- Influence and contagion in social media.

6. Subjectivity and Sentiment Analysis, Emotions and Irony detection

- Lexicons.
- Supervised e semi-supervised models.
- Natural language processing techniques.

7. Named-entity Recognition and Linking

- Probabilistic graphical models for named entity recognition.
- Linking methods: learning vs retrieval.
- Word Sense Disambiguation.

8. Visualization of social media data: open issues and techniques

Prerequisites

Basic knowledge of the principles of linear algebra, statistics, programming, relational databases, and NoSQL.

Teaching form

- Lectures
- Exercises
- Laboratories

- The course will be held in English

Textbook and teaching resource

- Rahman, Md Saidur. Basic graph theory. Springer, 2017.
- Reinhard Diestel. Graph Theory. Springer, 2017.
- Suliman Hawamdeh, Hsia-Ching Chang. Analytics and Knowledge Management. CRC Press. 2018.
- John Scott. Social Network Analysis. Sage, 2012.
- Bing Liu. Sentiment Analysis and Opinion Mining. Morgan & Claypool. 2016.
- Eneko Agirre, Philip Edmonds. Word Sense Disambiguation: Algorithms and Applications (Text, Speech and Language Technology). Springer. 2007.

Semester

First semester

Assessment method

- **Written exam** with exercices and open questions.
    - The written test aims at the extensive and intensive evaluation of the theoretical and theoretical-practical skills acquired during the course.
- **Group project** (with oral presentation).
    - The project aims to assess students' ability to translate the skills acquired during the course into real application fields, through the development and use of technological solutions for analyzing social media data.

- The written exam is assessed on a scale from 0 to 24.
- Students must obtain a grade of 12 or more in the written exam.
- The project, with relative oral discussion, is evaluated on a scale from 0 to 9.

The final grade will be given by the sum of the evaluation obtained in the written exam and by the evaluation related to the project.

Office hours

By appointment:

- Elisabetta Fersini (elisabetta.fersini@unimib.it)
- Marco Viviani (marco.viviani@unimib.it)