

CS 333: Natural Language Processing

Class Meetings: Tu F 9:55-11:10
Sci L039

Instructor: Carolyn Anderson
carolyn.anderson@wellesley.edu

Drop-in hours: M 4-5:30
Th 4-5
[By appointment](#)
Sci W422

Tutors: Caroline Witty
Ashley Sheng

Announcements: We will use the CS 333 Google Group for course announcements.

Course Description

Natural Language Processing (NLP) is the subfield of CS that focuses on language technology. Because language is one of the most complex human abilities, building computational technologies that involve language is both challenging and important.

This course introduces NLP methods and applications. Students will (1) learn core NLP algorithms and models; (2) explore the challenges posed by different aspects of human language; (3) learn to evaluate ethical concerns about language technology; and (4) complete a series of projects to implement and improve NLP models. We will cover a range of techniques, including n-gram models, Bayesian classifiers, regression, and large language models. Applications include sentiment analysis, machine translation, and language generation, as well as information retrieval tasks like summarization, topic modeling, and question-answering.

Prerequisites(s): CS 230 and 1 from {MATH 206, MATH 220, MATH 225}

Expectations

- You are expected to attend and participate in class. However, you **should not** come to class if you are feeling sick; I'll work with you to catch up after you are feeling better.
- You are expected to seek help when needed. Sometimes students are embarrassed to seek help. I am happy to help with **any course material**, no matter when it was taught.

Assignments

There will be **9 weekly homework assignments**, typically composed of:

- short math questions to reinforce the theory behind the techniques we are studying
- programming exercises that you ask to implement NLP algorithms or models
- evaluation questions that ask you to analyze a model that you have implemented

Assignments will be due on Thursdays at 10pm. We will use Gradescope for submission.

10 points per assignment will be **curiosity points**: I will award these for demonstrating intellectual curiosity by going beyond the minimum requirements of the assignment.

Midterms

Midterm 1 will be in class on **October 10th** and Midterm 2 will be in class on **November 7th**. If you have an accommodation that calls for extra time or a separate testing environment, please let me know two weeks before so that we can make alternative arrangements.

Final Paper

You will write a final paper which will be due at the end of the semester. This is an individual research project. You will work with me to select an appropriate topic.

Class Preparedness

You are expected to come to class each week having done the weekly reading. We will have a brief quiz at the start of class each week to assess your understanding of the weekly readings, to provide accountability for coming to class on time, prepared for the week.

Readings

Our textbook is *Speech and Language Processing* by Jurafsky & Martin, 3rd ed. **The PDF is free:** <https://web.stanford.edu/~jurafsky/slp3>. Look for the 2nd edition if you prefer a print book.

Grading

- | | |
|----------------------|-----|
| • Homework | 70% |
| • Class Preparedness | 5% |
| • Final paper | 10% |
| • Midterms | 15% |

Late Policy

- Assignments will typically be due at **10pm on Thursdays**.
- **You have 5 late days**. You can use them all for one assignment, or spread them out.
- If you run out, I will deduct 5 points (**5% of the assignment grade**) for each extra day.
- **I will not answer questions on late work during help hours.**
- **If you are falling behind, email me ASAP.** We will make a plan to support your learning and well-being. **You do not need to give any personal details to justify your request.**

Generative AI Policy

- **You may not use ChatGPT, Co-pilot, Codex, or any other generative AI models on CS 333 assignments (homework, exams, final paper).** If you disagree with this policy, please talk to me! I am interested in hearing why you think this would benefit your learning.
- **You must contact me before using GenAI for any class purpose.** This includes drafting emails, creating study guides or other review materials, etc. Send me an email describing how you would like to use GenAI, and I will review the request. **Using generative AI for any class purpose without explicit permission will be considered an Honor Code violation.**

Collaboration Policy and the Honor Code

- Collaboration:

- **Wait 30 minutes after discussions with other students** before writing your solution. This helps you know if you actually understand the solution.
 - **You may not share code** from homework problems with other students **at any point in the course** (even after the assignment deadline).
- Reference:
 - You may consult external reference resources for general concepts and techniques, but you **must cite them**.
 - You must **not** consult solutions to this or similar assignments from other NLP courses, books, or online resources.
- Code reuse and adaptation:
 - **For homework assignments, you must not reuse or adapt any code beyond what is given to you in the assignment and in CS333 lectures.**
- After the semester:
 - **Do not post homework, exam, or quiz solutions publicly** (i.e. on GitHub or your website). You may show your work to potential employers or others who aren't students. You may also share your final project publicly.

Seeking Help

- **If you go 20-30 minutes without making progress, please ask for help!**
- **Come to my drop-in hours!** You don't need to have a specific question. If the posted time doesn't work for you, you can book a one-on-one meeting using [this calendar](#).
- You can request a one-on-one tutor from the PLTC.

Disabilities and Accommodations

My job is to help you learn, and part of that is making accommodations for any disabilities you might have. **Please feel free to offer suggestions about how I can make the course more accessible to you.** I will keep the details of our conversations confidential. You are welcome to share anything that will be helpful, but you should not feel pressure to share details with me.

If you would benefit from reasonable academic adjustments, I recommend contacting [Accessibility and Disability Resources \(ADR\)](#) to get a letter outlining your accommodation needs as soon as possible, as some may take significant time to arrange. If you need immediate accommodations, please meet with me ASAP. If you are unsure but suspect you may have an undocumented need for accommodations, ADR can help with referrals for assessments.

Disclosures of Discrimination and Sexual Misconduct

Wellesley College considers diversity essential to educational excellence, and we are committed to being a community in which each member thrives. The College does not allow discrimination or harassment based on race, color, sex, gender identity or expression, sexual orientation, ethnic or national origin or ancestry, physical or mental disability, pregnancy or any other protected status under applicable local, state or federal law.

If you or someone you know experiences discrimination or harassment, support is available:

- Confidential reporting:
 - Health Services (781.283.2810), the Stone Center Counseling Service (781.283.2839), and Religious and Spiritual Life (781.283.2685) are not required to report allegations of sexual misconduct to the College.

- Non-confidential reporting:
 - You can let me know. **As a faculty member, I am a mandatory reporter:** I am obligated to report allegations of sex-based discrimination to the Title IX Office.
 - You can report directly to the Nondiscrimination/Title IX Office (781.283.2451) to receive support, and to learn more about your options for reporting.
 - You can report to the Wellesley College Police (Emergency: 781.283.5555, Non-emergency: 781.283.2121) if you believe a crime has been committed.