Optimization of Sustainable Energy Systems through Data Engineering

Meta Analysis of existing solutions regarding bais AI data training

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
McKayla Thomas

October 27, 2022

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Professor Wayland, Department of Engineering and Society

Brad Campbell, Computer Science

Introduction

AI usage is becoming more common as the technology industry continues to grow. Government and private companies are using AI in various fields such as policing, health, housing, workplace and education. Although AI is notoriously known for being biased against people of color and the disable community, severe consequences are often overlooked. There has not been much action to prevent further damages compared to the speed of technological growth. The issue is artificial intelligence is built by humans and therefore is inherently susceptible to the intolerance of minority groups. The data collected and used to train the AI are often unrepresentative of marginalized groups. The one sided training of the AI systems directly correlates to the exacerbation of preexisting prejudice [5]. By acknowledging the downfalls of unbalanced training datasets, I was able to identify trends of bias and design a dataset during an internship last summer.

The company I worked for was a sustainability energy company focused on the optimization of the buying, selling, production and movement of energy resources. My task was to compose and implement an EIA renewable energy generation dataset into the already existing data warehouse. I had to read through hundreds of zip files containing detailed information about the generators and create the appropriate models. As I was completing the task, I was researching how to properly design datasets. That is when I realized that it was common for datasets to be subjected to the bias of the person creating them. My topic is not limited to a certain section because the effects of the discriminatory practices of AI are felt in every sector of society. I want to present the danger so it is obvious how AI bias can grow and harm the progress of society as a

whole. For my STS topic, I will be exploring the effects of unfair dataset designs in AI and how different groups strategize to combat these issues.

Technical Problem

Unfortunately, government documents are usually scattered and vary in jargon. This makes it difficult to utilize the data for any type of project. For instance, the EIA has yearly filings(860A/860B) of the status of the energy generators throughout the US. The keywords and codes may have different meanings based on the year of the filing. This is an inconvenience to any customer looking for up to date information about the energy generation. The accessibility and centralization of the data allows for customers to visualize the possible usage of the dataset. This dataset was introduced to the data warehouse so users can easily access the information via a frontend application. The project required the usage of various tools and software such as React(Javascript), Django(Python), Docker, and MySQL. When designing this dataset, I had to relate the individual models to each in a way that did not project my own personal bias into the data. When data is sorted and organized without taking into account the effect of unfair datasets, minority groups are disproportionately affected.

STS Research Topic

I am studying different solution the problem of discriminatory datasets used to train AI. In addition I will be focusing on the negative effects of these datasets on minority groups. My developer prospective will offer a deeper dive into the technological issues along with the societal problems. My STS research topic challenges society to not only look at the convenience Artificial Intelligence has offered but also the effects these technologies have on minority communities. It has become a norm to let these obvious issues go unchecked. So in response the

question arises, What harm has improperly designed datasets already caused when used to train AI?

In the education system AI is now being used to detect cheating on assignments and exams. The problem of the usage of these testing software is students of color and students with disabilities are being inaccurately flagged and as a result locked out of exams due to the discriminatory guesses the AI makes. These guesses are developed by datasets that are organized based on the physical traits of the student taking the exam. Based on this information, the program stores the data as high risk leading to misguided assumptions [8]. This is a direct result of developers placing their own bias in their programs. To combat this problem, developers should allow many people with different backgrounds to assist in editing of the code. During the editing process, the programmers should focus on equality despite personal beliefs. This can be achieved by a consistent user flow for all students and allocating more attention to video processing to detect unauthorized materials. Issues like these cause the European Union and other organizations to design practicable solutions.

The European Union has an antidiscrimination law that is equipped with an appropriate doctrinal tool kit to face the fast growing evolution of AI technology. The law provides legal recognition of indirect discriminations, which no longer require certain proofs of causality, but put the focus on conspicuous correlations, instead [9]. Due to the focus of this law, it depends on knowledge about vulnerable groups, both on a conceptual as well as on a factual level. In addition to some interference from governmental system, organizations such as ICSR have dedicated themselves to combating AI discrimination in all sectors. ISCR was inspired by the

benefits AI offers to society but also realized the issues that currently plague the technology. ISCR often discusses the issue of the government gathering large amounts and how these dataset are organized in bias matter then filtered into AI systems. But ISCR's current focus is lack of correction of the companies spearheading the AI industry such as Google, Amazon and Meta(Facebook) [10].

Facebook uses AI for targeted advertisements. The idea behind targeted advertising is that Facebook's data helps advertisers predict which users will respond to their ads [3]. In the recruitment context, this means Facebook's tools help employers predict which users are most likely to apply for their jobs. The datasets used to train Facebook's recruitment software is similar to the company's ad software which is known for displaying alcohol and drug addiction services to minority groups. This directly translates to why the recruitment software disvalues people of color [7]. Amazon also uses bias datasets to train their facial recognition software. Amazon has spent millions of dollars developing their facial recognition system named Rekognition. The software has issues identifying darker-skinned women and men. Even after the AI program learned from data containing millions of pictures inputted by engineers, the program regularly misidentified women for men when the people had darker skin. It was later discovered that this was due to the software being fed more pictures of lighter skinned people [4]. Technology like Rekognition are being used in vital parts of the criminal justice system but the full usage is still limited when it comes to government decision making. The government usage seems to be centered more around statistical analyses. These statistics consist of data about cognitive abilities based on demographics. This data is distributed throughout the world to be

used for arguments that can be discriminatory[6]. To combat practices that benefit from prejudiced AI systems, many believe True AI could be a valuable tool.

It can be difficult to tell if a system is True AI. What is True AI? True AI consists of proactive characteristics, self-learning algorithms, and real-time decisioning. True AI requires the system to be completely autonomous, meaning it can function without human maintenance. The argument that True AI could be a way to implement the technology without bias does not take into account that these systems learn through data experiencing, invidious presumptions are possible. This is likely since society exhibits unfairness. So, the question is can AI ever be true? Questions like these must be answered before these systems are implemented further. Amazon's Rekognition software is not a true AI system but is advertised as minimal human involvement, and is being used by the criminal justice system. For instance, an African American man was wrongfully arrested for shoplifting. The reason for this was the facial recognition algorithm had matched the photo on the man's license to unclear camera footage.[2] Amazon produced AI is affecting two of the major sections of society, the first being criminal justice as I previously mentioned and the second being the employment arena. Despite the increase in difficulty to obtain jobs upon graduating college, Amazon decided to use AI in order to select candidates through their artificial intelligence-driven software. The software is designed to distinguish the most fit applicants for a job position. [1]The data that is feed to the software contains the race, sexuality, possible disabilities, and sex about the candidate. However, the software algorithm was biased against women and routinely concluded that men were preferable candidates.[4] Amazon has been aware of this bias but similar to other large corporations but has not done much to fix the issue. After considering the presented examples of the existing problems with AI,

the most clear patterns consist of the non representation of minority groups in training data and organizational discrimination.

These issues could undo much the progress society has made to rectify centuries of unfairness against marginalized groups. In order to prevent further usage of clearly flawed artificial intelligent software, developers need to identify the bias in the dataset and design the architecture of data warehouses in a preventative way in order to compensate for the existing bias. Also developers and companies should use the available resources about how to design datasets without as much bias.

Conclusion

Artificial Intelligence offers many benefits to societal development but like any system, minority groups must be considered to prevent unfair treatment. The datasets used to train AI systems need to promote equality. If the bias continues to plague the AI industry, it can very well become more dangerous for marginalized groups. Some of the existing negative effects were highlighted previously. The consequences are not limited to one aspect of society but many structures vital to the overall well-being of an individual. It is the responsibility of developers to design transparent datasets and research the proper way to do so. I followed this during my summer internship. When creating my Energy Information Administration dataset, I researched heavily how to limit bias as much as I could. This is when I found the ISCR and FDA information about identifying unfairness in data and how harmful it is if not addressed properly. I was able to build a template to assist me in the design. Through the design process, I consulted people with many different backgrounds. I used their input to adjust the organization of the data

to lessen the bias in the software. The software is still being used by many influential people to help them transition or create sustainable energy systems. My hope is as AI technology continues to advance, we implement ways to effectively use and train it using datasets that have been vetted by people from different backgrounds.

References

[1] Allen, R., Masters, D. Artificial Intelligence: the right to protection from discrimination caused by algorithms, machine learning and automated decision-making. ERA Forum 20, 585–598 (2020). https://doi.org/10.1007/s12027-019-00582-w

[2] Hoffman, S., & Podgurski, A. (2020). *Artificial intelligence and discrimination in health care.* Yale Journal of Health Policy, Law and Ethics, 19(3), 1-49.

[3] Kim, P. T. (2019). *Big data and artificial intelligence: new challenges for workplace equality*. University of Louisville Law Review, 57(2), 313-328.

[4] Lifshitz, B. (2021, May 6). Racism is systemic in artificial intelligence systems, too. Georgetown Security Studies Review. Retrieved October 9, 2022, from https://georgetownsecuritystudiesreview.org/2021/05/06/racism-is-systemic-in-artificial-intelligence-systems-too/

[5] Leavy, S., O'Sullivan, B., & Siapera, E. (2020, July 28). Data, power and bias in artificial intelligence.
arXiv.org. Retrieved October 27, 2022, from https://arxiv.org/abs/2008.07341

[6] Paul Henman (2020) *Improving public services using artificial intelligence: possibilities, pitfalls, governance,*
Asia Pacific Journal of Public Administration, 42:4, 209-221, DOI: 10.1080/23276665.2020.1816188

[7] Santos, I. (2021, July 13). *How artificial intelligence can deepen racial and economic inequities*: News & commentary. American Civil Liberties Union. Retrieved October 9, 2022, from https://www.aclu.org/news/privacy-technology/how-artificial-intelligence-can-deepen-racial-and-economic-inequities

[8] Silver, J. (2022, October 3). *Algorithms are making decisions about health care, which*

may only worsen medical racism: News & commentary. American Civil Liberties Union. Retrieved October 9, 2022, from https://www.aclu.org/news/privacy-technology/algorithms-in-health-care-may-worsen-medical-racism

[9] Tischbirek, A. (2020). *Artificial Intelligence and Discrimination: Discriminating Against Discriminatory Systems*. In: Wischmeyer, T., Rademacher, T. (eds) Regulating Artificial Intelligence. Springer, Cham. https://doi.org/10.1007/978-3-030-32361-5_5

[10] Watkins, S. C. (2022, March 17). How ai can help combat systemic racism. IDSS. Retrieved October 9, 2022, from https://idss.mit.edu/news/how-ai-can-help-combat-systemic-racism/