<div align="center">

Algorithms, Public Policy, and Ethics
POLI 179 – Winter 2020
Tuesday, Thursday 6:30pm – 7:50pm
Location: Warren Lecture Hall, Room 2111

</div>

Professor: Kirk Bansak
Contact: `kbansak@ucsd.edu`
Office: SSB 365
Office Hours: Wednesdays, 10:00 AM – 12:00 PM

TA: Kevin Flannagan
Contact: `kflannag@ucsd.edu`
Office: SSB 448
Office Hours: Mondays, 10:00 AM – 12:00 PM

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 predict things 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 pre-trial 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 and/or data scientists—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 and prediction in society more generally.

1.2 Prerequisites

In order to motivate and illustrate key concepts and themes throughout the course, there will be a limited amount of mathematical content, and many of the readings contain quantitative analyses and results. However, the course does not require a technical background other than basic arithmetic and algebra, as well as an elementary understanding of probability. There are no prerequisites other than upper-division standing.

1.3 Evaluation

Students will be evaluated across the following areas.

- **In-Class Participation: 35% of your grade.**

  - **Daily Quizzes**
    At the start of each class, there will be a short online quiz (accessible via any smart phone or computer) featuring a small set of straightforward, factual questions about the readings. These quizzes must be completed without the help of others.

  - **Participation in Discussions**
    At various points during class, particular issues will be raised for discussion in small groups or as a whole. Participation will take a variety of forms, including discussing prompts in groups, answering specific questions posed in class, raising new questions, bringing up issues or concepts that require additional clarification, and submitting responses to in-class surveys.

- **Group Policy Memo: 30% of your grade.**
  In groups of approximately four, 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. The precise format and guidelines for this paper are TBD.

– **Group Formation**: Each group must e-mail the instructors (`kbansak@ucsd.edu` and `kflannag@ucsd.edu`) a list of group members by **Tuesday, February 4th**.

– **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 **Tuesday, February 18th**.

– **Policy Memo Due Date**: The memo is due on **Tuesday, March 3rd**.

- **Group Presentation: 5% of your grade.**
  Each group will give a 10-15 minute in-class presentation on their policy memo, highlighting their proposal, evaluation, and final recommendation. The presentations will take place on March 5th and March 10th, and the order of presentations will be randomly drawn on each day.

- **Final Exam: 30% of your grade.**
  The final exam will take place on Tuesday, March 17th, at 7:00 – 10:00 PM.

2 Logistics

2.1 Office Hours

I will hold office hours from 10:00 AM to 12:00 PM on Wednesdays in my office, room 365 of the Social Sciences Building. Please make sure to sign up for office hours in advance using the Calendly link: calendly.com/kbansak/officehours. If you would like to meet but have class during my office hours, please e-mail me to arrange an alternative time.

2.2 Teaching Assistant

Kevin Flannagan (`kflannag@ucsd.edu`) will be the teaching assistant for this course. He will be holding office hours on Mondays in the Social Sciences Building office 448 from 10:00 AM to 12:00 PM.

2.3 Course Website

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

2.4   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*!

2.5   In-Class Technology

We will be using Mentimeter and Qualtrics surveys in order to conduct in-class surveys, take quizzes, and participate in discussion. These services will be accessible via any smart phone or computer. If you don't have either of these, please fill out your answers on a piece of paper and bring it to me after class.

2.6   Expectations

	What you can expect of me	**What I expect of you**
Attendance	I'll be there, and I will always begin on time!	Regular attendance is essential and expected. Please show up to class a few minutes early.
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 confusing points.	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 everyone else.	Complete all assignments on time and reach out to me if you are encountering troubles.

3 Course Outline

3.1 Introduction

- Tuesday, January 7, 2020

  Topic: **Course Introduction**

  Nothing to read before class.

- Thursday, January 9, 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.

Case Study 1

- Tuesday, January 14, 2020

  Topic: **Allocation of Health Care Services**

  Read before class:

  - David W. Bates et al., "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 et al., "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, January 16, 2020

  Topic: **Basics of 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, January 21, 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 et al., "An Introduction to Statistical Learning," Springer, 2013. Read pp. 15-26, 61-63.

- Thursday, January 23, 2020

  Topic: **Hands-on Example: Implementing a Predictive Algorithm**

  Read before class:

  – 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, January 28, 2020

  Topic: **Evaluating Predictive Performance and Uncertainty**

  Read before class:

  – 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, January 30, 2020

  Topic: **Algorithms in Criminal Justice**

  Read before class:

  – Nathan James, "Risk and Needs Assessment in the Federal Prison System," *Congressional Research Service*, July 10, 2018.

  – Julia Angwin et al., "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

3.3 Algorithmic Bias and Ethics

- Tuesday, February 4, 2020

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

  Read before class:

  – Brent Daniel Mittelstadt et al., "The Ethics of Algorithms: Mapping the Debate," *Big Data & Society*, Vol. 3, No. 2 (2016). Read only pp. 1–12.

  – Jon Kleinberg et al., "Discrimination in the Age of Algorithms," *Journal of Legal Analysis* Vol. 10 (2018). Read only pp. 113-132.

- Thursday, February 6, 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).

- Tuesday, February 11, 2020

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

  Read before class:

– Sam Corbett-Davies et al., "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 et al., "Algorithmic Decision Making and the Cost of Fairness," *Proceedings of the 23rd Conference on Knowledge Discovery and Data Mining* (2007).

- Thursday, February 13, 2020

  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 et al., "Discrimination in the Age of Algorithms," *Journal of Legal Analysis* Vol. 10 (2018). Read only pp. 132–153.

- Tuesday, February 18, 2020

  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 et al., "Discrimination in the Age of Algorithms," *Journal of Legal Analysis* Vol. 10 (2018). Read only pp. 154–164.

Case Study 3

- Thursday, February 20, 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).

3.4 From Prediction to Decision-Making

- Tuesday, February 25, 2020

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

  Read before class:
  - 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.
  - Berkeley J. Dietvorst, Joseph P. Simmons, and Cade Massey, "Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them," *Management Science* Vol. 64, No. 3 (2018).

- Thursday, February 27, 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).

- Tuesday, March 3, 2020

  Group POLICY MEMOS DUE in Class Today!

  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 et al., "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!

Group Presentations

- Thursday, March 5, 2020

  Presentation Day 1

- Tuesday, March 10, 2020

  Presentation Day 2

3.5   Conclusion

- Thursday, March 12, 2020

  Topic: **Conclusions and Review**

  Nothing to read before class.

4 Policies

4.1 Academic Integrity

Students agree that by taking this course all required papers will be subject to submission for textual similarity review to `Turnitin.com` for the detection of plagiarism. All submitted papers will 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 OSD Accommodations

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 via in-person advising (SSB 301) or the Virtual Advising Center.

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`