Algorithm Bias Influence and Negative Consequences: A Complex Dive into the Nature of Algorithms

Mary Margaret Stephenson
Blaez Jibben
Alan Ray
Brynn Lavender

September 7th->9th, 2024
Boone, North Carolina

Introduction Outline

Background Information:
The information interchange of the world today is driven by the internet, whether that be through social media posts, a blog site, or an online scholarly library. In order to transfer and spread information, algorithms are commonly used to organize web content into worthwhile information. If a company does not organize this information, their site could be taken over by misinformation and fraudsters [3]. This is why algorithms are useful- but inevitably, "search engines cannot passively and neutrally distribute third party content" in an unbiased manner. These algorithms can generate biases from incomplete data, or flawed information, emphasizing the human biases that may show up already [4]. These biases can result in disastrous consequences, including job bias, gender biases, and advertisement biases [4].

Background Information Notes:
One of the earliest examples of algorithmic bias in computer programming is described by S. Lowry and G. Macpherson [1]. In 1979, a computer program written for St George's Hospital Medical School began replicating programmer biases by reducing the chances of female and minority applicants.

Thesis:
Algorithm bias extends human biases and may extend these biases onto people through gender discrimination, promoting inaccurate information, and polarizing groups around the US. Steps including code transparency, open source alternatives, and an incentive to increase in general knowledge of risks and benefits of these algorithms could help mediate the bias that these algorithms hold.

Literature Review Outline:
Summary of Existing Research:
Algorithmic bias exists, and is capable of real-world harm. These are two conclusions overwhelmingly supported by existing research. When computer programmers began writing code to automate the sorting and selection of people, their personal biases found ways into that code [1]. In some cases, the algorithmic sorting of social media profiles merely reinforced pre-existing categorization [5]. The harm done by algorithmic bias can be as minor as less-visible social media posts [5], or as serious as rejection from medical school [1] or a longer prison sentence [2].

Identification of Gaps:
Research into algorithmic bias is limited by what researchers have access to. They, like us, are unlikely to have access to the code of major social media networks, search engines, or LLMs. Studies on this topic often feel like research into a black box, where algorithms must be guessed at by inputs and outputs.

Significance to Argument:
Evidence of algorithms adopting and worsening human biases can be found in the literature as far back as 1979 [1], and as recently as 2019 [2] - to the detriment of minority groups.
These algorithms are not just programmed with human biases - they feed off the user's pre-existing biases. Calice [7] details the way algorithms feed into existing societal divisions.
Lastly, we make the argument that a lack of knowledge with regard to algorithms allows people to be misled by viral misinformation. Goldman [3] explores the ways in which search engine results are purposefully skewed to the detriment of users. Cotter and Reisdorf name the knowledge gap between

more- and less-experienced users "a new dimension of (digital) inequality" [4]. Zhang and Chen, however, describe a less definitive study into algorithmic biases with regards to gender on Weibo [5].

Argument Development Outline
Argument: Algorithms can worsen human biases, especially against women and minorities; often, the usage of an algorithm is more detrimental to these groups than if there was no algorithm used at all.
Evidence: A computer program was created to streamline who was offered an interview at St George's Hospital Medical School. After it was implemented, it was discovered that the program "unfairly discriminated against women and people with non-European sounding names" [1]. Even something as innocent as one's name can be used by an algorithm to exclude them. This article went on to say that because of the program, fewer women and minorities were offered interviews for medical school [1]. This is only one example of algorithms showing bias against women and minorities. There are likely hundreds of other examples which have not been investigated.
Counterargument: There are many studied steps that programmers can take to lessen or eliminate these biases. It is easy to use these steps to dismiss the reality of gender and race discrimination altogether. One of the most significant proposals is to use more varied training data for algorithms [6]. When more training data is used, algorithms can sort more accurately. Therefore, it is vital to determine what training data can be used.

Argument: Algorithms have a tendency to feed into a user's personal bias or beliefs, especially when search engines personalize results based on past searches and data.
Evidence: One study found that search engines can omit critical information and rank information in a biased way [3]. These practices have only become more popular as time passes and companies compete for top spots on search engines. The issue grows as usage continues, and a feedback loop is created. This could be seen as a good practice with more accurate search results for users and more traffic to frequently searched websites. However, these large companies are given the perfect opportunity to shut out small businesses by taking the top spots in search results and leaving small companies lower on the list and less likely to be found.
Counterargument: While there is evidence of this happening, American Democrats and Republicans often exaggerate the issue for their own benefit. One study found that Democrats and Republicans viewed algorithm bias differently when the data was from their party instead of the other one [7]. This means that the media often exaggerates the impact of search results on different political parties or beliefs when, in reality, many algorithms have been created to counteract false information.

Argument: A lack of algorithm knowledge increases the rate at which misinformation is spread, and a greater understanding of these algorithms would allow people to critique what they read online to a higher extent.
Evidence: "Without knowledge of algorithms, individuals lack important context that could be mobilized in assessing the merit of information they encounter on platforms" [4]. This is a direct quote from the authors, where they argue that lack of knowledge of algorithms decreases ability to access such information. The authors later conducted a study, which suggests "that an algorithmic knowledge gap exists within the context of online search" while also finding that the there is a gap in knowledge

based on socio-economic status (SES), with the study finding that "higher SES populations exhibit greater knowledge about how algorithms work than lower SES populations" [4]. While we aren't arguing about SES, it's an important assessment which analyzes a knowledge gap between individuals, which could become greater as time continues.

Counterargument: Zhang and Chang argued that while algorithm knowledge in the mens model held up, it "fails to stop women from being discriminated against. Algorithm knowledge… was found to influence bias in our mens model only" [5]. This may suggest some outside factors that could have a play in algorithm bias overall, and how people are affected by this bias, but overall still suggests that some algorithm knowledge will increase assessment of info online.

Recommendations or Solutions Outline:
Proposed Solutions:
Argument #1: Require Discrimination Training and Team Collaboration for Selection Algorithms
- In regards to the St. George's Medical School selection algorithm, one of the big issues was that it was created by one person and was not extensively peer-reviewed. This person also had no discrimination training prior to his work on the algorithm. In order to avoid biased selection algorithms in the future, it would be best practice to involve multiple persons on the project, as well as to thoroughly educate them on what discrimination looks like and the risks involved. [1]

Argument #2: Make changes to Practices for Algorithms
- In terms of algorithms such as search engines, increasing exposure of marginalized websites could be an effective method to decrease bias in these sites. A 'random rank promotion' could increase exposure to such sites [3].

Argument #3: Disclose Information About Biases with the Public
- If users are made aware of the biases that exist within many of the algorithms that operate today, they can better select software products they want to use and it could help them to gauge the amount of time and effort they need to put into understanding and organizing any results they receive. [3]

Justification:
Argument #1:
- When working with an already existing set of data, discriminatory factors may not be obviously apparent to those not properly trained. With this training, people that are ignorant to what discrimination might look like in a data set will be able to more easily point it out and adjust their algorithms accordingly. Also, as is the case with most projects, it is better to have more than one person handling the creation of the selection algorithm. What one person may not notice at first glance, may be caught by another team member. A collective group of working minds is better than just one.

Argument #2:
- Changing algorithms to support randomized rank promotion may support more unbiased results from showing up. This does not only apply to only search engine algorithms, but it is easily applicable to such.

Argument #3:

- As stated in the argument, many individuals are completely unaware of the biases that exist within the computer algorithms they regularly use. If they were educated on, or at the very least made aware of, these biases, they could then make an informed decision on how to handle their results in the future.

Implications:

Argument #1:

- As stated in the argument, there are surely many other examples of biased selection (and other) algorithms that have not yet been caught. However, if more people can be made aware of this issue and trained properly for it, we can hopefully avoid producing harmful algorithms in the future. It is also safe to assume that the more awareness that can be brought to this issue, the more likely existing problematic algorithms will be caught.

Argument #2:

- Randomized and/or different ranking for algorithms could incentivise less bias, promote new material, but potentially support misleading information [3].

Argument #3:

- Many commonly used applications may have to adjust their algorithms to become more "personalized." Other applications may gain more relevance as users discover certain applications suit them better than the more commonly used ones. [3]

Conclusion:

Summary of Key Points:

In conclusion, algorithm bias is going to be inevitable, but with awareness and education we can do our best to keep it from affecting the results we gain from these algorithms. These results can be as harmless as search results from a search engine to as life-changing as prison sentencing and school admittance. Because of this broad range in influential factors, it is greatly important for computer users to understand that these biases exist and adjust accordingly. This is especially important for those that worry they are receiving false or misleading information from whatever application they are using. This proactive approach not only strengthens user trust in technology but also ensures fairer outcomes for all users.

Call to Action:

Education on algorithm bias is crucial for its avoidance. Being aware of it will allow users to make educated decisions on how to handle results and help developers properly analyze their data and remove discriminatory factors from within their algorithms. Those that employ algorithms that could have potentially biased results should offer transparency with users about it. Users deserve to understand how their applications really work, especially when those applications can have dramatic effects.

Closing Statement:

This white paper was important for our team to create because it allowed us to properly organize our research findings and practice plain language skills. It helped us to better condense our findings and organize it in a way that includes what is most necessary for our audience. When writing about this topic in the future, we can come back to this paper to find our notes in a well-written and easy-to-read format to help us.