Bias in AI and Machine Learning: The Impact of COVID-19 in African Healthcare Communities

Andrew Galvin*, Andrew Hogan

Wentworth Institute of Technology 550 Huntington Ave Boston, MA 02115

*Corresponding author

doi: https://doi.org/10.21467/proceedings.157.10

ABSTRACT

AI technology has become increasingly involved in a plethora of societal functions in recent years, but racial bias in AI algorithms has revealed a dangerous trend. With the rapid advancements in technology in general as well as artificial intelligence algorithms, bias is unknowingly developing in these algorithms due to the lack of attention towards it. However, recent efforts have been made to first recognize that bias exists in these algorithms as well as strategies to eradicate it. The implications of the research performed in this analysis go much further than a simple moral obligation to promote inclusiveness for marginalized groups in society; racial bias in AI algorithms has the potential to involve life and death consequences. Specifically in the provision of health, an unbiased algorithm may inherently contain bias due to factors outside of the algorithm itself. It is important to use diverse data sets in our algorithms to ensure that the data does not contain bias. Using a data set which is not diverse may lead to the algorithm developing bias over time, which may cause adverse impacts on patients. In addition, we will discuss how bias affects Africa in comparison to more developed countries. We will look into the future of how we can eliminate bias in artificial intelligence and advance the provision of health more equitably across the global community. Based on findings showcasing examples of racially biased AI technology used to combat the COVID-19 pandemic, efforts currently being taken to eradicate racial bias in AI are highlighted along with a discussion of future actions that should be performed. With more federal regulations surrounding AI algorithms along with an emphasis on promoting diversity in the personnel and data of the AI community, particularly in Africa, the future of AI can be one free of racially biased tendencies.

Keywords: Algorithmic Bias, Diversity in AI Community, Health Implications

1 Introduction

For an Artificial Intelligence algorithm to produce the high level of efficiency sought after in today's standard for the technology, it must be given targeted data in order to learn how to perform its desired task. It therefore is critical that the AI model is trained in a way conducive to the overall use case in question for anyone who may use it in the future.

Unfortunately, this is not the case for many of the AI algorithms currently being used today. There is a systematic issue of bias embedded in the AI tools and models governing our global society today on a racial, gender, and socio-economic level.

In order to prevent bias in artificial intelligence and ma- chine learning, we first have to admit that it exists. In Joy Buolamwini's documentary "Coded Bias," Joy utilized IBM, Microsoft, Face++, Kairos, and Google's facial recognition software. During Joy's benchmark tests on these large corporations' software, she determined that the software performed "significantly better on the lighter faces than the darker faces." (Buolamwini 2020) Joy describes data as "destiny" (Buolamwini 2020) in regard to artificial intelligence. Amazon, for example, has been subject to suspicion of algorithmic bias in its AI hiring technology. The algorithm learns from the pool of resumes submitted to the company over the past ten years, and since a

Proceedings DOI: 10.21467/proceedings.157; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-970666-0-3

vast majority of the applicants were male, the algorithm was trained to favor males over females in the hiring process (Shin 2020). Depicting an example of algorithmic bias from a gender perspective, the flaws in Amazon's AI hiring model suggest a dangerous trend for its other AI endeavors. In order to teach machines, they require big data to train and learn from. However, the data we are using is a "reflection of our history" (Buolamwini 2020) therefore the data sets used in our algorithms do not adequately address the diversity of our global community.

In this paper we discuss biased data and how it affects the provision of healthcare. Through this discussion, we will examine how healthcare disparities may be created based on the evolution of AI after clinical implementation. We show the challenges in building algorithms which do not contain bias and the importance of maintaining them.

We will consider recent efforts to prevent bias in the algorithms used in healthcare. We will discuss the challenges of algorithmic bias in the provision of healthcare as well as ways to prevent it. In addition, we will consider the communities which are the most impacted by bias.

It is important to consider third world countries when looking into the provision of healthcare. Due to limited resources, third world countries are not able to advance as fast as a more developed country. For this reason, we will look into how bias in artificial intelligence is affecting Africa and what steps more developed countries can take to ensure that Africa can utilize artificial intelligence accurately in the provision of healthcare.

Finally, we will look positively into the future of artificial intelligence in healthcare. We will examine possible processes which can be used to help reduce, or eliminate, human bias in our algorithms. It is important for us to develop processes in which we can ensure the least amount of bias in any algorithm that may inadvertently contain bias.

This analysis will aim to answer the question of what efforts should be made to eliminate the presence of algorithmic bias in today's AI tools that have significant ramifications on the people's lives that they are acting on. While it is widely accepted that the data being used to train the AI is the cause of any example of bias (facial recognition technology trained on white faces, hiring models trained on male applicants, etc), it is paramount that researchers develop a solution to prevent embedded bias in future AI technology. In order to discern the efforts that should be made to eliminate algorithmic bias in future AI models, the subsequent sections of this analysis will address the dangerous implications of bias in AI aiding the COVID-19 crisis, particularly in data sparse regions such as Africa.

2 Latent Bias

In this section we discuss the challenges created due to la- tent bias in an unbiased model. The increased recognition of biases in artificial intelligence has caused an increase in the development of unbiased algorithms. However, Matthew DeCamp and Charlotta Lindvall discuss that building an un- biased model "may be only half the challenge." (DeCamp, Lindvall 2020) An unbiased algorithm may evolve itself or through its implementation, which can cause a latent bias. They describe a latent bias as "biases waiting to happen." (DeCamp, Lindvall 2020) Matthew and Charlotta explain the importance of addressing latent biases before implementing AI in the provision of healthcare.

We will discuss two of the three challenges stated by Matthew and Charlotta and how we may be able to avoid them in the future. First and most importantly, an unbiased algorithm may learn biases on its own over time. (DeCamp, Lindvall 2020) This can be due to the algorithm creating correlations between disparity points in the data. For example, the bias correlations may increase over time which would result in less accurate predictions. One way to prevent this would be to re-train the model often with new data and compare the results with the previous model. The second challenge discusses the interaction between

Proceedings DOI: 10.21467/proceedings.157
ISBN: 978-81-970666-0-3

Series: AIJR Proceedings
ISSN: 2582-3922

humans and the algorithm. For example, a doctor may consider the outcome from the model as truth, without question. (DeCamp, Lind- vall 2020) This can be due to the doctor not having sufficient time to research bias in the outcome. One way to reduce the inevitable human error missing the bias in the model would be to ensure doctors are actively monitoring the results of the model and discussing any disparities that may occur.

3 COVID-19 and AI Bias

In order to understand the overall trend of AI algorithmic bias as well as what is currently being done to combat it, it is important to analyze a situation that continues to impact the entire global society. The COVID-19 pandemic has taken a toll on millions of people across the world, and in particular has disproportionately impacted minority groups and underdeveloped countries. For example, research conducted by the British Medical Association found that COVID-19 mortality rates are more than double for people of color and other minority groups in the country than the Caucasian population (Leslie *et al.* 2021). A similar trend can be observed in the United States along with several other developed countries. Unsurprisingly, the root cause of this development is not due to the severity of the virus itself, but rather racial bias in existing AI algorithms used in healthcare systems. Such algorithms have been used to combat the pandemic by aiding in resource distribution such as PPE and vaccinations while also optimizing ventilator and ICU use during the early surges of the virus (Jercich 2020). Since these AI models utilize existing healthcare data for its training data sets, the algorithms are unintentionally preventing the group most impacted by the pandemic from receiving the relief they desperately need. Minority groups in developed countries without access to healthcare are being unfairly neglected by AI models that are deployed on the front lines of COVID-19 relief, thus adding to the existing trend of racial bias in AI technology.

While AI bias is having a measurable impact on pandemic relief in countries with an abundance of data, it poses even more dangerous implications for data sparse regions of the world such as Africa. Global pandemics require a unified and timely response by world leaders with a well-defined plan on how to combat the crisis. With racial bias in existing AI technology, however, it becomes extremely difficult to effectively aid countries in Africa with their pandemic relief. Without large data sets to train AI models and funded re- search into the technology itself, biased algorithms become the only option for less advanced countries in the AI com- munity. As the world continues to utilize AI to tackle global crises such as the COVID-19 pandemic, researchers must focus their efforts on obtaining more consistent data from regions such as Africa while also improving racial bias in AI.

4 Supportive Studies

In this section we will discuss how bias in AI affects health-care and what we can do to prevent it. In Katherine J. Igoe's "Algorithmic Bias in Health Care Exacerbates Social In- equities – How to Prevent it," she discusses how bias in algorithms happen and how it can negatively impact patients as well as what we can do to prevent it. In her research, she discusses how the Framingham Heart Study cardiovascular risk score had a similar outcome to Joy Buolamwini's. This score performs significantly better for Caucasians, but does not perform well for African American patients. In this scenario, it is vital that we are able to create a solution to negate these instances. Katherine estimates that "Caucasians make up about 80 percent of collected data" (Igoe 2021) in the genomics and genetics. This is a major issue. We can infer from this fact that the majority of algorithms used in health-care are inherently biased on race. The future of utilization of AI in the provision of healthcare depends on creating unbiased algorithms.

What can we do to prevent algorithmic bias in healthcare? As Joy Buolamwini mentioned in her

documentary, AI was founded by a group of Caucasian men at Dartmouth College in the 1950s. However, it is still true to this day that the teams creating these algorithms predominantly consist of Caucasian men. In order to prevent algorithmic bias, as Mattie discusses in Katherine's article, it is vital to "address bias at every step" (Panch, Mattie, Atun 2019) of the algorithm creation. Since we as humans will have underlying bias in our day-to-day life, each algorithm will have a small amount of bias. In order for us to create algorithms with as little bias as possible, we need to ensure that the teams working on the creation of the algorithm be as diverse as possible. (Igoe 2021) Along with this, it is important that we are collecting more data for marginalized groups. Specifically, collecting a more diverse data set to train the algorithms used in the provision of healthcare can positively impact the healthcare community as a whole. For example, when we are able to create an algorithm with as little bias as possible it will greatly impact the provision of healthcare in third world countries who may not have access to the data to create these algorithms.

5 Impact in Africa

In this section, we will discuss Ayomide Owoyemi, Joshua Owoyemi, Adenekan Osiyemi and Andy Boyd's article, "Artificial Intelligence for Healthcare in Africa." Based on their article, Africa does not have enough digital medical records in order to create an efficient unbiased algorithm. (Owoyemi, Owoyemi, Osiyemi 2020) However, it is possible for Africa to utilize the algorithms made in other countries. The only issue is the data sets used in these algorithms more often than not do not contain people who are similar to Africans physiologically. (Owoyemi, Owoyemi, Osiyemi 2020) As we have mentioned before, these data sets may contain unintended bias which could result in incorrect out-comes when utilized in Africa.

In order to better the provision of healthcare in Africa by using artificial intelligence, we have to first locate where the issue stems from. The main issue for Africa is that there are not enough digital data for algorithms to be used effectively. As for the data used in medical artificial intelligence, each data point needs a valid result. This results in a high cost to obtain data for training medical models, since each data set must contain labeled data which can only be created by medical experts. (Owoyemi, Owoyemi, Osiyemi 2020) How can we help Africa create a more accurate provision of health with the help of artificial intelligence? If more developed countries focus on obtaining data that contains more people who are physiologically similar to the people of Africa, we can create a more accurate model for the people of Africa to use. Due to the limited access to artificial intelligence in Africa, it is imperative that more developed countries take on this task and help better the future for the people of Africa.

6 Eradicating AI Bias

6.1 Reducing bias in the future

In this section, we will discuss how we can reduce bias in artificial intelligence in the future. In Brian Uzzi's "A Simple Tactic That Could Help Reduce Bias in AI" he concludes that," it's easier to program bias out of a machine than out of a mind." (Uzzi 2020) In his article, Brian discusses how "blind taste tests" could help prevent bias from intruding into our algorithms. (Uzzi 2020) For example, the famous "Pepsi Challenge" from the mid-1970s when people tried Coca-Cola and Pepsi 'blind' – no labels on the cans – the majority preferred Pepsi over its better-selling rival. (Uzzi 2020).

The idea is that we would repetitively train the data set with all of the data, and then train a data set and remove one variable and then compare the results. (Uzzi 2020) If the results from the first data set are different than the second data set, then we can conclude that either the missing variable is important to include in the training data set, or by including this variable, has directly created bias in the resulting model. (Uzzi 2020) Implementing this ideology into the training methods of AI would create a thorough way to ensure that human bias is eliminated from our algorithms. This could be especially helpful for the provision

Proceedings DOI: 10.21467/proceedings.157
ISBN: 978-81-970666-0-3

Series: AIJR Proceedings
ISSN: 2582-3922

of health- care as it would be a process that could counter the risks of bias in algorithmic development.

6.2 Challenges

Just as human bias can be observed in many functions of society, the same bias will continue to impact the AI models enlisted to aid these functions. As a result, the current trajectory of practical AI implementation will have detrimental effects on minority groups and people living in data sparse regions in regard to critical components of life (healthcare, education, employment, housing, etc). It can be said that with any modern algorithm, the bias put in equates to the bias produced. Therefore, it is increasingly difficult to imagine unbiased data without unbiased humans overseeing the creation of data sets. While this ideal implies a societal shift on a much larger scale than algorithm development, it is paramount that researchers focus their efforts on eradicating racial bias in all future AI technology. The global impact of the COVID-19 pandemic and the AI used to help com- bat the virus has brought the issue of racial bias in AI to the forefront of the AI community. Several efforts have already been made to eradicate AI bias and push towards a future in which everyone can benefit from AI technology, including those in data sparse regions.

6.3 Successes

Since the COVID-19 pandemic has spread across the globe, the dependency on AI has also increased in the last two years. Michael Horikawa analyzed this trend of AI usage during the pandemic in his paper aimed at revealing the connection between the pandemic and AI bias. As a result of the increased usage of AI during the pandemic, Congress has started looking into legislation encompassing racial bias in AI. Indeed, in May of 2020, several members of Congress raised their concerns to House and Senate leadership surrounding evidence of racial bias in AI uncovered by the U.S. pandemic response. The letter postulates that the several functions of AI technology in regard to the pandemic, including monitoring the spread of the virus and allocating health care resources, "is literally a matter of life and death" (Clarke *et al.* 2020). Although the healthcare industry has observed evidence of racial bias in its AI models in the past, the pandemic has brought the issue to the forefront given the potentially fatal consequences brought upon by the virus.

The members of Congress suggested as part of their call to action that any health care provider being federally funded as well as any business earning over a specific revenue threshold must obtain written approval that their AI algorithm has been bias tested before being put into use. This is truly a significant development in terms of eradicating racial bias in AI in the United States brought upon by the pan- demic, but it does not directly address AI development in data sparse regions. The path towards truly removing bias from AI technology requires not only for the leaders of the industry to diversify their own AI endeavors, but also to push towards involving less developed areas in the AI community. One such endeavor to address the latter criteria can be seen from a 2018 initiative by Google. The company opened the first AI research lab on the African continent, in Accra, Ghana (Cisse 2018). Not only does this effort by Google spark interest in AI research among African communities, but it is also a start for giving them a means to advance AI involvement across the continent. As previously mentioned, eradicating racial bias in AI in the United States does not address areas of the world such as Africa, and increased representation of data sparse regions will prevent future models from unintentionally exhibiting bias against them due to a lack of data.

7 Research Novelty

With the COVID-19 pandemic just emerging in 2020, research on the virus itself has been confined to the last two years. Within that time frame, there is not an abundance of publications that have considered the

connection between the pandemic and efforts to eradicate racially biased AI. Those that have, such as the works referenced in this analysis, lack a holistic viewpoint towards the issue of biased AI. In most cases, research on the topic focuses on developed areas in the AI community without mention of a procedure for improving representation as well as racial bias of AI in data sparse regions like Africa. With the findings presented in this publication, the call to action in the future work is founded on the hope that researchers will begin to prioritize data sparse regions while continuing to solve the overarching issue of racially biased AI.

8 Conclusion/Future Work

In conclusion, it is evident that bias in artificial intelligence negatively impacts the provision of health to non- Caucasians. However, over the past few years this bias has become more recognized in the artificial intelligence com- munity and hence more effort is being put into solving the problem of bias. Given the limited resources in third world countries, the success of eradicating this bias hinges upon the willingness of developed countries to assist by obtaining more diverse data to further the advancement of artificial intelligence for the provision of healthcare.

The current trajectory of AI technology is one in which almost all aspects of future society will somehow be aided by AI. As a result, eliminating racial bias in AI algorithms is not a matter of choice, but rather a necessary action to take in order to ensure that the future of AI does not damage those that it is intended to help. The efforts currently being taken to eradicate racial bias in AI are important stepping stones to achieve this reality, but much more work must be done. While eliminating the inherent bias that humans have requires significant societal changes on a global scale, there are several other areas for future work by AI researchers on this issue. At the most fundamental level, there should be a more unified system of regulations surrounding the criteria for AI algorithms in countries that utilize the technology. Just as the racially biased AI being used in response to the COVID-19 pandemic prompted members of Congress to bring proposed regulations such as bias checking of algorithms into legislation, similar regulations should be considered worldwide. Unified bias tests would provide a non- negotiable form of evaluation of an AI model for any potential racial bias that could detrimentally impact groups of people subject to the technology before it is permitted for public use. By having algorithmic bias tests in place, one key component of eradicating racial bias in AI would be achieved.

In addition to establishing a coordinated bias test system to catch any potential bias in AI tools before it is too late, there should also be an emphasis on improving the early stages of algorithm development. This would ensure that the potential for biased models is reduced from the very start. This includes propositions such as diversifying key leadership positions in AI development along with pushing for more representation in data sparse regions. Morgan Livingston, a member of the University of California Berkeley, proposed a federal diversity initiative for AI technologists in her 2020 entry in the Journal of Science Policy & Governance. Her initiative involves a baseline standard for inclusive hiring in AI development teams. As a result, the likelihood of racial bias in AI technology would be greatly reduced from the onset of a project due to a diverse team of engineers training AI models in diverse ways (Living- stone, 2020). By eliminating unintentional bias in development teams and data sets, AI tools have a far greater potential to be bias free before even reaching the bias tests.

Finally, data sets used to train AI algorithms should include more of an emphasis on data sparse regions such as Africa. Even if racial bias is effectively removed from existing AI, the models will still unintentionally discriminate against areas of the world where data for the algorithm is not as abundant and therefore cannot effectively classify the people living there. Researchers should aim to provide these areas with more resources for AI research, such as Google has attempted in Ghana. In addition, data sets used

Proceedings DOI: 10.21467/proceedings.157
ISBN: 978-81-970666-0-3

Series: AIJR Proceedings
ISSN: 2582-3922

to train AI should be required to utilize a specific number of entries from regions that can be defined as data sparse. As a result, AI models will become more effective at classifying these regions as time goes on and exhibit less bias. With a strong emphasis on combating racial bias in both existing AI models from leaders in the industry as well as in less developed data sparse regions, the future of AI technology has the potential to change its current trajectory towards unbiased tools that aid all of its users equally.

9 Publisher's Note

AIJR remains neutral with regard to jurisdictional claims in institutional affiliations.

How to Cite

Galvin & Hogan (2024). Bias in AI and Machine Learning: The Impact of COVID-19 in African Healthcare Communities. *AIJR Proceedings*, 67-73. https://doi.org/10.21467/proceedings.157.10

References

[1] Owoyemi A, Owoyemi J, Osiyemi A and Boyd A. 2020. Artificial Intelligence for Healthcare in Africa *Front. Digit. Health* 2:6.

[2] DeCamp M., and Charlotta Lindvall. 2020. Latent bias and the implementation of artificial intelligence in medicine *Journal of the American Medical Informatics Association: JAMIA* 27:12.

[3] Igoe, K. A. 2021. Algorithmic Bias in Health Care Exacer-bates Social Inequities — How to Prevent It *The Harvard T.H. Chan School of Public Health: Health Literacy Studies Web Site.*

[4] Panch, T., Mattie, H. and Atun, R. 2019. Artificial intelligence and algorithmic bias: implications for health systems *Journal of Global Health* 2019:9

[5] Buolamwini,J. 2020.Coded Bias *Netflix* https://www.netflix.com/watch/81328723?trackId=14277281

[6] Uzzi, B. 2020. A Simple Tactic That Could Help Reduce Bias in AI *Harvard Business Review*

[7] Cisse, M. 2018. Look to Africa to advance artificial intelligence. *nature* 562, 461.

[8] Clarke, Y., Wyden, R., Markey, E., Beyer, D., Cardenas, T., Carson, A., Garcia, J., Lee, S., Jayapal, P., Johnson, H., Lieu, T., Moulton, S.Takano, M. 2020. *Congress of the United States*.

[9] Jercich, K. 2020. AI bias may worsen COVID-19 health dis- parities for people of color. *Healthcare IT News*.

[10] Leslie, D., Mazumder, A., Peppin, A., Wolters, M K., Hagerty, A. 2021. Does "AI" stand for augmenting inequality in the era of covid-19 healthcare? *BMJ* 2021;372:n304.

[11] Livingstone, M. 2020. Preventing Racial Bias in Federal AI. Journal of Science Policy & Governance 16-2.

[12] Shin, T. 2020. Real-life Examples of Discriminating Artificial Intelligence. *Medium.*