Medical Machine Learning technologies as an example for (necessary?) ethical trade-offs in ML

Arthur Jaques - 5998179
Julian Petruck - 3857386
Voulgaris Sotiros - 6013093
Thomas Albrecht - 5733587

We use the application of Machine Learning to healthcare as a case study of ethical trade-offs. We concentrate on trade-offs between privacy and predictability in the use of patients' data, between group fairness and individual fairness in the attempt to make ML-based systems "fair", and between fairness and prediction accuracy when applying fairness constraints to the ML systems. Firstly, we examine and discuss whether those tradeoffs are unavoidable, and relate them to moral dilemmas in moral philosophy. Secondly, we examine the results that are obtainable with regards to those tradeoffs (where do we want to lie on the Pareto frontier?). In the case of the trade-off between group fairness and individual fairness, we dive into the conflict between the aggregate and the individual, between the population level view of the "average man" and the concrete individuals that are affected by the ethical policies. In our critical analysis, we relate the existing best practices in medicine and their existing literature (as an example, the four principles proposed by Beauchamp and Childress), and the fairness tools and analyses provided by the ML community. As a consequence, we suggest what the communities could learn from each other and what differences need to be resolved.

In general 3 areas that we should focus:
AI-ML // Healthcare // Fairness and ethics

Other topics: Accountability(explainable ML?), privacy, ???

1. Find the connections between fairness and ethics from a general technological perspective (don't try to get into philosophy too much unless julian wants to)
2. How AI/ML is applied to current health care setting (this can range from automated health decision in a hospital to health insurance to clinical trials to diagnosis to human classification in a very broad sense)

Tradeoffs:
check how the trade off are handled right now

1) Privacy against predictability in medical sciences(the cost of privacy)
2) Group fairness vs individual fairness
3) Fairness vs accuracy

and hopefully a specific real use case we can look at at more detail
Compare "beauchamp and childress" principles to existing approaches to ML fairness
(also julian proposed some papers on this topic)

and of course point out the problems (this can be a new topic but don't only focus on the obvious race problem)

3. How ethics in such settings can guide us to a more fair/ethical ML

And finally the big question is how do we combine all that at the end. What is our aim? What do we want to say? What is our critic about?
Some shallow ideas:
- Opinion on why and if it is a good idea.
- Regulations and what are the tools we need

From the facct conference: (anything related and maybe I am missing something)
*(Dissecting Racial Bias in an Algorithm that Guides Health Decisions for 70 Million People) https://dl.acm.org/doi/10.1145/3287560.3287593 Also Data For Black Lives
*A Taxonomy of Ethical Tensions in Inferring Mental Health States from Social Media (not so relevant)
*Algorithmically Encoded Identities: Reframing Human Classification
*Ethics on the Ground: From Principles to Practice
*Fairness, Accountability, Transparency in AI at Scale: Lessons from National Programs
*Reasoning About (Subtle) Biases in Data to Improve the Reliability of Decision Support Tools
*https://www.youtube.com/watch?v=yJtzU8YbTxc (Health, Technology, and Race)
*Tutorial: From Publishing to Practice: Bringing AI Model Monitoring to a Healthcare Setting
*More from John D Dickerson and Liz O'Sullivan??

Konstantin Genin: If we would work harder, certainly we would not need to have a trade off! Philosophers do not know about trade-offs - trade-offs are moral dilemmas.

Maybe related recent paper also by Genin: Randomized Controlled Trials in Medical AI: A Methodological Critique (http://philmed.pitt.edu/philmed/article/view/27)

Further articles:
The Ethics of Machine Learning in Medical Sciences: Where Do We Stand Today?(https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7640783/)

Ensuring Fairness in Machine Learning to Advance Health Equity

Arthur's notes from the meeting:

Tradeoff <-> moral dilemma (some philosophers doubt their existence)
Pareto frontier
2 questions: does there exist a tradeoff in the first place? What weights do we assign?
2 axes to develop:

1) Tradeoffs (see above)
2) Going down from the aggregate to the individual. “The average man” and the population level view. The clash between policies (aggregate level) and the carriers of the ethical consequences (individual level). Examples where a policy modifies the people who receive harm (eg vaccine).

Compare best practices in medicine with ML solutions