Weapons of Math Destruction Book Review

I chose this book because of my experience in data and working with other data professionals. I am very aware of the laissez faire attitude of many data scientists and people within the tech industry regarding data privacy and the implications of their work. In one conversation, my friend justified getting more data by saying “Well I won’t be doing anything bad with it”. This didn’t bother me as much as it should have at the time but upon reflection this attitude is at the heart of the problem. As O’Neil points out WMDs or weapons of math destruction can come out of good intentions.

The book goes over many different examples of WMDs and breaks down why they are bad and what impact they have had on society. She goes over three different criteria for a WMD which are opacity, scale, and damage, all of which are used to see if something can be considered a WMD. Keep in mind that a model doesn’t have to be a WMD to be dangerous or even harmful, it will just have less of an impact unless it covers all three aspects. O’Neil goes into several different examples of these in action such as college admissions, job applications, as financial instruments, and prison and crime prevention.

Regarding the first aspect these models tend to be black boxes for various reasons. Sometimes they need to be very complicated or sometimes the company or industry will keep them a secret. She later makes a point about how ”Many people are unfortunately intimidated by math”. This further detaches people who are affected by these models from the actual operation as many people are not in a place to contest or even think they would understand the inner workings.

Another interesting point is how bias transfers over to the models and just because it is a “machine” making the decisions, or classifications, that doesn’t separate inherent human biases or flaws. A great example of this from the book is the implementation of a crime prevention algorithm. On the surface it sounds like a great idea, policing could be made more efficient by being at the right time and dividing resources in a data driven way. In reality, the model is only as good as the data that is fed into it, which can lead to many unintended or perhaps intended consequences. Because poor neighborhoods have had a higher police presence there tends to be more data points, this in turn will lead to the model recommending that patrols be sent out to those neighborhoods. This has the potential to turn into a positive feedback loop. Because more people are caught committing crimes in a certain neighborhood more patrols will be sent, meaning even more people will be caught in the future. This can be a problem of discrimination especially because crimes such as possession of marijuana are disproportionately enforced in these neighborhoods while the consumption of marijuana isn’t that different among rich and poor.

I think my main takeaway is to always be vigilant and continuously be aware of how my work will affect others. I haven’t met any data scientists who are actively looking to do harm but I have met many, especially those who haven’t come from an academic background who don’t have as much concern about the bigger impact. Like in her example of Google mislabelling three African American people as gorillas in an automatic tagging system, “it was most likely faulty machine learning (and probably not a racist running loose in the GooglePlex)”. One more example is when she is referring to her teaching an ethics class to data scientists. They are discussing various features that could be fed into a model and which might prove to be problematic, with race being an obvious omitted feature due to discrimination concerns. She then moves onto Zip code with many of the students saying they thought it would be fine to use. The problem with using Zip code as stated by O’Neil is that it takes into account historical prejudice which could also include a racial component. The point being that you need to think carefully about all your assumptions and the implications that they bring into the model.

Anothre relevant takeaway from this book that I had not really recognized before is how these models feed into inequality. Not necessarily because they are designed that way, although some actually are like pay day lone targeting campaigns, but because the people who are designing them and paying for them come from wealthier and more educated populations. They can be expensive and difficult to implement, with Data Scientists being paid very well for their services. This means that the goals of the models carry with them those inherent biases. In a recent statement Joseph Stiglitz warned that the society is measuring success by the wrong metrics. He said that “If we measure the wrong thing, we will do the wrong thing”, which sounds like the opening statement made in a machine learning 101 class. In this context he means that wealth inequality, climate disasters, and democratic crisis will continue to grow worst if we are optimizing for the wrong things such as GDP and profit.

In conclusion I agree with her final chapter that data professionals should hold themselves and their peers to a higher standard, to recognize their impact and strive to be better. I have seen this already in the Data science community here in Vancouver. Looking at the Datacamp scandal, where the CEO was involved in a sexual misconduct incident and an attempted cover up, many instructors and community members migrated away from the platform with courses being put up for free on other platforms as a protest. Much as sustainability and social metrics are just starting to be used in financial metrics there is a push to quantify these and have them impact other models as well. Many of the problems highlighted in the book can be attributed to bad data science and statistics as a means to reach the desired goal, only if we change the goal and hold data professionals, corporations, and regulatory bodies accountable will we start to see changes to these systems.

References:

ONeil, C. (2017). Weapons of math destruction: how big data increases inequality and threatens democracy. Great Britain: Penguin Books.

Queally, J. (2019, November 25). 'Everything Is Not Fine': Nobel Economist Calls on Humanity to End Obsession With GDP. Retrieved from https://www.commondreams.org/news/2019/11/25/everything-not-fine-nobel-economist-calls-humanity-end-obsession-gdp.