Ethical Algorithms for the Modern Clinician

Overview

Ethical Algorithms for the Modern Clinician is a short course intended to teach future clinicians the basics of machine learning (ML) and artificial intelligence (AI) as they pertain to clinical practice.

This resource does ***not*** discuss how machine learning algorithms work or how to build ML models. Instead, we introduce ML from clinical practitioner's perspective and discuss what ML does right, where it falls short, and how it will impact patient care. We focus on five key aspects:

1. **Introduction to Machine Learning**: What is machine learning? How is it similar to and different from conventional software?
2. **Bias and Fairness**: How can algorithms be biased against different patient groups? How can we quantify, detect, and reduce bias in clinical decision making algorithms?
3. **Privacy and Anonymization**: How can we anonymize patient data? Why does anonymization often fall short in protecting patient identities? How can we ensure that clinicians maintain patient privacy?
4. **Algorithmic Interpretability**: What does mean for an algorithm to be interpretable? Is it important for us to be able to explain how an algorithm works in order to use it in clinical practice?

5. **Generative AI**: What is generative AI, and how might it be used for patient care? What are the new challenges and opportunities associated with generative AI models?

Just as epidemiology was introduced in medical school curricula alongside the rise of evidence-based guidelines[1], **it is crucial for future clinicians to have a working understanding of machine learning algorithms** as they become increasingly prevalent in clinical practice. Much of the curriculum has been inspired by the (much longer) ethical algorithms course at Penn taught by Michael Kearns. *Ethical Algorithms for the Modern Clinician* is intended as foundational knowledge for medical students **of all backgrounds** - prior experience with machine learning is *not* required.

Additional (optional) context to help motivate interesting discussions will be marked by little information icons like this in each of the modules.

How to Use This Book

Everyone learns differently. Some find it most helpful to self-study the material, which usually can be done in a weekend according to prior students. We believe the most effective and fruitful way to learn the content herein is in **small, discussion-based** classroom environments. Many of the discussion questions included in each module have *no single right answer*, and it is often helpful to learn how peers are thinking about challenging topics, too. To facilitate this learning environment, the content has been broken down into five modules that can each be covered in one-hour sessions.

Optional follow-up discussion questions and food-for-thought are marked with a lightbulb to help you explore topics further.

What This Resource Is *Not*

There are a lot of machine learning resources out there already - what makes *Ethical Algorithms for the Modern Clinician* different? This resource is ***not*** any of the following:

1. **A introduction to theory of and program-**

[1] Optional citations for digging deeper into interesting topics will be included in the margin like this: Fowkes FG, Gehlbach SH, Farrow SC, et al. Epidemiology for medical students: A course relevant to clinical practice. Int J Epidemiol 13(4): 538-41. (1984). doi: 10.1093/ije/13.4.538. PMID: 6519897

ming for machine learning: If you're interested in building ML models from scratch, the MedML@Emory Club has put together a fantastic tutorial here, and Andrew Ng also has a great curated list of technical tutorials here.

2. **An overview of the computational techniques used to build machine learning models**: If you're looking for something like this, we recommend checking out recent work by Pfob et al. (2022) and Sidey-Gibbons et al. (2019).
3. **A mathematically rigorous foundation for assessing algorithmic fairness and privacy**: If this is something you're interested in, we highly recommend checking out The Ethical Algorithm and the Ethical Algorithm Design course offered at Penn every year.

About the Authors

Michael Yao is an MD-PhD candidate at the University of Pennsylvania, and has formerly worked at companies such as Microsoft Research, Scale AI, Glass Health, and Hyperfine. His current research is in trustworthiness and robustness for deep learning and offline optimization problems. Michael is broadly interested in developing methods that leverage prior knowledge to help algorithms better generalize to new distributions.

Allison Chae is an MD candidate at the University of Pennsylvania researching how AI algorithms can be used in real-world clinical workflows. She currently works alongside radiologists to better understand how physicians and computer systems interact with one another. Allison's research is supported by the AΩA Society and the Perelman School of Medicine.

Expert Collaborators

Ethical Algorithms for the Modern Clinician is backed by experts in both clinical medicine and machine learning.

1. **Mark Yatskar, PhD** is an Assistant Professor of Computer and Information Science at the University of Pennsylvania. His lab studies the interplay between visual and language systems in machine learning, and how such systems can inherit human biases.
2. **Walter Witschey, PhD** is an Associate Professor of Radiology at the University of Pennsylvania. He leads an interdisciplinary lab studying MRI imaging techniques and machine learning for clinical radiology.
3. **James Gee, PhD** is Professor of Radiologic Science in Radiology at the University of Pennsylvania. His lab conducts research in biomedical imaging analysis and computational methods for both ex-vivo and in-vivo medical imaging.

Institutional Partners