

NLP for Data Science in Context

- Topic Course: Computational Models of Social Meaning<sup>1</sup> -

Instructor: Smaranda Muresan

Research Scientist

Center for Computational Learning Systems

This applied Natural Language Processing course will focus on computational methods for extracting social and interactional meaning from large volumes of text and speech (both traditional media and social media). Topics will include:

1. **Sentiment Analysis:** automatic detection of people's sentiment towards a topic, event, product, or persons. Practical applications in various domains will be discussed (e.g., predicting stock market prices, or presidential elections)
2. **Emotion and Mood Analysis:** automatic detection of people's emotions (angry, sad, happy) by analyzing various media such as books, emails, lyrics, online discussion forums. Practical applications in various domains (such as predicting depression, categorization of songs)
3. **Belief Analysis and Hedging:** automatic detection of people's beliefs (committed belief and non-committed beliefs) from social media. Analysis of the use of hedging as a communicative device in various media: online discussions, scientific writing or legal discussions.
4. **Deception Detection** (e.g., detecting fake reviews online, or deceptive speech in court proceedings)
5. **Argumentation Mining:** automatic detection of arguments from text, such as online discussion or persuasive essays. Practical application for various domains (e.g., political, legal or education (e.g., improving students' skills in writing persuasive essays)
6. **Social Power:** automatic detection of power structure in organizations by analyzing people's communications such as emails.
7. **Extracting Social Networks** from text, such as networks of characters from novels, or networks from social media (e.g., people holding particular opinions, or network of friends).
8. **Personality and Interpersonal Stance**

The class consists of lectures and discussion of research papers led by students. Each topic will include 2-3 research papers for discussion. Students will learn state-of-the art approaches for the topics covered in class, understand fundamental concepts in linguistics and natural language processing and acquire skills for successful completion of research projects, including writing of research papers.

Grading will be based on data analysis assignments (10% of grade), discussion of research papers (30% of grade) and final project (60%).

---

1