Course Description

Text as Data focuses on applications of natural language processing, guided by psychological theories, for identifying various social and cognitive properties evident in human related big data. In this course, we will survey state-of-the-art techniques, and applications of such techniques, for investigating various aspects of human cognition. The intended audience for this course is psychology and computer science PhD students, and more broadly graduate students in social sciences, who are interested in using machine-learning techniques for analysis of data. Also, this course may be of interest to PhD students in communications and the business school.

Learning Objectives

This course is designed to survey current state of research in automated analysis of language within the domain of psychology. It should be noted that the purpose of this class is not to teach text analysis, nor social psychology, but to survey how the established methods are used within the social sciences. Optional reading material will be provided for students unfamiliar with topics discussed.

- **Prerequisite(s):** Instructor permission

- **Recommended Preparation:** For Non-Engineering majors: Psych 625 or a similar course, for Engineering students: CSCI 544 or a similar course.

Course Notes

Students are not allowed to use laptops or smartphones during class, unless used for class presentation. Homework assignments will be posted on Blackboard.
Required Books


Description and Assessment of Assignments

1. Paper presentation. Each student will present a set of papers related to one of the topics discussed in class.

2. Reaction paragraphs. Students are asked to write a short note, one or two paragraphs in length, about their reaction to the reading assignments of the week. These can be a quick summary of the material, comments about the subject area, or a critique of a particular theory or experiment. I will read these paragraphs carefully before each class, and will use them to guide the discussion in class. Simply reading the first page of a paper and writing a summary of it will not count as a reaction paragraph.

3. Class Project. This class is project oriented, and group-based. The goal of the project is for students to get experience working in interdisciplinary groups to tackle specific social scientific problems, and bring together theory from the social sciences and NLP techniques from computer science to tackle that problem. This will include a project proposal presentation, three project update presentations, final project presentation, and a report. For project proposals, students will present a problem and a data collection method and/or dataset for which they want to analyze. Each presentation should be about 10-15mins. The goal of the project update presentations is to inform the class about the state of the project and brainstorm with other students on how to solve the remaining issues. Each update presentation should be around 10 minutes. For the final project presentation, each student/group will give a 15-20min presentation on their project. Students are expected to spend at least 80 hours working on their final project. The project report will be around 20 pages.

Grading Policy

- 15% Participation
- 20% Paper Presentations
- 25% Reaction Paragraphs
- 15% Project Status Updates
- 10% Final Project Presentation
- 15% Final Project Write up
Assignment Submission Policy

All assignments are due on Thursdays at 10am. Assignments turned in any later than 10:10am will be considered late. Students will be allowed a total of four late days that can be used on the assignments. In exceptional circumstances, arrangements must be made in advance of the due date to obtain an extension. Once you have used up your four late days, one additional day late will result in a 25% reduction in the total score, two additional days late will yield a 50% reduction, and no credit will be given for three or more additional days late. Late days are in units of days, not hours, so using up part of a day uses up the whole day. The final project report, plus code used, will be due on the day of the final exam.

Schedule and weekly learning goals

The schedule is tentative and subject to change.

**Week 01, 08/26:** Introduction to Computational Social Sciences 1/2


**Week 02, 09/02:** Introduction to Computational Social Sciences 2/2


**Week 03, 09/09:** Dictionary Methods 1/2

Week 04, 09/16: Dictionary Methods 2/2


- Choose two:

Week 05, 09/23: Differential Language Analysis & Project Proposals


Week 06, 09/30: Distributed Dictionary Representations


Week 07, 10/07: Neural Networks 1/2


Week 08, 10/14: Fall Recess

5/14
Week 09, 10/21:  Neural Networks 2/2


- Choose two:

Week 10, 10/28:  Bias in NLP & Project Update I


- Choose two:


**Week 11, 11/04:** Other methods


- Wang, S.-Y. N. and Inbar, Y. (2021b). Re-examining the diffusion of moralized rhetoric from political elites: Effects of valence and ideology. Under-review


**Week 12, 11/11:** Clinical & cognitive applications & Project Update II


• Choose two:
  


**Week 13, 11/18:** Ethics


• Watch in class: Friends You Haven’t Met Yet

**Week 14, 11/25:** Thanksgiving Holiday

**Week 15, 12/02:** Final project presentations
Statement on Academic Conduct and Support Systems

Academic Conduct

Plagiarism — presenting someone else’s ideas as your own, either verbatim or recast in your own words — is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Section 11, Behavior Violating University Standards https://scampus.usc.edu/1100-behavior-violating-university-standards-and-appropriate-sanctions/. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, http://policy.usc.edu/scientific-misconduct/.

Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the Office of Equity and Diversity http://equity.usc.edu/ or to the Department of Public Safety http://capsnet.usc.edu/department/department-public-safety/online-forms/contact-us. This is important for the safety whole USC community. Another member of the university community — such as a friend, classmate, advisor, or faculty member — can help initiate the report, or can initiate the report on behalf of another person. The Center for Women and Men http://www.usc.edu/student-affairs/cwm/ provides 24/7 confidential support, and the sexual assault resource center webpage sarc@usc.edu describes reporting options and other resources.

Support Systems

A number of USC’s schools provide support for students who need help with scholarly writing. Check with your advisor or program staff to find out more. Students whose primary language is not English should check with the American Language Institute http://dornsife.usc.edu/ali, which sponsors courses and workshops specifically for international graduate students. The Office of Disability Services and Programs http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html provides certification for students with disabilities and helps arrange the relevant accommodations. If an officially declared emergency makes travel to campus infeasible, USC Emergency Information http://emergency.usc.edu/will provide safety and other updates, including ways in which instruction will be continued by means of blackboard, teleconferencing, and other technology.
 References


