

Machine-Learning, Artificial Intelligence & Society

by
Christensen Raveendran

Master's Seminar
Techno-Intimacies
Tom Engels

Royal Academy of Fine Arts, KASK
Ghent
Belgium
2021

In 2002, *Minority Report* a film based on a book written in 1956 was released. It tells us, the story in the near future, where there exists a special police division, PreCrime, that as the name suggests, prevents crime before it has even been committed. This happens with the help of the three siblings, *Agatha*, *Arthur* and *Dash*, who all have individual special powers but are amplified when *connected* with one another, to predict disturbances in peace and quite literally 'see' the crime taking place in their visions.

These visions are connected to a special screen where the police are able to watch along and then go the place of future crime and essentially wait for the soon to be perpetrator and arrest him/her before the crime officially takes place. This sounds science-fiction, and well it is, but in a world not too far away similar things are happening with AI. Replace the three siblings with computers, data and machine learning and world of predicting if you can be allowed to own a house, get a loan or even gain employment in a job is right here.

The promises of Artificial Intelligence and its current reality.

Artificial Intelligence, from now on referred to as A.I, and Machine Learning (M.L) are not new technologies as the systems have been used as far as back as the 1950s (Kaplan & Haenlein, 2019). But A.I is talked about more now than ever as the difference is, in the early applications it was quite novel (used to simulate a human in the playing of games, chess or checkers) but today we are using it for a variety of more life impacting applications (apologies to professional chess and checkers players). Sundar Pichai, the CEO of Alphabet Inc and Google, has even described AI as "more profound than electricity and fire" (Ignatius, 2020)

AI is being used for crime prevention, banking, hiring employees, offering drivers directions to avoid high traffic, managing physical health and much more. It is being used by governments, multi-national companies and with the advent of 'no-coding' platforms like Amazon SageMaker and Microsoft Azure, even smaller companies and individuals are using it as it no longer requires data scientists to create complex machine learning algorithms.

The advantage of AI and ML is that it has the ability to breakdown and drink up unimaginable amounts of data to produce an output but it does not guarantee a usable output especially if the data it is drinking is tainted.

Let's look at the story of Mikhail Arroyo and the problems he faced with AI.

In 2016, Mikhail and Carmen (his mother) were applying for an apartment together in Connecticut, USA. But quickly after the application, Carmen found it was denied and that outcome was caused by a flagging in the AI system, Crimcheck, that the leasing company uses to decide who to lease an apartment (Lecher, 2019).

Through this system, Mikhail was deemed DOES NOT MEET THE CRITERIA FOR THE COMMUNITY in status as he had obtained a logged criminal record for a past offense. Carmen, understandably as we will find out later in this essay, was confused and felt this should be overturned. Reason being Mikhail had been logged, many years ago, just one record in the system for shop-lifting items of under USD150.00 in value (the charge was later pulled) and in 2015 had an accident leaving him in a coma and eventually in a physical state where he was unable to speak, walk or care for himself.

Carmen, his mother, was his legal care-giver, but due to the AI system was denied the ability to move into an apartment together. The leasing company did not overturn the decision despite the cries from Carmen and instead told her to contact the company that controls the AI system, Crimcheck, instead.

Without any other option, and after months of attempting to reach Crimcheck, Carmen and Mikhail had lost the apartment as it was rented out to someone else.

Mikhail's problematic coming together with AI is a not a new one, there have been many people wronged by AI and as AI continues to converge on different parts of everyday life, the question of how to tune out problems of having no human interaction, bias and stereotype continue to plague us.

Crimcheck, founded in 1991, is an American company that provides an automated tool that processes and interprets criminal records (Crimcheck Holdings, LLC, 2021). They provide an actively updating record of criminal activity that companies employ its systems use to determine a certain specified output.

In the case of leasing companies and landlords, they use it to decide if someone is able to lease a home without even meeting the potential applicant and since this is seen as the first channel to remove 'unwanted' applicants, it can be done completely algorithmically.

What it boils down to is that It determines if you are fit for something based on a set of parameters.

There are two types of AI, a strong and a weak. What we see in movies like the Terminator, Ex Machina and such is strong AI (unsupervised AI) and scientists say this type of AI that takes over is highly speculative (Neville-Neil, 2016) and not within current horizons of possibility. Weak AI (supervised AI) is what we see now with algorithm based softwares like voice-activated personal assistants in Google's Alexa and Apple's Siri, Spotify and YouTube's 'suggestions based on previous plays' and to Crimcheck, systems that enable companies to make specified decisions given data. Supervised AI gives the ability to present a human-like experience via an algorithmic simulation

AI screening tools like Crimcheck, allow landlords to remove themselves from the decision making, to screen out the 'absolute no's' leaving them more time to speak to clients with a so called 'higher potential'. The problem with the breaking down of the individualized system is that it bypasses human forms of critical decision making it more likely to turn complex matters, like someone's life and abilities in the case of Mikhail, into simple pass-fail mechanisms working on maths. Extenuating circumstances, like Mikhail's health, physical ability and severity of past crime, is not captured in the rigid mathematical algorithms that power the system.

This was, is and will continue to be problems alienating people who otherwise would be accepted in a traditional form of acceptance.

In computer science, the concept that has flawed inputs (data) will result in flawed outputs is what popular term "garbage in, garbage out" (Kilkenny & Robinson, 2018) is and ML is extremely prone to falling in that hole. One of the most glaring examples of "garbage in, garbage out" is when in 2016, Microsoft launched a twitter bot that aimed

to 'converse' with Twitter users. In just one day, it changed its 'ethical mind' from saying "Humans are super cool" to praising Hitler and posting misogynistic remarks (Kleeman, 2016). All of this was due to the data that was (quite literally) Twitter users teaching the bot words and concepts. It illustrated the truth, albeit scary, about machine learning that it learns exactly what it is taught.

In 2016, ProPublica, posted an extensive study that proved the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), an AI/ML system that makes recommendations for criminal sentences used extensively in Florida was biased against African Americans (ProPublica, 2020). The system recommended African Americans twice as more as whites to be 'higher-risk' but never actually re-offend, in return on the flip side also wrongly labelling whites, twice as more as blacks, as 'lower-risk' but actually re-offend.

Racial stereotypes have existed for centuries so perhaps to avoid this bias, we should just leave variables like 'race' out of the algorithm and perhaps this will rectify bias? It sounds good but it has also proven that bias can still leak into the data.

In 2018, Amazon uncovered that their ML recruiting program was biased against women despite not having 'gender' as a variable. Because the system was trained with data consisting of resumes submitted to the company over a ten-year period and observing the patterns of successful and unsuccessful applicants (Dastin, 2018). The flaw was that there is a historical gender bias in the tech industry, as the industry has been highly male centric (White, 2020). Essentially the system taught itself to penalize words like "women's" in terms like "women's chess club" and even penalized all women's colleges.

All while it amplified 'masculine terminologies' with resume's that consisted of words such as "captured" or "executed" (Dastin, 2018).

Most recently, Joy Buolamwini, a researcher at MIT wrote a landmark paper highlighting the error rates in 3 of the major commercially available facial-analysis programs. She revealed the error rates in analyzing gender of a face was less than 0.8% for white men in all programs. However, that rate was a lot higher in the determination of darker-skinned females, more than 20% in one and 34% in the other two (Raji & Buolamwini, 2019). Buolamwini says these techniques that analyze gender based on facial recognition are also being used to identify a person when looking for a criminal suspect and other applications where accuracy is vital. Having AI facial recognition be so imbalanced in accuracy is problem as these systems are already in place despite the known errors.

The cause of this error is again in the data fed in, if a disproportionate number of facial images are white men, then the result will be more accurate for white men.

“Most people creating AI and the data being used do not reflect the vast majority of the world. As a result, there are major blind spots in the technology being created,” as Buolamwini says is a call to action in claiming and reclaiming technology that has been inherently biased in a world where social media and technology is a cornerstone of the industrialized civilization. Technologies like AI, ML and facial recognition are being used, now, in mass government approved surveillance in many parts of the world and in every continent. We are using inherently biased and stereotypical systems to facilitate national security. What happens if a person gets misidentified as being a criminal? These are

real problems being faced by society and it is caused by technology that knows it is biased but is it being implemented. We have a tendency to overplay effects of technology in the long run with the thoughts of AI becoming 'human' when the case is, it is unlikely to achieve 'movie-level AI'. But we also have the tendency to downplay the impact of technology in the short run, implementing under-tested and biased software and allowing the general public to 'beta-test' whilst genuinely affecting many and society as a whole.

The stakes are getting higher and if trends continue, AI and ML will be increasingly used to help make decisions that affect companies positively in revenue as lesser employees are needed but affect the every day person negatively as they are shunned out. Bias imposes a serious cost on society by magnifying stereotypes and inequalities and unfairly denying opportunity. The applications of AI are done with good intentions, for the most part, but the knowledge that these systems are faulty, at this moment, is important. Mikhail was flagged by the AI system but he really was denied by the leasing company because they could have overturned it, instead they pushed the blame to the AI company. As of 2020, the lawsuit is against the AI agency and not the leasing company, this is a potential flaw in the system as it legitimizes bias and stereotypes. In a case where a human makes openly bias decision, it can be blamed to an individual but what happens when a machine makes the bias? who's to blame? If untouched or investigated about then mechanized stereotypes on gender, race, sexual orientation and more will become a part of life that will unnoticed.

Bibliograph:

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–25. <https://doi.org/10.1016/j.bushor.2018.08.004>

Ignatius.Bowskill-Dutkiewicz@digital-startup.com. (2020, July 6). "More Profound than Fire and Electricity." AI by Alphabet. *Disruption Banking*. <https://disruptionbanking.com/2020/01/24/more-profound-than-fire-ai-by-google/>

Lecher, C. (2019, February 1). Automated background checks are deciding who's fit for a home. *The Verge*. <https://www.theverge.com/2019/2/1/18205174/automation-background-check-criminal-records-corelogic>

Crimcheck Holdings, LLC. (2021, January 7). FCRA Certified Background Checks & Screenings. *Crimcheck*. <https://crimcheck.net>

Neville-Neil, G. (2016). The Chess Player who Couldn't Pass the Salt. *Queue*, 14(6), 14–18. <https://doi.org/10.1145/3028687.3029797>

Kilkenny, M. F., & Robinson, K. M. (2018). Data quality: "Garbage in – garbage out." *Health Information Management Journal*, 47(3), 103–105. <https://doi.org/10.1177/1833358318774357>

Kleeman, S. (2016, March 24). Here Are the Microsoft Twitter Bot's Craziest Racist Rants. *Gizmodo*. <https://gizmodo.com/here-are-the-microsoft-twitter-bot-s-craziest-racist-ra-1766820160>

ProPublica. (2020, February 29). Machine Bias. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

Dastin, J. (2018, October 11). Amazon scraps secret AI recruiting tool that showed bias against women. *U.S.* <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>

White, S. K. (2020, January 23). Women in tech statistics: The hard truths of an uphill battle. *CIO*. <https://www.cio.com/article/3516012/women-in-tech-statistics-the-hard-truths-of-an-uphill-battle.html>

Raji, I. D., & Buolamwini, J. (2019). Actionable Auditing. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 1–2. <https://doi.org/10.1145/3306618.3314244>