STAT 498: Introduction to the Ethics of Algorithmic Decision Making

This syllabus is subject to change.

Course Information

Instructor: Peter Gao (petergao at uw dot edu)

Time: Tuesdays and Thursdays, 2:30 - 3:50pm

Location: THO 335

Office hours: T 12:30 - 2:20 via Zoom

Note: When emailing the instructor, please include "[STAT 498]" at the beginning of the email header! It will help us respond to you faster.

Overview

Using examples from medicine, education, and criminal justice, this course surveys ethical & social implications of the design, implementation, & interpretation of statistical decision-making algorithms. Students will examine how algorithms interact with social categories including race, class, & gender, preserving or reshaping existing inequities. Students will evaluate statistical frameworks for balancing fairness & privacy with efficiency.

Course objectives

1. Identify and discuss ethical considerations related to the design, implementation, and results of decision making algorithms.

2. Examine historical uses of data in decision making and draw connections to modern algorithmic decision making systems in fields including medicine, education, and criminal justice.

3. Understand and critique statistical approaches to understanding fairness and privacy, especially as they pertain to data-driven decision making.

4. Practice communicating about ethical considerations in data science through writing, discussion, and presentation.

Outline

In this course, we will cover the following:

- Introduction to ethics and algorithms
- Algorithmic fairness and bias
- Facial recognition technologies
- Algorithms in the public and private spheres
- Privacy and surveillance
- Paths forward
- Final Presentations

Course structure

- **In-class Discussion (10%)**: You will be expected to attend class regularly and to participate in class activities and discussions. Please try to bring something to each class: a question, an idea, a connection, etc. In addition, you will serve as a designated discussant for **two** class sessions during the quarter. As a discussant, you will prepare discussion questions in response to readings and be called upon to

respond to questions that arise during class. More detailed instructions will be provided at the start of the quarter.

- **In-class Work (10%)**: In order to give students the flexibility to miss class as personal/health matters arise, attendance will not contribute to final grades. Instead, students will be asked to complete brief in-class assignments, which students may submit in person or online if they are unable to attend. You will receive zero, half, or full credit based upon completeness and depth of your responses.

- **Long Responses (45%)**: Throughout the quarter, students will complete three writing assignments. These assignments ask students to write brief memos that could be used to communicate about issues related to algorithmic decision systems to a general public. These responses will be graded according to completeness and depth of your work.

  - Response 1 asks students to study an algorithm and identify potential ethical concerns.
  - Response 2 concerns statistical frameworks for assessing algorithmic bias and fairness.
  - Response 3 concerns forms of resistance to algorithmic decision systems.

- **Final Project (35%)**: In groups or by yourself, you will select a topic from class and extend our discussion (ex. By further examining the human impacts of a given algorithm). Halfway through the course, you will submit a final project proposal. During the 10th week, you will submit a written report and self assessment and present a brief talk (< 10 minutes) in-class.

Obtaining a good grade in this class will depend on your ability to stay organized and complete readings and assignments on time, but if you attend class with the goal of helping us to build a healthy discussion, you can and should be successful.

In general, the late policy is as follows: Any assignment that is received late but less than 24 hours late will receive a grade penalty of 25%. Any assignment that is received 24–48 hours late will receive a grade penalty of 50%. Assignments will not be accepted more than 48 hours late. **That said, if you communicate directly with me before an assignment is due, I will often be willing to relax a deadline.**

Expectations

The COVID-19 pandemic has and will continue to present many of us with unforeseen difficulties. I encourage all of you to prioritize the health and safety of yourselves and those around you and would be happy to make accommodations that help you to do so. In addition, I hope that we can be patient with one another as we begin transitioning back to in-person learning, as it is likely that not everything will go to plan. Please feel free to reach out to me via email at any point to discuss any concerns you may have about the course.

Keeping each other safe

I am thrilled to be teaching in person and I hope you are excited to be back on campus as well. As we return to physical classrooms, please be respectful of your classmates' boundaries and precautions–we are all readjusting to in-person learning. In addition, I hope we will all make every effort to keep ourselves and our classmates safe. If you test positive or are exposed to possible infection, I encourage you to err on the side of caution with regards to attending classes and would be happy to make accommodations that allow you to do so. I will make high-quality recordings of lecture and additional office hours available to students that are absent due to quarantine.

Classroom environment

Names & Pronouns

Everyone deserves to be addressed as they would like. Feel free to send us your preferred name and correct pronouns at any time.

Feedback

I encourage and appreciate your feedback throughout the quarter. You are welcome to provide feedback on any aspect of the course at any time via email or in person. If you would prefer to do so confidentially, you can do so through the form [here]

Participation expectations

Above all, we are trying to build a space for spirited discussion and learning. Come to class prepared to share what you have learned and to ask questions of us and your classmates. To be more specific:

- Read and review the assigned texts before class. Every class, try to bring something to contribute. What kind of something? Bring a question to ask the class, an idea to share, or a connection you drew to current events or other courses you have taken. If you loved a reading–great! We want to hear about it. If you hated a reading or think that a writer has it all wrong, also great!
- We hope that you will feel comfortable taking risks in discussion. You do not always have to have the "right" answers or to prepare exactly what you want to say before you raise your hand. We are here to figure it out together, but you must take the first step by beginning to speak.
- Listen to your classmates and really consider what they have to say. We are all here to learn–challenge each other by asking questions or building on each other's ideas. Along these lines: don't hog the mic! Give each other time to think and speak. It is our collective responsibility to help each other feel comfortable sharing our thoughts and ideas in the class.
- You may sometimes feel that you do not have anything to contribute to a discussion. This is okay! We trust your judgment: sometimes the best way to contribute is by listening and taking some time to process your own thoughts.

Academic misconduct

Academic integrity is essential to this course and to your learning. On certain assignments, collaboration is allowed and encouraged when following the collaboration policy outlined above. Violations of the academic integrity policy include but are not limited to: copying from a peer, collaborating where it is not allowed, copying from an online resource, using a solutions manual, and using resources from a previous iteration of the course. Anything found in violation of this policy will be automatically given a score of 0 with no exceptions. If the situation merits, it will also be reported to the UW Student Conduct Office, at which point it will be out of my hands. If you have any questions about this policy, please do not hesitate to reach out and ask.

The university's policy on plagiarism and academic misconduct is a part of the Student Conduct Code, which cites the definition of academic misconduct in the WAC 478-121. (WAC is an abbreviation for the Washington Administrative Code, the set of state regulations for the university. The entire chapter of the WAC on the student conduct code is here http://www.washington.edu/admin/rules/policies/WAC/478-121TOC.html) According to this section of the WAC, academic misconduct includes:

"Cheating"—such as "unauthorized assistance in taking quizzes", "Falsification" "which is the intentional use or submission of falsified data, records, or other information including, but not limited to, records of internship or practicum experiences or attendance at any required event(s), or scholarly research"; and "Plagiarism" which includes "[t]he use, by paraphrase or direct quotation, of the published or unpublished work of another person without full and clear acknowledgment."

The UW Libraries have a useful guide for students at http://www.lib.washington.edu/teaching/plagiarism Students found to have engaged in academic misconduct may receive a zero on the assignment (or other possible outcome).

Conduct

The University of Washington Student Conduct Code (WAC 478-121) defines prohibited academic and behavioral conduct and describes how the University holds students accountable as they pursue their academic

goals. Allegations of misconduct by students may be referred to the appropriate campus office for investigation and resolution. More information can be found online at https://www.washington.edu/studentconduct/.

Disability Resources

Your experience in this class is important to me. It is the policy and practice of the University of Washington to create inclusive and accessible learning environments consistent with federal and state law. If you have already established accommodations with Disability Resources for Students (DRS), please activate your accommodations via myDRS so we can discuss how they will be implemented in this course.

If you have not yet established services through DRS, but have a temporary health condition or permanent disability that requires accommodations (conditions include but not limited to; mental health, attention-related, learning, vision, hearing, physical or health impacts), contact DRS directly to set up an Access Plan. DRS facilitates the interactive process that establishes reasonable accommodations. Contact DRS at http://depts.washington.edu/uwdrs/

Diversity, equity and inclusion

Diverse backgrounds, embodiments, and experiences are essential to the critical thinking endeavor at the heart of university education. Therefore, I expect you to follow the UW Student Conduct Code in your interactions with your colleagues and me in this course by respecting the many social and cultural differences among us, which may include, but are not limited to: age, cultural background, disability, ethnicity, family status, gender identity and presentation, citizenship and immigration status, national origin, race, religious and political beliefs, sex, sexual orientation, socioeconomic status, and veteran status.

Religious accommodations

Washington state law requires that UW develop a policy for accommodation of student absences or significant hardship due to reasons of faith or conscience, or for organized religious activities. The UW's policy, including more information about how to request an accommodation, is available at https://registrar.washington.edu/staffandfaculty/religious-accommodations-policy/.

Accommodations must be requested within the first two weeks of this course using the https://registrar.washington.edu/students/religious-accommodations-request/

Student privacy

Note that the software used in this class (e.g. Canvas, Zoom, Panopto) when used with our UW Net IDs, are FERPA compliant (https://registrar.washington.edu/students/ferpa/). This means they do not monitor student use of their service and they do not share student data with third parties.

Sharing recordings and other class materials outside of class that include personally identifiable student information without the written consent of those students is a violation of FERPA. State law requires consent from people to be recorded (https://apps.leg.wa.gov/rcw/default.aspx?cite=9.73.030), please note that (1) that your participation in this class indicates your consent for course activities to be recorded, (2) you are not permitted to make your own recordings without consent from the instructor and everyone else involved, and (3) that the instructor's recordings will be available for later playback only to students taking the course. For more information about privacy concerns, review the UW Privacy Office policies (https://privacy.uw.edu/policies/best-practices-online-conferencing/), or contact Helen Garrett, the UW's FERPA Officer.

Links and other resources

- Canvas Page