AI Seminar (3 ECTS, optional course, Artificial Intelligence Master)

Academic course 2015-16

Topic: *Social media content analysis*

Seminar coordinator: Dr Antonio Moreno, URV. Specialists in the field will participate in the seminar as invited lecturers.

Dates: May 9th-13th 2016, 15:00-19:00 (with a short break around 17:00). Lecturers of other optional courses in the Master will be informed of the seminar date so that they can reserve this week for the seminar.

Place: Classroom A5101, UPC Campus Nord, Barcelona.

Brief description:

The evolution of the Web 1.0 towards the Social Web has led to an enourmous generation of content through social media. The analysis of this information on real time may provide valuable insights on many different domains: relevance of current events, detection of hot topics, public opinions on products/services, spatial movement of users, disease tracking, etc. The application of Artificial Intelligence techniques to this kind of analysis, in particular those based on semantic information, is currently a hot research topic. This seminar, that will combine theoretical lectures with hands-on sessions in which students will interact with analytic tools, will provide a general introduction to the area of semantic analysis of textual content on social media, and several invited speakers that are specialists in this area will describe different techniques used in the field and the specific tools and systems they have developed.

Seminar sessions:

- **Monday, May 9th - Dr. Diana Maynard, Univ. of Sheffield, UK**

  An introduction to social media content analysis.

  This session will provide an introduction to the task, challenges and solutions for the automatic processing of social media with Natural Language Processing (NLP) tools. After a short introduction to NLP and social media, we will cover tools to processing such content, discuss available evaluation datasets, show some examples of real applications, and outline remaining challenges. Each section of the tutorial will contain practical exercises for the participants to try out examples for themselves and see the results: for example, experimenting with different methods, tools and resources for the same task to see how the results differ. The session will consist of four main sections: challenges of social media annotation, tools for social media analysis, semantic annotation of social media and practical applications.

- **Tuesday, May 10th – Dr. Antonio Moreno, Univ. Rovira I Virgili (URV)**

  Topic detection on Twitter. Semantic analysis of brand communication through social media.

  In this session, we will explore two hot topics of research. The first one is the automatic, unsupervised identification of the main topics within a set of tweets. Three families of methods (probabilistic, document-pivot and feature-pivot) will be presented. The second one is the analysis of how destinations communicate their brands through social media and how tourists (as well as local people) perceive this information.

- **Wednesday, May 11th – Dr. Javier Béjar, Barcelona Tech (UPC)**

  Application of Artificial Intelligence techniques to the spatial and temporal analysis of Location-Based Social Networks.

  Location-Based Social Networks are a source of information about the behaviour of users inside constrained geographical areas likecities or metropolitan areas. Different Machine Learning techniques can be used to discover recurrent patterns that combine spatial and temporal behaviour. In this session of the seminar we will focus on unsupervised algorithms including clustering, social networks analysis, association rules and frequent subgraph discovery. As an application of the described techniques we will explore different datasets extracted from Twitter and Instagram for several European cities.

- **Thursday, May 12th - Dr. Patty Kostkova, University College London and City University London, UK**

  Analysis of health-related information on Twitter.

  Twitter, crowdsourcing and other medtech inventions producing real-time geo-located stream of personalized data using mobile technologies have changed the way we think about health. Twitter's strength is its two-way communication nature – both as a health information source but also as a central hub for the creation and dissemination of media health coverage. Health authorities, insurance companies, marketing agencies and individuals can leverage the availability of large datasets from Twitter to improve early-warning and preparedness, aid disease prevalence mapping and provide personal targeted health advice, as well as influence public sentiments about major health interventions. However, despite the growing potential there are still many challenges to address to develop robust and reliable systems integrating Twitter streams to real-world provision of healthcare. This seminar will combine a theoretical and practical part allowing students to work in groups and apply the knowledge on examples from health.

- **Friday, May 13th - Dr. Eugenio Martínez, Univ. of Jaén**

  Sentiment analysis in Twitter.

  People have to make decisions every day. The demand of the experience of other people through the request of their opinion is a common tool to ease the decision making process. An opinion is the expression of a personal state, which can be related to a thought, a view concerning a specific topic, a past experience or even an emotional state.

  The NLP task that is concerned with the processing of opinions is Sentiment Analysis. Its aim is is the identification, processing, classification and visualization of opinions published on different media. The advent of Web 2.0 and specifically social networks has favoured the publication and the access to opinions about any topic. Twitter is a microblogging social network in which opinions about topics take place in real time. The live character of Twitter is its strength and the source of the relevance of the task of analysing and extracting the information about the opinions that are constantly posted.

  This session will cover the fundamentals of Sentiment Analysis; moreover, a practical exercise will be developed so that students may apply the knowledge acquired in the first part of the session.

Evaluation: the students enrolled in the Master will have two/three weeks after the seminar to prepare a short individual (or in pairs, if there is a high number of enrolled students) written report on a question related to the content of the seminar.