Psych/CSCI 626: Text as Data

Morteza Dehghani

Fall 2023

E-mail: mdehghan@usc.edu
Office Hours: Mon 10-12
Office: SGM 607

Web: mola-lab.org
Class Hours: Th 2-6pm
Class Room: SGM 911

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 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. 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, about two paragraphs in length, about their reaction to the reading assignments of the week. These can be 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

- **25%** Discussion Participation
- **20%** Paper Presentations
- **20%** Reaction Paragraphs
- **15%** Project Status Updates
- **20%** Final Project Presentation and 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.

Use of LLMs

I expect you to use LLMs in this class, mainly for making your writing more fluent, and for brainstorming. Learning to use AI is an emerging skill, but please keep in mind the following:

- AI tools are permitted to help you brainstorm topics or revise work you have already written.

- If you provide minimum-effort prompts, you will get low-quality results. You will need to refine your prompts to get good outcomes. This will take work.

- Proceed with caution when using AI tools and do not assume the information provided is accurate or trustworthy. If it gives you a number or fact, assume it is incorrect unless you either know the correct answer or can verify its accuracy with another source. You will be responsible for any errors or omissions provided by the tool. It works best for topics you understand.

- AI is a tool, but one that you need to acknowledge using. Please include a paragraph at the end of any assignment that uses AI explaining how (and why) you used AI and indicate/specify the prompts you used to obtain the results. Failure to do so is a violation of academic integrity policies.

- Be thoughtful about when AI is useful. Consider its appropriateness for each assignment or circumstance. The use of AI tools requires attribution. You are expected to clearly attribute any material generated by the tool used.

Schedule and weekly learning goals

The schedule is tentative and subject to change.

Week 01, 08/24: Introduction to Computational Social Sciences


Week 02, 08/31: Jamie


Week 03, 09/07: Morteza’s Favorites


Week 04, 09/14: NLP + Experimentation


**Week 05, 09/21: Debates**


• Youyou, W., Yang, Y., and Uzzi, B. (2023a). A discipline-wide investigation of the replicability of psychology papers over the past two decades. *Proceedings of the National Academy of Sciences*, 120(6):e2208863120


**Week 06, 09/28: Large Language Models I: Potential for Social Sciences**


**Week 07, 10/05:** Large Language Models II: Cognitive abilities


• Crockett, M. and Messeri, L. (2023). Should large language models replace human participants?

• TBA

**Week 08, 10/12:** Fall Recess

**Week 09, 10/19:** Large Language Models III & Project Proposal Presentations


**Week 10, 10/26: Large Language Models IV: Subjectivity**


• TBA (2023). Shaping the perspective: Evaluating chatgpt’s capability in subjective annotation tasks. *TBA*

**Week 11, 11/02: Clinical & cognitive applications & Project Update I**

• Cognitive papers:


• Clinical papers:


**Week 12, 11/09:** Bias in NLP


**Week 13, 11/16:** Ethics & Project Update II


• Watch in class: Friends You Haven’t Met Yet
Week 14, 11/23:  Thanksgiving Holiday

Week 15, 11/30:  Final project presentations
Statement on Academic Conduct and Support Systems

Academic Integrity

The University of Southern California is a learning community committed to developing successful scholars and researchers dedicated to the pursuit of knowledge and the dissemination of ideas. Academic misconduct, which includes any act of dishonesty in the production or submission of academic work, compromises the integrity of the person who commits the act and can impugn the perceived integrity of the entire university community. It stands in opposition to the university’s mission to research, educate, and contribute productively to our community and the world.

All students are expected to submit assignments that represent their own original work, and that have been prepared specifically for the course or section for which they have been submitted. You may not submit work written by others or “recycle” work prepared for other courses without obtaining written permission from the instructor(s).

Other violations of academic integrity include, but are not limited to, cheating, plagiarism, fabrication (e.g., falsifying data), collusion, knowingly assisting others in acts of academic dishonesty, and any act that gains or is intended to gain an unfair academic advantage.

The impact of academic dishonesty is far-reaching and is considered a serious offense against the university. All incidences of academic misconduct will be reported to the Office of Academic Integrity and could result in outcomes such as failure on the assignment, failure in the course, suspension, or even expulsion from the university.

For more information about academic integrity see the student handbook or the Office of Academic Integrity’s website, and university policies on Research and Scholarship Misconduct.

Please ask your instructor if you are unsure what constitutes unauthorized assistance on an exam or assignment, or what information requires citation and/or attribution.

Students and Disability Accommodations

USC welcomes students with disabilities into all of the University’s educational programs. The Office of Student Accessibility Services (OSAS) is responsible for the determination of appropriate accommodations for students who encounter disability-related barriers. Once a student has completed the OSAS process (registration, initial appointment, and submitted documentation) and accommodations are determined to be reasonable and appropriate, a Letter of Accommodation (LOA) will be available to generate for each course. The LOA must be given to each course instructor by the student and followed up with a discussion. This should be done as early in the semester as possible as accommodations are not retroactive. More information can be found at osas.usc.edu. You may contact OSAS at (213) 740-0776 or via email at osasfrontdesk@usc.edu.

Support Systems

Counseling and Mental Health - (213) 740-9355 - 24/7 on call
Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.
988 Suicide and Crisis Lifeline - 988 for both calls and text messages - 24/7 on call
The 988 Suicide and Crisis Lifeline (formerly known as the National Suicide Prevention Lifeline) provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week, across the United States. The Lifeline is comprised of a national network of over 200 local crisis centers, combining custom local care and resources with national standards and best practices. The new, shorter phone number makes it easier for people to remember and access mental health crisis services (though the previous 1 (800) 273-8255 number will continue to function indefinitely) and represents a continued commitment to those in crisis.

Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-9355(WELL) - 24/7 on call
Free and confidential therapy services, workshops, and training for situations related to gender- and power-based harm (including sexual assault, intimate partner violence, and stalking).

Office for Equity, Equal Opportunity, and Title IX (EEO-TIX) - (213) 740-5086
Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

Reporting Incidents of Bias or Harassment - (213) 740-5086 or (213) 821-8298
Avenue to report incidents of bias, hate crimes, and microaggressions to the Office for Equity, Equal Opportunity, and Title for appropriate investigation, supportive measures, and response.

The Office of Student Accessibility Services (OSAS) - (213) 740-0776
OSAS ensures equal access for students with disabilities through providing academic accommodations and auxiliary aids in accordance with federal laws and university policy.

USC Campus Support and Intervention - (213) 740-0411
Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity, Equity and Inclusion - (213) 740-2101
Information on events, programs and training, the Provost’s Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency - UPC: (213) 740-4321, HSC: (323) 442-1000 - 24/7 on call
Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety - UPC: (213) 740-6000, HSC: (323) 442-1200 - 24/7 on call
Non-emergency assistance or information.

Office of the Ombuds - (213) 821-9556 (UPC) / (323-442-0382 (HSC)
A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.
Occupational Therapy Faculty Practice - (323) 442-2850 or otfp@med.usc.edu
Confidential Lifestyle Redesign services for USC students to support health promoting habits and routines that enhance quality of life and academic performance.