ARTIFICIAL INTELLIGENCE FOR PUBLIC POLICY: PPHA 38829
Fall quarter 2023

Instructor:	Jens Ludwig
Contact:	jludwig@uchicago.edu
Course meeting details:	Tuesdays, 3:30-6:20 Room 1022 Harris School of Public Policy Keller Center, 1307 East 60th Street

Course Description

It is hard to name a sector that will *not* be dramatically affected by artificial intelligence (or machine learning), from the private sector to government and nonprofits. There are many excellent courses that teach you the *mechanics* behind these innovations -- helping you develop an engineering skill set, like the R or Python programing skills required to build these algorithms.

This course takes a different approach. It is aimed at people who want to *deploy* these tools, whether that's in a start-up company, a medium-sized NGO or a large government agency. While this requires some knowledge of how these tools work, that is only a small part of the equation, just as knowing how an engine works is a small part of understanding how to drive. What is really needed is an understanding of what these tools do well, and what they do badly. This course focuses on giving you a *functional*, rather than mechanistic, understanding. By the end, you should be an expert at identifying ideal (and problematic) use-cases and thereby should be well-placed to create new policy or other applications that use artificial intelligence.

Objectives and Goals

This course aims to equip students with a functional, rather than a mechanical, understanding of AI. Through a series of interactive lessons students will develop a better intuition for AI applications, helping students become expert at identifying ideal use-cases and thereby well-placed to create new products, businesses and policies that use artificial intelligence.

Our goal is to make students smarter consumers of AI. The class is intended as a complement to, not substitute for, standard machine learning classes that focus on the nuts and bolts of how to be an algorithm *producer* ('what's the advantage of a support vector machine over a gradient-boosted decision tree?' or 'how do I program this up in Python?') But there is a distinct set of intuitions that are under-developed in standard machine learning classes, and perhaps under-appreciated even by people who engage in industrial-strength machine learning as their profession. We hope that by the end of the quarter each of you will be positioned to tell even a sophisticated private-sector user of AI like Netflix, with their giant team of data scientists, something that they currently only dimly understand -- or if they understand it, they and their C-suite executives do not fully appreciate how critical it is to the company's entire future as a business (or similarly for, say, the office of a big-city chief information officer).

More generally by the end of the course, students will have:

- A *functiona*l framework for thinking about what AI does.
- The ability to ask questions to determine whether a potential idea is actually a feasible AI project, especially understanding the hidden risks.
- How to evaluate whether an AI system is doing the job you imagine it is.
- How a strategic thinker (rather than an engineer) can help *build* AI systems
- The ability to explain at a high level how AI algorithms work, to facilitate communication with both coders and those who don't know anything about AI at all
- An improved ability to find new opportunities to apply AI that are actually feasible

Students should also know that this course will NOT teach:

- *How to code AI algorithms*
- *How to estimate data models*
- *The mathematical formalisms behind AI or machine learning algorithms*
- *A nitty-gritty understanding of the kinds of specific algorithmic classes out there (e.g. model architectures in convolutional neural networks, etc.).*

The class, therefore, does not require a background in programming.

The class does, however, require enough understanding of key statistical concepts such as mean, variance, sampling, correlation and regression. The course is set up to avoid heavy mathematical notation and formalism; thus imposing a heavy demand for critical thinking.

Prerequisites and Auditing

There are no formal prerequisites for this class although we assume students have familiarity with basic concepts in statistics and regression analysis.

Format

We will rely on:

1. Offline lectures. Videos for each lecture will be posted a week in advance. Watch these early. Please note any questions you have and include them in your homework submissions.
2. In-person class meetings
   a. The first part will be to answer your submitted questions. The goal is to clarify what was unclear in the lecture. Or expand on any points you wanted clarified. **This part will only be as good as the question you submit.**
   b. The second part will be a series of group discussions (a mix of whole-class discussions of cases, as well as smaller breakout groups). The goal will be to reinforce what you have learned in the video lectures.

3. Readings. These will be assigned each week.

Assignments and Grading

There will be no exams. Students will complete two big projects:

- **Explain** *(15 points).* A great way to learn something is to teach it (make whatever inference you wish about why this class is being taught). The goal of this project is to pick one of the concepts from class and to teach it. Specifically, you will put together either a slide deck or essay:
  - A slide deck (with voice over written-out). This will be as if you were giving a 10-15-minute presentation to teach this concept to fellow students.
  - An essay. This will be a 2000-word essay again as if you were putting together a good Medium-post. The audience is again people with your level of knowledge.
  - This will be due the day of the final class
  - You will be graded on: (i) accuracy, (ii) clarity and (iii) the effort you put into making the concept your own. At a minimum, you should not use any examples or explanations we used.
  - This is a solo assignment. We encourage you to practice your talk in groups or circulate your essays for feedback. But every part of what you submit is meant to be entirely yours.

- **Apply** (*25 points*). Another great way to learn something is to apply it. Both as consumers and from your jobs you have a wealth of experience. Use it and find an AI application. The application must be feasible. We will grade first and foremost on feasibility and the thought you have put into how you would build it. Within that constraint, you are obviously looking for lucrative and impactful applications.
  - The final output will be a memo describing the application. It should spell out the basic idea, the potential flaws, how will you address them and your strategy for building it out. You can spend some time on the value but that is not the central point – this is not a strategy, pitching or market-sizing exercise. It is an AI-build exercise.
  - The final output is also a solo assignment just as with the "Explain" assignment. But we realize that brain-storming is hard so we'll encourage you to work in groups of 4-6 (if you have trouble finding a group contact a TA). The groups are solely for finding and generating ideas. They are not meant for the homework assignments and vetting of ideas.

- **Submitted questions**
  - These will be submitted by Sunday at 5PM along with your homework below. Submitting thoughtful questions is part of your class participation grade.

- **Homework assignments** (60 points).
  - Each one is due by **Sunday at 5PM**

    - Each homework will be addressing a question for the lecture you just saw and we discussed.
    - If you are confused about something submit a question – otherwise you will do badly on the homework associated with that lecture.
    - There will be 6 homeworks, each graded on a 10-point scale
- **Grading of homeworks:**
    - If you misunderstood something, how we penalize you depends on whether you asked a question about it.
    - If you did ask a question about the thing you got wrong, then you will be penalized less: obviously I failed to explain it well
    - If you didn't ask a question, you will be penalized heavily.

Late assignments are not accepted and will simply receive a 0.

Course Outline

Below you will find a short overview of the topics and order in which we will cover them. Guesses for how many topics each lecture will take are in brackets.

September 26 Introduction

Watch video lectures in advance of class

Required reading

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig and Sendhil Mullainathan (2018) "Human decisions and machine predictions." *Quarterly Journal of Economics.* 237-293.

Angwin, Julia, Jeff Larson, Surya Mattu and Lauren Kirchner (2016) "Machine bias." *ProPublica*. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

October 3 Factory Tour & Blueprints

Watch video lectures in advance of class

Homework 1 due

Submit student questions by 5pm October 1

Required readings

Will Parker and Konrad Putzkier, "What went wrong with Zillow? A real-estate algorithm derailed its big get. The company had staked its future growth on its digital home-flipping business, but getting the algorithm right proved difficult." *Wall Street Journal*.

Marco De Nadai and Bruno Lepri (2018) "The economic value of neighborhoods: Predicting real estate prices from the urban environment."

<u>October 10 Datafication</u>

Watch video lectures in advance of class

Homework 2 due

Submit student questions by 5pm October 8

Required readings

Sara B. Heller, Benjamin Jakubowski, Zubin Jelveh and Max Kapustin (2022) "Machine learning can predict shooting victimization well enough to help prevent it." Cambridge, MA: NBER working paper 30170.

<u>October 17 Finding Opportunities, & When ML is the wrong choice</u>

Watch video lectures in advance of class

Submit homework 3

Submit student questions by 5pm, October 15

Required readings

Erik Brynjolfsson, Tom Mitchell and Daniel Rock (2018) "What can machines learn and what does it mean for occupations and the economy?" *American Economic Association Papers & Proceedings*. 108: 43-47.

Erik Brynjolfsson and Tom Mitchell (2017) "What can machine learning do? Workforce implications." *Science*. 358(6370): 1530-4.

<u>October 24 Bad data and how to fix it</u>

Watch video lectures in advance

Submit homework 4

Submit student questions by 5pm October 22

Required Readings

Sendhil Mullainathan and Ziad Obermeyer (2017) "Does machine learning automate moral hazard and error?" *American Economic Review: Papers & Proceedings*. 107(5): 476-480.

<u>October 31 Bad data and how to fix it (cont'd) and evaluating the algorithm</u>

Watch video lectures in advance

Submit homework 5

Submit student questions by 5pm October 29

Required readings

Ben Dickson (July 29, 2020) "Why deep learning won't give us level 5 self-driving cars." *TechTalks*. https://bdtechtalks.com/2020/07/29/self-driving-tesla-car-deep-learning/

<u>November 7 Common issues with ML algorithms</u>

Watch video lectures in advance

Submit homework 6

Submit student questions by 5pm November 5

Required Readings

Sam Corbett-Davies, Emma Pierson, Avi Feller and Sharad Goel (2016) "A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear." *Washington Post*. https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/

Sendhil Mullainathan (12/6/2019) "Biased algorithms are easier to fix than biased people." *The New York Times*. https://www.nytimes.com/2019/12/06/business/algorithm-bias-fix.html

Ziad Obermeyer, Brian Powers, Christine Vogeli and Sendhil Mullainathan (2019) "Dissecting racial bias in an algorithm used to manage the health of populations." *Science*. 366(6464): 447-53.

Jeffrey Dastin (2018) "Amazon scraps secret AI recruiting tool that showed bias against women." *Reuters*. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

November 14 Behavioral science and algorithmic blueprints

(No video lecture this week)

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan and Manish Raghavan (2023) “The inversion problem: Why algorithms must infer mental state and not just predict behavior.” *Perspectives on Psychological Science*.

Kahneman, Daniel (2011) *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

November 21 (No class, Thanksgiving break)

November 28 Some of my favorite applications

Watch video lectures in advance

Peter Bergman, Danielle Li and Lindsey Raymond (2023) “Hiring as exploration.”

Peter Bergman, Elizabeth Kopko and Julio Rodriguez (2023) “A seven-college experiment using algorithms to track students: Impacts and implications for equity and fairness.”

Marie-Pascale Grimon and Chris Mills (2022) “The impact of algorithmic tools on child protection: Evidence from a randomized controlled trial.”