

Final project

*CSCI 3358 – Foundations of algorithmic (un)fairness**J. Finocchiaro*

Summary There are so many important and fascinating challenges that we simply won't have time to cover in class. This project is your opportunity to get some hands-on experience with something that piques your interest. I hope this project can be something you add to your portfolio as you are all applying for jobs or graduate positions in the near future.

Learning goals This assignment is designed to develop skills in...

- *Thoughtful* algorithmic modeling and data engineering that acknowledges shortcomings and emphasizes what matters for your question of interest.
- Translating from social to technical challenges and back
- Collaborating in a team. You'll be doing it in your jobs (more likely than not.) The practice now will hopefully expose you to other viewpoints and experiences that you are bringing in to your presentation.
- Developing technical communication skills, both to technical and non-technical audiences.

Evaluation This case study will be 40% of your final grade, broken down into a project proposal, updates, and final presentation. The case study should have three parts, but how you go about each part has a lot of flexibility.

1. **Project proposal, due Sept 13** (5 pts) Your class project will be in groups of 2-3 on a topic of your choosing.

The [AI Incident Database](#) might be a great place to find inspiration for your project if you need ideas.

Possible topics include fairness or discrimination in...

- Generative AI (e.g., ChatGPT)
- Computer Vision
- Precision Medicine
- Geographic fairness (e.g., facility placement)
- Selection mechanisms (e.g., how does Team USA choose 4 athletes for its overall gymnastics teams?)
- Recommender systems
- Specific algorithmic audits

For this, please submit a 2-3 paragraph proposal with your team members' names, idea for the project, plan of approach, estimated timeline, and some related links and motivation that will serve as your starting point.

2. **Project updates** (10 pts) Throughout the semester we will have 3 project updates to make sure I'm able to provide some insight and feedback into the direction you're going and suggestions.

For updates you have 2 options:

- (a) submit a writeup of (a) what you've done since the last update, (b) roadblocks you're facing/faced and how you've overcome them, (c) your plan for the next sprint, and (d) any specific feedback you'd like from me.

(b) schedule a 15-minute meeting with me, and be prepared to discuss (a)—(d) above.

- **Update 1, due Sept 27, 5pm** At this point, you should understand what the challenges are that you're going to face and how you plan to overcome them. If you want to collect or use a specific dataset, you should identify what this is and have a plan to acquire it.
- **Update 2, due October 25, 5pm** At this point, you should have access to data that you will need for your project, if necessary. If applicable, experiments should be planned. If your project is not data-centric, a good goal at this point is to have your model formalized and an idea of your conjectures that you want to prove.
- **Update 3, due Nov 8, 5pm** You should have some preliminary results, at the very least. If your project is more modeling-oriented, then proof ideas and counterexamples of why the statement is needed are great here. In this update please confirm your plans for your final presentation with me just to make sure I'm on board.

3. **Final presentation** (10 pts) During the final exam period, project presentations will be given in lieu of a written exam. Final project presentations will be about 10-12 minutes, with a few minutes after for questions and discussion. The final presentation can take the form of a slide deck, short film, podcast, or something else! (Just run it by me first!) **This presentation should not be targeted to computer scientists. Instead, pick and share a non-technical audience who you want to reach with your presentation.**

Presentations should give enough background context for your classmates to understand why the problem is important, who it affects, and the “origin story” of how the problem emerged, as well as a (palatable) summary of your findings and their implications.

I am happy to give feedback on slides and presentations with at least a week's notice.

4. **Technical writeup** (15 pts) The technical writeup is due at the beginning of the exam period and uploaded through the Canvas, and should communicate results on a more technical level. The writeup should be 4-8 pages in NeurIPS style ([.docx version](#) or [.sty file for L^AT_EX](#)) and be submitted via Canvas.

In general, the technical writeup should outline

- Problem motivation and background
- Review previous work around the problem
- Describe the modelling approach, including a discussion of its shortcomings
- Description of technical results
- Discussion and conclusion of takeaways and recommendations

but it is not restricted to these points.

Best practices for communicating technical material

- Identify your target audience and prototype someone in it; have they had their coffee? Have they taken a machine learning class? Is their livelihood affected by this area?
- Follow [Dan Larremore's principle](#) of (a) Tell them what you're gonna tell them, (b) tell them, and (c) tell them what you told 'em. Inevitably, someone will zone out, and this repetition will help drive home the takeaway message.
- Please do not throw graphs at me without discussion, and make sure your axes are labeled. For every visualization you add, you should mentally have a one sentence description of what you want the audience to learn from it. Make sure the visualization emphasizes that point. Please see [Sam Way and Dan Larremore's recommendations](#) of data visualization.
- Outline your presentation before you start it. This seriously saves so much time in the long run, and makes sure everything you share has a point.