CS 4973/6983: Research in Human-Centered NLP Syllabus

Course Information

Computer Science 4973/6983
Research in Human-Centered Natural Language Processing (NLP)
Spring 2024
4 credits

Dates: January 6, 2025 - April 15, 2025
Time/Date: Mon & Wed 2:50 PM- 4:30 PM
Location: Forsyth Building, Room 129

Canvas: https://northeastern.instructure.com/courses/202331

Course Prerequisites: While previous experience with specific technologies is not required, being proficient and comfortable with extensive programming is a fundamental prerequisite for this course. Having *some* prior experience with Natural Language Processing (NLP) or Human-Computer Interaction (HCI), through courses and/or applied projects, is preferred but not required. The coursework will cover a crash course on NLP and the basics of using Large Language Models (LLMs). Previous research experience is not expected for this course. If you are not comfortable with programming, please talk with the instructor.

Instructor Contact Information

Course Instructors

Instructor: Hye Sun Yun (she/her pronouns)
Office: EXP 760 & 177 Huntington 22nd Floor
Email address: yun.hy@northeastern.edu
Website: https://www.hyesunyun.com/
You may call me "Hye Sun" (pronounced like hey sun!) in writing and when we talk.
Feel free to contact me either in-person during office hours or via email. I will respond to your emails within 24 hours.

Student Hours (Office Hours):
Fridays 11:30am-1:30pm at EXP 760
or by appointment (email me to set up a time)

Join your instructor and peers to discuss the material being covered in class, questions or concerns you might have, and other related issues. Please join even if you don't have any questions. You can listen in on the conversation (which might spark questions for you), or we can use the time to get to know each other. If these hours do not work with your schedule, please let me know and I will try to work out a time to meet you either in person or Zoom.

Course Description

Welcome to Research in Human-Centered Natural Language Processing (NLP)!

This course aims to provide a human-centered perspective toward modern language technologies and highlights research topics shared between Human-Computer Interactions (HCI) and Natural Language Processing (NLP). In this course, you will learn and apply research methods for designing and developing human-centered NLP applications and systems. As part of the course, you will read research papers at the intersection of NLP and HCI and practice critically evaluating them to identify future research directions. The course should help students gain the skills and confidence to pursue research at the intersection of HCI and NLP.

The structure and content of this course has been inspired by Sherry Tongshuang Wu's Human-Centered NLP course, and Diyi Yang's Human-Centered LLMs course.

Course Learning Goals

Through active engagement and completion of course activities, you will be able to:

1. identify how standard NLP research fails at meeting user needs, preferences, and requirements.
2. practice common research methods for designing and evaluating technology and user interactions.
3. design and develop a user-centric language technology or an interactive tool in a capstone project.
4. critically evaluate research in the intersection of NLP and HCI through reading and presenting research papers in class.

Prerequisites

I welcome everyone who is passionate about Human-Centered NLP.

The recommended pre-requisites are either a course on HCI or NLP or any equivalents. However, this is just a recommendation and not having this prerequisite does not prevent you from taking this course.

However, all students are expected to:

1. be familiar with Python or similar programming language (mostly for completing the capstone project)
2. know basic ML/NLP concepts – such as train/dev/test set, features, supervised learning, etc.

If it is unclear if you have the right background for this course, please talk to the instructor.

Course Structure

This course will be facilitated through a combination of in-person class meetings and preparatory material posted on Canvas.
Our Canvas website is: https://northeastern.instructure.com/courses/202331. All readings, slides, and assignments will be posted on Canvas. Announcements will also be made through the Canvas site, and any questions should be posted to the **Discussions** tab. Homework assignments, project proposals, and papers should be submitted via the **Gradescope** tab.

- **Lectures**: Our class will meet in person, but I will allow attendance by Zoom if circumstances prevent you from attending in person. Since class sessions will be highly interactive, they will not be recorded, but all slides will be posted on Canvas. Class sessions will be a combination of mini-lectures and in-class activities where you will discuss & solve problems with your classmates.
- **Research Paper Presentations & Discussions**: For the second half of the course, we will focus our attention on reading and critically evaluating research papers in HCI and NLP. Before each class session, I expect everyone to read the assigned research paper and prepare for the discussion. During class, one student will present the research paper, and then we will proceed with class discussions to deepen our understanding and conversations.

Course Materials

Technology: You will need to have access to a device that connects to the internet so that you can access email, Canvas, and Gradscope. All course details and materials will be posted on our Canvas course site. Students can borrow equipment and access other learning technology from Information Technology Services. For tech support, see ITS Northeastern Tech Resources.

Readings: The primary readings from this course are drawn from textbooks on either NLP or research methods and published research papers, which are all available

electronically from Snell Library or made available in Canvas. Only select chapters (NOT full texts) from the textbook will be assigned.

- Jurafsky & Martin (2024) ***Speech and Language Processing****: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*. 3rd edition. [Link to online manuscript]
- Lazar, Feng, & Hochheiser (2017) ***Research Methods in Human-Computer Interaction***. Elsevier Science & Technology [Link to EBook via Snell Library]
- All other assigned readings can be found in Reading Materials for each Module on Canvas

Coursework and Grading Scheme

The class schedule in the syllabus highlights assignments and due dates so that you can plan your work for the semester accordingly. The table below summarizes the kinds of work you will be asked to do and their contributions to your course grade, on either grading scheme. I anticipate the readings taking 1 to 3 hours per week while each assignment taking anywhere from 2 to 4 hours.

Subject to change, based on student feedback and input!
We will spend some time during the third week of the semester to discuss the grading scheme and make any adjustments based on the class input and consensus.

Component/Assignment	**% of course letter grade**	**Course Learning Goal**	**Notes**
Identifying Issues of Language Technologies: Explore real-world applications of language technologies and critically analyze their potential issues, limitations, or unintended consequences. **DUE: Jan 17**	10%	LG1	• Short report with citation
CITI human subjects research training: university's online training requirement for conducting human subjects research. **DUE: Jan 24**	5%	LG2	• Graded for completion
Evaluating Model Outputs with NLP metrics & human feedback: comparing NLP metrics with human feedback. **DUE: Feb 14**	10%	LG2	• Involves programming in Python and conducting small human subjects study

Component/Assignment	**% of course letter grade**	**Course Learning Goal**	**Notes**
Usability Testing Study Design: designing a usability testing study. **DUE: March 14**	10%	LG2	• Short report answering questions
Research Paper Presentation: present one assigned research paper during class. **DUE: March 10~April 7**	15%	LG4	
Capstone Project: involves designing, developing, or evaluating some sort of human-centered NLP system or data. **DUE: April 18**	40%	LG4	● Proposal – 5% ● Mid-progress presentation – 5% ● Final presentation – 15% ● Final Report – 15% ● ***All components must be submitted on time***
Attendance and participation in class: attendance and engagement during class (activities, questions, etc.) and discussions.	10%	-	● 2 absence 'passes', no questions asked ● Arrival >15 mins late counts as an absence

Note: we do NOT have a final or a midterm exam in this course. The final grade will be computed from the grades you receive in individual assignments.

There will be other activities and assignments that will not be graded such as lab work.

The final letter grade is calculated using the following grading schema:

Grade	**Minimum Percent**
A	94
A-	90
B+	87
B	84
B-	80
C+	77

C	74
C-	70
D+	67
D	64
D-	60
F	0

Course Schedule

Subject to change, based on student feedback and input!
Major assignments are **bolded and in blue**.

Week & Topic	Work to Complete Before Class	Class Activities	Work to Complete After Class	Holidays/ Key Dates
Week 1: Introduction to Human-Centered NLP & NLP Crash Course	Pre-course Survey Lab prep Jurafsky & Martin. "Chapter 3: N-gram Language Models." *Speech and Language Processing* (2024) pp. 1-11.	Lab: n-gram language modeling	**Assignment: CITI Human Subjects Research Training** **Assignment: Identifying Issues of Language Technologies** Paper Presentations Signup	
Week 2: NLP Crash Course Continued	Transformers (how LLMs work) explained visually (Video) Attention in transformers, visually explained (Video) – (OPTIONAL) Schulhoff et al. "The Prompt Report: A Systematic Survey of Prompting Techniques." (2024) pp 1-7.	Labs: Huggingface (LLM inference & fine-tuning)		

	Lab prep			
Week 3: Design Thinking for Innovative Solutions	Abras et al. "User-centered design." *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications* (2004) Friedman et al. "Value Sensitive Design and Information Systems." (2013) pp 55-61 and 72-79. Lazar et al. "Chapter 8 Interview & Focus Groups." *Research Methods in Human Computer Interaction* (OPTIONAL)			MLK Day (Monday)
Week 4: Collecting, Curating, & Preparing Data	Bommasani et al. "On the opportunities and risks of foundation models." (2021) pp 101-104. Sambasivan et al. "Everyone wants to do the model work, not the data work": Data Cascades in High-Stakes AI." (2021)		**Assignment: Capstone Project Proposal**	University deadline: Last day to add/drop a course (Monday)
Week 5: Evaluating Models with Human Feedback	Schuff et al. "How to do human evaluation: A brief introduction to user studies in NLP." (2022) Lab setup Lazar et al. "Chapter 4 Statistical Analysis." *Research*	Lab: R for statistical analysis	**Assignment: Evaluating Model Outputs with NLP metrics & human feedback**	

	Methods in Human Computer Interaction (OPTIONAL)			
Week 6: Building Human-AI Systems & Interactions	Gao et al. "A Taxonomy for Human-LLM Interaction Modes: An Initial Exploration." (2024) Jakob Nielsen "Ten Usability Heuristics" (2005) https://lawsofux.com/ (OPTIONAL)	Lab: prototyping interfaces		
Week 7: Evaluating Human-AI Interactions	Lazar et al. "Chapter 10 Usability Testing "*Research Methods in Human Computer Interaction* (pp 263-265, 271-293; 10.1-10.2 & 10.5) Lee et al. "Evaluating Human-Language Model Interaction." (2022) pp 1-12 & pp 29-32. (OPTIONAL) Lazar et al. "Chapter 3 Experimental Design" *Research Methods in Human Computer Interaction*. pp 45-55. (OPTIONAL)		**Assignment: Usability Testing Study Design**	Presidents' Day (Monday)
Week 8: Capstone Project Progress & Learning to Read/Present Research Papers	Fong. "Reading a computer science research paper" (2009) Keshav. "How to read a paper." (2007)	**Capstone Project Progress Presentation**		
Spring Break (No Class)				Spring Break

Week 9: Human Feedback for Model Alignment	Wang et al. "Putting humans in the natural language processing loop: A survey." (2021) Ouyang et al. "Training language models to follow instructions with human feedback." (2022)	Research Paper Presentation & Discussions		
Week 10: Human-AI Collaboration (Writing & Productivity)	Lee et al. "Coauthor: Designing a human-ai collaborative writing dataset for exploring language model capabilities." (2022) Nguyen et al. "How Beginning Programmers and Code LLMs (Mis)read Each Other." (2024)	Research Paper Presentation & Discussions		
Week 11: Inclusive Language Technologies	Naouset al. "Having beer after prayer? measuring cultural bias in large language models." (2023). Ryan et al. "Unintended Impacts of LLM Alignment on Global Representation." (2024) Yin et al. "Including Signed Languages in Natural Language Processing." (2021)	Research Paper Presentation & Discussions		
Week 12: Safety, Privacy, & Trust	Mei et al. "ASSERT: Automated Safety Scenario Red Teaming for Evaluating the Robustness of Large Language Models."	Research Paper Presentation & Discussions		

	(2023) Zhang et al. "It's a Fair Game", or Is It? Examining How Users Navigate Disclosure Risks and Benefits When Using LLM-Based Conversational Agents." (2024) Zhou et al. "Rel-A.I.: An Interaction-Centered Approach To Measuring Human-LM Reliance" (2024)			
Week 13: Agents & Robots/Work on Capstone Project	Kim & Lee et al. "Understanding Large-Language Model (LLM)-powered Human-Robot Interaction" (2024)	Research Paper Presentation & Discussions		
Week 14: Capstone Project		**Capstone Project Presentation**	**Capstone Project Paper Submission**	

Course Expectations

What you can expect from me

I am here to guide your learning and will challenge you to actively engage in the learning process through class activities, assignments, and more. I will do my best to give you the tools, feedback, and support to succeed, and will always welcome suggestions for improvement. Learning is a never-ending process, so I hope to motivate students to seek out more information on topics we don't have time to cover. I highly encourage everyone to visit me during student (office) hours or to set up a meeting, even if you don't feel that you have questions. I want to get to know you and support you in this learning experience! The best way to reach me is by email (see contact information), and you can expect me to respond within 24 hours.

What I expect from you

I expect you to take an active role in your learning by coming to our class meetings prepared and ready to collaborate with your classmates. Keep in mind that each

member of this class has different ideas and perspectives that will enrich the experience for us all. I expect all of us to speak and listen with compassion and not make assumptions about others. I expect what happens in class to stay in class, to protect the privacy of class members and other rights of the university. Never hesitate to email me, join me in my office hours, or set up a meeting. This class should challenge you, but I believe everyone has the ability to succeed with effort and dedication.

Best Practices for Course Learning

Learning how to learn effectively is a skill unto itself! To get the most out of our course learning experience, I recommend the following:

- If possible, attend all class meetings and be fully present and engaged.
- Take notes on what you read, either directly in the text or somewhere else, and jot down all the questions you have.
- Ask questions! During class, during student (office) hours, and asynchronously over email or Canvas.
- Feel free to collaborate with your peers for assignments and capstone projects. Collaboration is an integral part of conducting research and an important skill to learn and practice.

Course Policies

Attendance and Participation

Thinking through difficult issues is most productive and most fun when done out loud in the company of others! This is why attendance and participation are part of your course grade (10%).

- You get **2 'passes'** to be absent from class or section, no questions asked.
- For additional absences, please let me know ahead of time and discuss ways of covering the material you might have missed.
- Even if an absence is excused, any work due after that class/section is subject to the policies below.
- Students who anticipate persistent challenges to participating in class or submitting work on time should share this with me as soon as possible.

Deadlines and Late Submissions

Assignment deadlines are listed in the course schedule to enable you to effectively plan and balance your academic work and other commitments. Despite the best planning, however, we know that life happens! So:

- You get **2 'passes'** to submit a non-capstone project assignment up to 48 hours late, no questions asked.
- Beyond that, your grade for an assignment goes down a third of a letter grade

(e.g., A- to B+) for every 24 hours the assignment is late.

Language Model Use Policy

The goal of this course is to think critically about current computing research on improving our use of language technologies. Language models (LMs) and chat assistants (e.g., ChatGPT) are some of the technologies we will be discussing in class. For that reason, we are allowing the use of LMs to assist you in completing assignments. However, the use of any LM comes with the following requirements:

- The output of the LM **cannot be the final output** that you submit for the assignment.
- You must **disclose** the use of the LM on the assignment.
- When you disclose LM use, also include 2-3 sentences **describing** why you used the LM, what you used it for, what you found helpful about using the LM, and what were limitations you had to work around.

NU Academic Integrity

Northeastern University is committed to the principles of intellectual honesty and integrity: the NU Academic Integrity Policy can be found on the website of the Office of Student Conduct and Conflict Resolution (OSCCR).

Academic Accommodations

If a student has a disability that needs accommodations in the course (e.g., accessible learning materials, extended time testing, etc.), please file a disability accommodation request with the Northeastern University Disability Resource Center.

The university's Disability Resource Center works with students and faculty to provide students who qualify under the Americans With Disabilities Act with accommodations that allow them to participate fully in the activities at the university. Ordinarily, students receiving such accommodations will deliver teacher notification letters at the beginning of the semester. Students have the right to choose whether to disclose their specific disabilities to instructors but must provide a letter to receive accommodations.

Student athletes who anticipate challenges in being able to participate in class or submit assignments on time should speak to with me as soon as possible about available alternatives or allowances.

Resources

Title IX Protections and Resources

Title IX of the Education Amendments of 1972 protects individuals from sex or gender-based discrimination, including discrimination based on gender-identity, in

educational programs and activities that receive federal funding. Any NU community member who has experienced such discrimination, sexual assault, relationship violence, stalking, coercion, and/or sexual harassment, is encouraged to seek help. Confidential support and guidance can be found through University Health and Counseling Services, the Northeastern Center for Spirituality, Dialogue, and Service, and the Office of Prevention and Education at Northeastern (OPEN). Note that faculty members are considered "responsible employees" at Northeastern University, meaning they are required to report all allegations of sex or gender-based discrimination to the Title IX Coordinator. For additional information and assistance please see the Office of Institutional Diversity and Inclusion webpage.

WeCare

WeCare offers supports for students during times of difficulty or challenge. You can find WeCare at 226 Curry Student center Monday - Friday from 8:30-5:00, call at 617-373-7591, or email wecare@northeastern.edu.

Mental Health Resources

In addition to mental health resources available through Northeastern's University Health and Counseling Services Northeastern has added Find@Northeastern, which is a 24/7 mental health consulting line and can be reached at 1-877-223-9477.

The Writing Center

The Northeastern University Writing Center offers free and friendly tutoring for any level of writer, including help with conceptualizing writing projects, the writing process, and using sources effectively. Currently, the Writing Center has virtual appointments only. To make an appointment, or learn more about the Writing Center, visit https://www.northeastern.edu/writingcenter, or email WritingCenter@northeastern.edu. Advance and same-day appointments are available.

Peer Tutoring

The Peer Tutoring Program offers a wide range of tutoring services to meet the academic needs of undergraduate students. If you need academic assistance, contact the Peer Tutoring Program Monday through Friday from 9:00am to 5:30pm. Peer tutoring services are free and open to all NU undergraduate students. Peer tutoring begins the second week of classes and ends the last day of classes. The Peer Tutoring

Program is in 1 Meserve Hall. Call 617- 373-8931, email peertutoring@northeastern.edu, or visit the weblink above.

International Tutoring Center

The International Tutoring Center (ITC) provides current Northeastern University international students with free, comprehensive English language and academic support. The ITC includes English as a Second Language Tutoring (ESL), Language and Culture Workshops, and Reading Workshops. For more information on available workshops and tutoring opportunities please visit the ITC weblink above.

Snell Library

Snell Library offers a variety of resources for undergraduate research, including subject-specific Research Guides, help with citation and bibliography, and 24/7 chat support. The library also houses the Digital Media Commons, which offers a variety of resources for instructors and students for multimedia projects.