Nels Schulzke

Apr 2, 2024

Dr. Neeley Tass

STAT 386

The Ethical Algorithm

"The Ethical Algorithm" claims that the ethical concerns and social implications of algorithmic decision-making are significant and need to be taken into account in algorithm design. The authors present principles for socially-aware algorithm design, including fairness, privacy, and transparency, and discuss different approaches to addressing the challenges of algorithmic discrimination and bias, such as algorithmic auditing and accountability.

A few of the insights that stood out to me most were the use of randomized polling, differential decision making, and the impact of user game theory on the training and application of machine learning (like how they talk about in the navigation example. For me, the most interesting of all of them was probably randomized polling and how that preserves both aggregate measures and personal privacy. I think that is so valuable and I guess I had not thought of it that way before. That also makes it researcher-blinded, which is so powerful.

The ethical concerns that stood out most were the questions of data privacy. Especially the example used in the Governor of Massachusetts’ medical data being found. That made it so evident to me that there is a need for serious consideration of data protection and privacy on all fronts.

Proposed solutions to the issue of privacy (this is the issue that impacted me the most) in the book are the differential decision making, supported by the use of randomized polling in the collection of data. This was very enlightening to me.

The authors view technology as both a tool for empowerment and a source of concern. Throughout the text, I see them acknowledge the potential benefits of algorithmic systems for improving efficiency and highlight the ethical and social risks associated with algorithmic training and application and how that can lead to discrimination and bias.

An example that reflects the themes of the book is the use of predictive algorithms in the criminal justice system to assess the likelihood of defendants reoffending or being a flight risk. These algorithms have been criticized for reproducing racial biases and perpetuating inequalities in the criminal justice system. This highlights the challenges of quantifying and mitigating algorithmic bias and the need for more transparent and accountable approaches to algorithmic decision-making. It also underscores the wider ethical concerns regarding the use of algorithms in domains where decisions can have significant social and legal consequences, and the need for interdisciplinary collaboration and public engagement in addressing these concerns.

The book offers several hopeful insights, I will be honest, it was pretty hopeful all in all and I didn’t feel anything particular stands out above everything else. The book is very solution-oriented in my mind, which is hopeful inherent.

I would rate the book 7 out of 10. I get bothered by pandering to social politics, where, for example, the authors even reference, in a very tongue in cheek way, identity politics and intersectionality, which in a philosophical and political context has been de-legitimized broadly. The book seems very focused on the equality of outcome, and that is not a positive thing. Equality of opportunity is relevant, but I felt it dwells a bit too much on equality of outcome, which is distinctive. But, as a positive, I really enjoyed its consideration of data privacy and especially its mathematical explanation of different approach to resolving that data privacy. Additionally, the discussion on the impact of game theory decision making in the development and training of auto regressive machine learning models was really worthwhile. This also addresses why I would recommend the book to someone else. I specifically would and have now recommended this book to my dad, who is a statistician and professor at another university. His research is based a lot in game theory and algorithm training, so this seemed particularly valuable to him.