Algorithms, Public Policy, and Ethics
POLI 170 – Fall 2020
Tuesday, Thursday 11:00am – 12:20pm
Location: `https://ucsd.zoom.us/j/98158397497`

Professor: Kirk Bansak
Contact: `kbansak@ucsd.edu`
Office Hours: Wednesday, 10:00am – 12:00pm (sign-up details below)
Location: `https://ucsd.zoom.us/j/8599340921` (different from lecture)

TA: Dora Hu
Contact: `xyhu@ucsd.edu`
Office Hours: Monday, 10:00am – 12:00pm
Location: `https://ucsd.zoom.us/j/7178050921`

1 Overview

With machine learning becoming more pervasive and data availability improving over time, decision-makers in various realms of public policy are increasingly utilizing predictive algorithms to help inform or optimize their decisions. For instance, in the area of criminal justice, algorithms are often used to make predictions like a criminal defendant's likelihood of failing to appear at court or reoffending in the future. These predictions are then used as risk assessments to inform various decisions, such as whether to grant a defendant pretrial release. The goal of incorporating these types of tools and data into public policy has been to address shortcomings in existing decision-making processes. However, the ability of algorithms to produce fair decisions and improve policy has been the subject of intense debate in recent years. Policymakers, academic researchers, and the popular media alike have scrutinized the increasing deployment of such tools.

As members of society—and potentially future policymakers, data scientists, lawyers, etc.—we have a stake in how public policy is being shaped and executed. What are the benefits, limitations, dangers, and ethics of the use of algorithms in public policy? How can we formulate balanced perspectives on these issues so that we can carefully evaluate different use cases? How can the deployment of such tools be improved in the future?

1.1 Objectives

This course will cover a broad range of topics on the use of predictive and related algorithms in public policy. This will include specific case studies, background material on how data are used in these tools, their possible benefits relative to status quo procedures as well as their limitations, and the potential harms and ethics surrounding their use (e.g. issues of algorithmic bias). As the goals of this course, students will:

1) Be introduced to a broad range of topics related to the use of algorithms in public policy.

2) Learn about specific case studies in which algorithms have been deployed in real-world public policy processes.

3) Develop key conceptual perspectives to think critically about, and weigh the pros and cons of, algorithms in specific public policy use cases. For example, perspectives on:

   - Prediction and predictive efficacy
   - Algorithmic bias
   - Applied notions of fairness
   - Human-algorithm interaction

4) Be introduced to various concepts and understandings that are useful for thinking about data, prediction, analytics, and ethics in society more generally.

5) Develop programming abilities in `R`.

1.2 Prerequisites

There are two prerequisites for the course. The first is upper-division standing. The second prerequisite is POLI5/POLI5D/ECON5 (Data Analytics for the Social Sciences), in particular as that course introduced the statistical software `R`. We will make use of `R` during class and in homework. See section regarding the use of `R` on the syllabus below.

In order to motivate and illustrate key concepts and themes throughout the course, there will also be some mathematical content, and many of the readings contain quantitative analyses and results. However, the course does not require a mathematical background other than basic arithmetic and algebra, as well as an elementary understanding of probability.

1.3 Evaluation

Students will be evaluated across the following areas.

- **Data Analysis Assignments: 50% of your grade.**

  Data Analysis Assignment 1: 10%

  Data Analysis Assignment 2: 20%

  Data Analysis Assignment 3: 20%

There will be three data analysis assignments that you will complete using `R` and `R Markdown` (see section regarding the use of `R` on the syllabus below). The first assignment will serve mainly as a refresher on using `R` and an introduction to `R Markdown` for anyone who has not used it before. The second and third assignments will be longer and more substantial.

You are allowed to work on data analysis assignments together in small groups of no more than three if desired, **but you must write up your code and answers on your own!** Submitting any code or answers that are copied from another student is unacceptable and in violation of academic integrity. If you work together, you must indicate on your assignments who your co-workers were.

- **Policy Memo: 30% of your grade.**

  Group Component: 20%

  Individual Component: 10%

  In groups of 3-4, you will write a policy memo in which you propose, discuss, and evaluate the hypothetical deployment of an algorithmic tool to improve some process or function at UCSD. It will be up to your group to identify an area of campus life, university decision-making, or any other official UCSD business that could theoretically be improved via the assistance of a data-driven algorithm. Using the various themes and concepts learned throughout the course, you will explain how such a tool could be integrated into (or replace) an existing process, and you will evaluate the possible benefits, risks, and viability of its deployment. Based on this evaluation, you will make a final recommendation as to whether deployment would ultimately be desirable. In addition to the group memo that your group will write together, there will also be an individual written component that each student will complete separately and independently. More information on the precise format and guidelines for this assignment will be provided later.

  - **Group Formation**: Each group must e-mail the instructors (`kbansak@ucsd.edu` and `xyhu@ucsd.edu`) a list of group members or request to be assigned to a group by **Tuesday, November 3rd**.
  - **Proposal Approval**: Prior to writing the policy memo, each group must get the instructors' approval of the proposed UCSD algorithm use case that will be highlighted in the memo. Groups are encouraged to get approval as early as possible, and the deadline for doing so is **Thursday, November 19th**.
  - **Assignment Due Date**: The group memo and individual component are both due on **Friday, December 4th**.

- **Final Exam: 20% of your grade.**
  You will complete a cumulative final exam, administered online and on the date delineated under UCSD's Fall Quarter exam schedule.

2 Logistics

2.1 Class Meetings

All class meetings will be conducted online at our scheduled class time (Tuesday, Thursday 11:00am – 12:20pm Pacific Time) via the video-conferencing platform Zoom. It is recommended that everyone tries to connect to Zoom in advance of our first class meeting to ensure the software is working properly. Please note that only authenticated users will be able to join our Zoom meeting room, which means you will need to ensure you have signed into your UC San Diego Zoom account. Our Zoom link for class meetings is the following:

`https://ucsd.zoom.us/j/98158397497`

In addition, all class meetings will be recorded and posted on our `canvas.ucsd.edu` page. However, students are highly encouraged to attend class at the scheduled meeting time in order to have the opportunity to participate and ask questions in real time. Note that in-class participation is not required and will not factor into grading. Individuals who do not want to have their surroundings visible are encouraged to use Zoom's virtual background feature, if feasible, or to participate without video. Please also be mindful of others who may not wish to be visible or recorded in the background.

2.2 Office Hours

I will hold office hours from 10:00am to 12:00pm on Wednesdays on Zoom. Please note that the Zoom room for office hours is different from the room used for lecture! Office hours will be held here:

`https://ucsd.zoom.us/j/8599340921`

Please make sure to sign up for office hours in advance using the following Calendly link:

`https://calendly.com/kbansak/officehours`

If you would like to meet but have class or another important commitment during my office hours, I am happy to make accommodations. Please e-mail me to arrange an alternative time.

2.3 Teaching Assistant

Dora Hu (`xyhu@ucsd.edu`) will be the teaching assistant for this course. Dora will be holding office hours on Mondays from 10:00am to 12:00pm on Zoom here:

`https://ucsd.zoom.us/j/7178050921`

2.4 Course Website

We will use our `canvas.ucsd.edu` page to distribute readings, submit assignments, and access Zoom recordings.

2.5 Readings

There is no textbook for this course. All readings will be uploaded to the Canvas page or will be available online via the links listed on the syllabus. Pay attention to page numbers when noted on the syllabus, which indicate the portions of a reading that are required. If there are no page numbers listed, then the document should be read in its entirety. However, anything *highlighted in yellow* in the readings uploaded onto the Canvas page is *not required*!

2.6 R

Any homework requiring the use of `R` should be completed and submitted using `R Markdown`, a markup language for producing well-formatted HTML documents with embedded `R` code and outputs. `R Markdown` requires installation of the `rmarkdown` and `knitr` packages. We strongly recommend using `RStudio`, an IDE for `R`, which is set up well for the creation of `R Markdown` documents.

To download/install `R`:
`https://cran.r-project.org/`

To download/install and learn more about `RStudio`:
`http://www.rstudio.com`

For more on `R Markdown`:
`http://rmarkdown.rstudio.com`

2.7 Assignment Schedule Summary

Due Date	**Assignment**
Before Each Class Meeting	Assigned Readings
Thursday, October 15	Data Analysis Assignment 1
Thursday, October 29	Data Analysis Assignment 2
Tuesday, November 3	Group Formation for Policy Memo
Tuesday, November 17	Data Analysis Assignment 3
Thursday, November 19	Approval of Policy Memo Proposal
Friday, December 4	Policy Memo (Group & Individual Components)
Wednesday, December 16	Final Exam

2.8 Expectations

	What you can expect of me	What I expect of you
Lecture	I will hold lectures online during our normally scheduled class time, and I will post the lecture recordings on Canvas.	I encourage everyone to join the lectures synchronously if possible to have the opportunity to participate and ask questions in real time. However, this is not required.
Emails	I try to respond to emails as soon as I can, and at least within 24 hours. Don't hesitate to send me a reminder if I haven't responded within that window.	Before writing to me with a question, check the syllabus to see if it's addressed there first. Please don't leave questions until the last minute.
Availability	I'll always be available during the office hours stated above.	Please use my office hours if you would like to discuss something. If you have class during my office hours, please e-mail me to arrange an alternative time.
Reading	The content in some of the readings I have assigned can be challenging or dense, so it's OK if you don't immediately understand everything. We will use class and office hours to clarify.	Please do all the readings. With academic articles, read outside-in: the introduction and conclusion first, then the middle.
Late Submission	I am willing to make reasonable accommodations and understand that difficult situations can arise. However, I will not make exceptions for one person that are not available to everyone else.	Complete all assignments on time, and reach out to me if you are encountering troubles.

3 Course Outline

3.1 Introduction

- Thursday, October 1, 2020

  Topic: **Course Introduction**

  Nothing to read before class.

- Tuesday, October 6, 2020

  Topic: **Algorithms and Public Policy: The Big Picture**

  Read before class:

  - Jacob Brogran, "What's the Deal With Algorithms?" *Slate Magazine*, February 2, 2016. Available at:
    https://slate.com/technology/2016/02/whats-the-deal-with-algorithms.html
  - Ben Buchanan and Taylor Miller, "Machine Learning for Policymakers: What It Is and Why It Matters," Belfer Center for Science and International Affairs, 2017.

- Thursday, October 8, 2020

  Topic: **A Refresher on R**

  Read/Do before class:

  - Make sure `R` and `RStudio` are installed on your computer (see section in the syllabus on `R` above).
  - Read section 2 ("2. Defining tidy data") in the following article:
    Hadley Wickham, "Tidy data," *Journal of Statistical Software* Vol. 59, No. 10 (2014).
  - Review your `R` materials from POLI5/POLI5D/ECON5.
  - Optional, for your reference: Additional refresher tutorial videos can be found at the links below.
    `https://www.youtube.com/playlist?list=PLOU2XLYxmsIK9qQfztXeybpHvru-TrqAP`
    `https://www.youtube.com/playlist?list=PLjgj6kdf_snYBkIsWQYcYtUZiDpam7ygg`

Case Study 1

- Tuesday, October 13, 2020

  Topic: **Allocation of Health Care Services**

  Read before class:

- David W. Bates, Suchi Saria, Lucila Ohno-Machado, Anand Shah, and Gabriel Escobar, "Big Data in Health Care: Using Analytics to Identify and Manage High-Risk and High-Cost Patients," *Health Affairs* Vol. 33, No. 7 (2014).
- Natalia Rodriguez, "Infographic: How to Read a Scientific Paper," *Elsevier Connect*, August 5, 2015. Available at: https://www.elsevier.com/connect/infographic-how-to-read-a-scientific-paper
- Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan, "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations," *Science* Vol. 366, No. 6464 (2019).

3.2 Better Understanding Data and Prediction

- Thursday, October 15, 2020

  Data Analysis Assignment 1 Due

  Topic: **Data, Data Sets, and Measurement**

  Read before class:

  - John D. Kelleher and Brendan Tierney, "What are Data, and What is a Data Set?" in *Data Science*, MIT Press, 2018.
  - Daniel T. Kaplan, "Data: Cases, Variables, Samples" in *Statistical Modeling: A Fresh Approach*, Project MOSAIC Books, 2009.

- Tuesday, October 20, 2020

  Topic: **Demystifying Prediction**

  Read before class:

  - Ajay Agrawal, Joshua Gans, and Avi Goldfarb, "Prediction Machine Magic," in *Prediction Machines*, Harvard Business Review Press, 2018.
  - Gareth James, Daniela Witten et al., *An Introduction to Statistical Learning*, Springer, 2013. Read pp. 15-26, 61-63, 71-74 (Recommended but optional: pp. 63-71, 75-82).

- Thursday, October 22, 2020

  Topic: **Implementing Prediction**

  Read before class:

  - Gareth James, Daniela Witten et al., *An Introduction to Statistical Learning*, Springer, 2013. Read pp. 82-92, 109-119.
  - Dan Kopf, "This is How Computers 'Predict the Future,'" *Quartz*, September 5, 2018. Available at: https://qz.com/1261817/predictive-algorithms-are-not-all-that-complicated/

- Tuesday, October 27, 2020

  Topic: **Evaluating Predictive Performance and Uncertainty**

  Read before class:

  - Gareth James, Daniela Witten et al., *An Introduction to Statistical Learning*, Springer, 2013. Read pp. 29-33, 176-183.
  - Georgios Drakos, "How to select the Right Evaluation Metric for Machine Learning Models: Part 1 Regression Metrics," Towards Data Science, *Medium*, August 26, 2018. Available at: https://towardsdatascience.com/how-to-select-the-right-evaluation-metric-for-machine-learning-models-part-1-regrression-metrics-3606e25beae0
  - Georgios Drakos, "How to select the Right Evaluation Metric for Machine Learning Models: Part 3 Classification Metrics," Towards Data Science, *Medium*, September 12, 2018. Available at: https://towardsdatascience.com/how-to-select-the-right-evaluation-metric-for-machine-learning-models-part-3-classification-3eac420ec991

Case Study 2

- Thursday, October 29, 2020

  Data Analysis Assignment 2 Due

  Topic: **Algorithms in Criminal Justice**

  Read before class:

  - The Sentencing Project, "Report of the Sentencing Project to the United Nations Special Rapporteur on Contemporary Forms of Racism, Racial Discrimination, Xenophobia, and Related Intolerance: Regarding Racial Disparities in the United States Criminal Justice System," 2018.
  - Nathan James, "Risk and Needs Assessment in the Federal Prison System," *Congressional Research Service*, July 10, 2018.
  - Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner, "Machine Bias," *ProPublica*, May 23, 2016. Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
  - Sam Corbett-Davies, Sharad Goel, and Sandra González-Bailón, "Even Imperfect Algorithms Can Improve the Criminal Justice System," *New York Times*, December 20, 2017. Available at: https://www.nytimes.com/2017/12/20/upshot/algorithms-bail-criminal-justice-system.html

Election Day

- Tuesday, November 3, 2020

  Group Formation for Policy Memo Due

  No Class! Go Vote!

3.3 Algorithmic Bias and Ethics

- Thursday, November 5, 2020

  Topic: **Ethics, Algorithms, and the Law**

  Read before class:

  – Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, Luciano Floridi, "The Ethics of Algorithms: Mapping the Debate," *Big Data & Society*, Vol. 3, No. 2 (2016). Read only pp. 1–12.

  – Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan, and Cass R. Sunstein, "Discrimination in the Age of Algorithms," *Journal of Legal Analysis* Vol. 10 (2018). Read only pp. 113-132.

- Tuesday, November 10, 2020

  Topic: **Applied Definitions of Fairness**

  Read before class:

  – Sahil Verma and Julia Rubin, "Fairness Definitions Explained," *2018 IEEE/ACM International Workshop on Software Fairness (FairWare)* (2018).

- Thursday, November 12, 2020

  Topic: **Choosing Fairness Priorities and Impossibility Results**

  Read before class:

  – Sam Corbett-Davies, Emma Pierson, Avi Feller, and Sharad Goel, "A Computer Program Used for Bail and Sentencing Decisions was Labeled Biased against Blacks. It's Actually Not that Clear," *Washington Post*, October 17, 2016. Available at:
  https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/

  – Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq, "Algorithmic Decision Making and the Cost of Fairness," *Proceedings of the 23rd Conference on Knowledge Discovery and Data Mining* (2017).

- Tuesday, November 17, 2020

  Data Analysis Assignment 3 Due

  Topic: **Sources of Bias**

  Read before class:

  - Solon Barocas and Andrew D. Selbst, "Big Data's Disparate Impact," *California Law Review* Vol. 104 (2016). Read only pp. 671–693.
  - Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan, and Cass R. Sunstein, "Discrimination in the Age of Algorithms," *Journal of Legal Analysis* Vol. 10 (2018). Read only pp. 132–153.

- Thursday, November 19, 2020

  Approval of Policy Memo Proposal Due

  Topic: **Dealing with Bias: Human and Algorithmic**

  Read before class:

  - Nicol Turner Lee, Paul Resnick, and Genie Barton, "Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms," *Brookings Institution*, Wednesday, May 22, 2019. Available at: https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/
  - Sendhil Mullainathan, "Biased Algorithms Are Easier to Fix Than Biased People," *New York Times*, December 6, 2019. Available at: https://www.nytimes.com/2019/12/06/business/algorithm-bias-fix.html
  - Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan, and Cass R. Sunstein, "Discrimination in the Age of Algorithms," *Journal of Legal Analysis* Vol. 10 (2018). Read only pp. 154–164.

Case Study 3

- Tuesday, November 24, 2020

  Topic: **Improving Refugee Integration via Algorithmic Assignment**

  Read before class:

  - James Vincent, "Could an Algorithm Help Find the Right Place to Resettle Refugees?" *The Verge*, January 18, 2018. Available at: https://www.theverge.com/2018/1/18/16905962/algorithm-resettle-refugees-machine-learning-research-employment
  - Kirk Bansak et al., "Improving Refugee Integration through Data-Driven Algorithmic Assignment," *Science* Vol. 359, No. 6373 (2018).

Thanksgiving

- Thursday, November 26, 2020

  No Class! Thanksgiving Holiday!

3.4 From Prediction to Decision-Making

- Tuesday, December 1, 2020

  Topic: **Prediction, Decision-Making, and Human-Algorithm Interaction**

  Read before class:

  – Angèle Christin, "Algorithms in Practice: Comparing Web Journalism and Criminal Justice," *Big Data & Society* Vol. 4, No. 2 (2017).

  – Ryan Kennedy, Philip D. Waggoner, and Matthew M. Ward, "Trust in Public Policy Algorithms," *Journal of Politics*, Forthcoming (2020).

  – Ben Green and Yiling Chen, "Disparate Interactions: An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments," *Proceedings of the Conference on Fairness, Accountability, and Transparency* (2019). Read only pp. 1-10.

- Thursday, December 3, 2020

  Topic: **Limits of Prediction-Based Action**

  Read before class:

  – Michael Luca et al., "Algorithms Need Managers, Too," *Harvard Business Review*, January–February, 2016. Available at: https://hbr.org/2016/01/algorithms-need-managers-too

  – Susan Athey, "Beyond Prediction: Using Big Data for Policy Problems," *Science* Vol. 355, No. 6324 (2017).

- Friday, December 4, 2020

  Policy Memo (Group and Individual Components) Due

- Tuesday, December 8, 2020

  Topic: **Getting to Deployment & Other Issues for Algorithms in Practice**

  Read before class:

  – Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan, "A Guide to Solving Social Problems with Machine Learning," *Harvard Business Review*, December 8, 2016. Available at: https://hbr.org/2016/12/a-guide-to-solving-social-problems-with-machine-learning

- Bruno Lepri, Nuria Oliver, Emmanuel Letouzé, Alex Pentland, and Patrick Vinck, "Fair, Transparent, and Accountable Algorithmic Decision-making Processes," *Philosophy & Technology* Vol. 31, No. 4 (2018).
- Article 29 Data Protection Working Party, "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679," February 6, 2018. Just skim through this!

3.5 Conclusion

- Thursday, December 10, 2020

  Topic: **Conclusions and Review**

  Nothing to read before class.

- Wednesday, December 16, 2020

  Final Exam

4 Policies

4.1 Academic Integrity

Students agree that by taking this course, all required assignments may be subject to submission to `Turnitin.com` for textual similarity review and the detection of plagiarism. All submitted assignments may be included as source documents in the `Turnitin.com` reference database solely for the purpose of detecting plagiarism of such papers. Use of the `Turnitin.com` service is subject to the terms of use agreement posted on the `Turnitin.com` site.

4.2 Students with Disabilities

Students requesting accommodations for this course due to a disability must provide a current Authorization for Accommodation (AFA) letter issued by the Office for Students with Disabilities (`https://osd.ucsd.edu/`). Students are required to discuss accommodation arrangements with instructors and OSD liaisons in the department well in advance of any exams or assignments. The OSD Liaison for the Department of Political Science is Joanna Peralta; please connect with her as soon as possible via the Virtual Advising Center:

`https://stark.ucsd.edu/students/vac/`

4.3 UC San Diego Principles of Community

The University of California, San Diego is dedicated to learning, teaching, and serving society through education, research, and public service. Our international reputation for excellence is due in large part to the cooperative and entrepreneurial nature of the UC San Diego community. UC San Diego faculty, staff, and students are encouraged to be creative and are rewarded for individual as well as collaborative achievements.

To foster the best possible working and learning environment, UC San Diego strives to maintain a climate of fairness, cooperation, and professionalism. These principles of community are vital to the success of the University and the well being of its constituents. UC San Diego faculty, staff, and students are expected to practice these basic principles as individuals and in groups.

For the complete UC San Diego Principles of Community, see:
`https://ucsd.edu/about/principles.html`

5 Resources

Library Help and Research Tools:	https://library.ucsd.edu/ask-us/triton-ed.html
Writing Hub:	https://commons.ucsd.edu/students/writing/index.html
Supplemental Instruction:	https://commons.ucsd.edu/academic-support/supplemental-instruction/si-students.html
Tutoring:	https://commons.ucsd.edu/academic-support/content-tutoring/index.html
Mental Health Services:	https://caps.ucsd.edu
Community Centers:	Learn about the different ways UC San Diego explores, supports, and celebrates the many cultures in our diverse community. https://students.ucsd.edu/student-life/diversity/index.html
Accessibility:	https://disabilities.ucsd.edu/
Basic Needs:	Any student who has difficulty accessing sufficient food to eat every day, or who lacks a safe and stable place to live, and believes this may affect their performance in this course, is encouraged to contact: foodpantry@ucsd.edu and basicneeds@ucsd.edu. For more information, see https://basicneeds.ucsd.edu/