CS 232: Artificial Intelligence

Class Meetings:

T F 9:55-11:10 Sci E101 Instructor:

Carolyn Anderson carolyn.anderson@wellesley.edu

Drop-in hours: M 4-5:15 F 3:30-4:30 By appointment Sci W422 Tutors:Lyra Kalajian & Lepei ZhaoTutor hours:TBAGrader:Jess Yao

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

Course Description

What is artificial intelligence (AI) and should humans fear it as one of "our biggest existential threats"? In this course, we will grapple with these difficult questions and investigate them in different ways. We will discuss the development of the field from the symbolic, knowledge-rich approaches of 20th century AI (e.g., rule-based systems), to statistical approaches that rely on increasingly large amounts of data, including an overview of contemporary deep learning techniques. We will explore how to apply these techniques in several AI application areas, including robotics, computer vision, and natural language processing, and consider ethical issues around AI in society. By the end of the semester, students should be able to answer the starting questions in-depth and with nuance.

Distributions: MM Prerequisites(s): CS111 and CS 230

Learning Goals

- Learn about a range of AI approaches, including rule-based/symbolic systems, statistical approaches, machine learning, and deep learning.
- Learn to recognize and choose appropriate techniques for a range of problems, including search, classification, and generation.
- Implement AI techniques in a variety of domains, including game-playing, robotics, computer vision, and natural language processing.
- Evaluate AI models with respect to scientific validity and social impact.
- Critically consider the ethical consequences of current and future AI technology.

Course Work

Expectations

- You are expected to come to class and to participate. 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.
- When you have questions about **any** course material, you can submit questions using the <u>Anonymous Question Form</u>. I strive to post answers promptly in the <u>Q&A Document</u>.
- You are expected to seek help when you have trouble with the course material. I am happy to help you with **any course material**, no matter how long ago it was taught.

Assignments

There will be 10 homework assignments, typically composed of:

- short answer questions that ask you to reflect on the weekly readings
- short math questions to reinforce the theory behind the techniques we are studying
- programming exercises that you ask to implement AI algorithms or models
- evaluation questions that ask you to analyze the performance of a model that you have implemented

You should expect roughly **75% of each assignment to be Python programming.**

Assignments will be due on Mondays at 10pm.

Midterm

We will have a midterm exam in class on March 1st. 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.

Class Preparedness

You are expected to come to class each week on time, ready to learn and to discuss the weekly reading. You will not be graded for correctness, but you will lose points if you do not participate.

Final Project

In lieu of a final exam, there will be a final project. I have chosen a theme, and you will pick a topic within that theme. You will pursue research on your topic and write a report summarizing your findings. You will also **present your findings** for the class on the last day of class.

Assignment Submission

We will use Gradescope for homework submissions. Some homework problems will involve writing text, rather than code. You may intersperse these answers as comments in your Python program, or submit them as a separate PDF.

Readings

The main textbook for this course is **You Look Like A Thing and I Love You** by Janelle Shane. **Let me know if you have difficulty obtaining a copy**. Technical readings will come from the 3rd edition of *Artificial Intelligence: A Modern Approach* by Peter Norvig and Stuart Russell. **There is a PDF online**. Other readings will be linked to on the course schedule.

Course Policies

Late Policy

- Assignments will typically be due at **10pm Monday each week**.
- 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 fall behind, email me right away. We will figure out a plan that supports your learning and well-being. You do not need to give any personal details to justify this.

Collaboration Policy and the Honor Code

You may not use ChatGPT, Co-pilot, Codex, or any other generative models on your homework assignments. If you disagree with this policy, please talk to me! I am interested in hearing why you think this would benefit your learning.

- 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 0 point in the course (even after the assignment deadline).
- Reference:
 - You may consult external reference resources for general concepts and 0 techniques, but you **must cite them.**
 - You must **not** consult solutions to similar assignments from other AI courses, books, or online resources.
- Code reuse and adaptation:
 - You must not reuse or adapt any code beyond what is given to you in the assignment and in CS232 lectures.
- After the semester:
 - Do not post homework solutions publicly (i.e. on GitHub or your website). You 0 may show your work to potential employers or others who aren't students.

Some of these rules will be relaxed for the final assignment, which is a group project.

Grading

- 70% Homework 15%
- Midterm
- Final project:
 - Probe task 4%
 - Presentation 3%
 - 8% • Report

This course will comply with the <u>Wellesley College grading policy</u>.

Flexibility and Feedback

I welcome feedback throughout the course. You can email me, or, if you prefer to be anonymous, you can use the Anonymous Question Form.

Seeking Help

Lots of help is available in this course!

- If you go 20-30 minutes without making progress, please ask for help!
- You are encouraged to post questions to the <u>Anonymous Question Form</u>. However, you may not post code. For code-specific questions, email me directly.
- Come to tutor hours! You don't need a specific question to attend.
- **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 <u>this calendar</u>.
- You can request a one-on-one tutor from the PLTC.

Disabilities and Accommodations

My job is to help you learn. Part of that is making accommodations for any disabilities you might have. **Feel free to suggest ways that I can make the course more accessible for you**. I will keep 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 <u>Accessibility and Disability Resources (ADR)</u> to get a letter outlining your needs right away, as this can take time to arrange. If you need immediate accommodations, please meet with me. If you are unsure but suspect you may have an undocumented need, ADR can help with referrals.

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.