Ethical Algorithms for the Modern Clinician

Zoe I Barinaga[1]   **Allison Chae, MD**[2]   **Daniel Hashimoto, MD**[1]   **Joy Jiang, PhD**[3]
Emily L Leventhal[3]   **Sabrina Meng**[3]   **Morgan Purcell**[1]   **Michael S Yao, PhD**[1,§]
[1]University of Pennsylvania   [2]Main Line Health   [3]Icahn School of Medicine at Mount Sinai
[§]Correspondence to `michael.yao@pennmedicine.upenn.edu`

Overview.   Ethical Algorithms for the Modern Clinician (EAMC) is a short series of educational modules intended to teach current and future clinicians the basics of machine learning (ML) and artificial intelligence (AI) as they pertain to clinical practice. This resource introduces ML from a clinical practitioner's perspective and discusses what ML does right, where it falls short, and how it may impact patient care. EAMC consists of 5 standalone modules that are independent of one another.
1. **Introduction to Machine Learning**: What is machine learning? How is it similar to and different from conventional software? When should I consider using machine learning?
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 maintain patient privacy?
4. **Algorithmic Interpretability**: What does it *mean* for an algorithm to be interpretable? Is it important for us to be able to explain how an algorithm works 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?

Description of the Material.   All standalone modules described above consist of the following materials:
- **Written Notes**: Each module consists of short, interactive notes made accessible as both **digital handouts** and PDF **lecture notes**.
- **Lecture Slides**: EAMC is also made available to be taught in small discussion-based groups typical of medical school learning environments. We provide interactive **lecture slides** and accompanying clinical cases to adapt each module to this setting. A sample **course syllabus** is also made available to accompany the set of lecture slides.

Furthermore, some of the modules feature the additional, **optional** supplementary materials:
- **Interactive Tutorials**: Introduction to Machine Learning, Bias and Fairness, and Algorithmic Interpretability feature optional, interactive tutorials for interested readers to explore the topics discussed in the module firsthand. **No pre-existing knowledge of Python or coding is required.**
- **Lecture Recordings**: The Introduction to Machine Learning, Bias and Fairness, and Algorithmic Interpretability modules include sample lecture recordings made available on YouTube.

Target Audience.   EAMC is designed for **current and future clinicians** *regardless of technical background*. Learners that may particularly benefit from our educational materials include medical students, resident physicians, attending physicians, nurses, and hospital administrators.

Expected Reading Time.   Each of the 5 modules requires no more than **10-15 minutes** of reading time.

Preliminary Reception.   Preliminary versions of EAMC were previously piloted at the University of Pennsylvania School of Medicine and Icahn School of Medicine at Mt. Sinai, and was positively received by both participating students and faculty. For the purposes of running these pilots at the authors' home institutions, early versions of the course material were made available internally at eamc-penn.github.io. However, **these materials were never published or advertised online to the general public**.

---

*Authors are listed in alphabetical order by last name. All authors contributed equally and may list themselves as co-first authors.