International Relations and Diplomacy, Jan.-Feb. 2024, Vol. 12, No. 1, 40-50
doi: 10.17265/2328-2134/2024.01.005

Digital Disparities: How Artificial Intelligence Can Facilitate Anti-Black Racism in the U.S. Healthcare Sector

Anthony Victor Onwuegbuzia
California State University Dominguez Hills, Carson, the United States

This paper delves into the intricate interplay between artificial intelligence (AI) systems and the perpetuation of Anti-Black racism within the United States medical industry. Despite the promising potential of AI to enhance healthcare outcomes and reduce disparities, there is a growing concern that these technologies may inadvertently/advertently exacerbate existing racial inequalities. Focusing specifically on the experiences of Black patients, this research investigates how the following AI components: medical algorithms, machine learning, and natural learning processes are contributing to the unequal distribution of medical resources, diagnosis, and health care treatment of those classified as Black. Furthermore, this review employs a multidisciplinary approach, combining insights from computer science, medical ethics, and social justice theory to analyze the mechanisms through which AI systems may encode and reinforce racial biases. By dissecting the three primary components of AI, this paper aims to present a clear understanding of how these technologies work, how they intersect, and how they may inherently perpetuate harmful stereotypes resulting in negligent outcomes for Black patients. Furthermore, this paper explores the ethical implications of deploying AI in healthcare settings and calls for increased transparency, accountability, and diversity in the development and implementation of these technologies. Finally, it is important that I prefer the following paper with a clear and concise definition of what I refer to as Anti-Black racism throughout the text. Therefore, I assert the following: Anti-Black racism refers to prejudice, discrimination, or antagonism directed against individuals or communities of African descent based on their race. It involves the belief in the inherent superiority of one race over another and the systemic and institutional practices that perpetuate inequality and disadvantage for Black people. Furthermore, I proclaim that this form of racism can be manifested in various ways, such as unequal access to opportunities, resources, education, employment, and fair treatment within social, economic, and political systems. It is also pertinent to acknowledge that Anti-Black racism is deeply rooted in historical and societal structures throughout the U.S. borders and beyond, leading to systemic disadvantages and disparities that impact the well-being and life chances of Black individuals and communities. Addressing Anti-Black racism involves recognizing and challenging both individual attitudes and systemic structures that contribute to discrimination and inequality. Efforts to combat Anti-Black racism include promoting awareness, education, advocacy for policy changes, and fostering a culture of inclusivity and equality.

Keywords: Bias in algorithms, Racial disparities in U.S. healthcare, Discriminatory healthcare practices, Black patient outcomes, Automated decision-making and racism, Machine Learning, Natural language processing

---

Anthony Victor Onwuegbuzia, M.A. Africana Studies Instructor, The Mervyn Dymally African American Political and Economic Institute, California State University Dominguez Hills, Carson, California, the United States.

Introduction

In exploring the annals of the American healthcare system, it becomes painfully evident that the pages are stained with a history of systemic injustice, specifically directed towards Black patients. The roots of anti-Black racism run deep, intertwining with the very fabric of healthcare institutions. As we delve into this narrative, it is crucial to confront the uncomfortable truths that have perpetuated harms and disparities in health outcomes and access for generations. Historical records of these uncomfortable truths go as far back as the early 18th century. During this time there was the emergence of a colonial medical complex beginning in the 1760s, which was facilitated by a robust medical experimental culture particularly in the British and French West Indies. One notable experiment was the smallpox inoculation of a population of 850 slaves conducted by a British doctor, John Quier, a plantation physician in rural Jamaica. This inoculation of slaves grew into an industry of sort, as many owners of slaves would begin to inoculate those on their plantations. By the 20th century, Blacks in the U.S. were still commonly subjected to racialized medical practices. For example, the U.S. Public Health Services in conjunction with the Centers for Disease Control and Prevention (CDC) infected over 400 Black men with syphilis over a span of 40 years. This became known as the infamous "Tuskegee Experiment", where medical researchers from Tuskegee University (back then Tuskegee Institute) were allowed to withhold information and to intentionally bypass the consent of the Black male patients whom they were intentionally infecting with syphilis. To add insult to injury, these so called "medical practitioners" intentionally withheld treatment that would have cured these cases and opted to facilitate their deaths. The 20th century also saw the emergence of the Eugenics Movement and racist beliefs led to the involuntary sterilization of Black women in the United States. For example, in North Carolina, data collected from 1937 to 1966 revealed that black women were the most targeted group to be forcibly sterilized. This was largely fueled by racism and the stereotype of black women being "unfit" mothers fueled the public's desire to ensure this population cannot further reproduce. Such two examples are simply meant to highlight the ills that both genders of Black society have historically endured in sanctioned medical practices.

Fast forward to the digitization era of medicine, racialized healthcare continues to plague Black patients, resulting in various disparities in treatment, diagnosis, and even access to medical care. However, the digitization of the medical industry has created new dynamics for which Anti-Black racism is practiced in medicine. The most recent advancements have been the role of Artificial Intelligence in healthcare, which undoubtedly became very prevalent during the global COVID-19 pandemic. Today, studies are still examining various medical outcomes of these AI systems, however, there have been a multitude of scholars that have published research revealing discriminatory practices that AI has intentionally/unintentionally taken on. One recent notable study includes the Kidney Algorithm findings, whereby researchers at Penn Medicine, Massachusetts General and Brigham and Women's found that after examining the health records of 57,000 patients with chronic kidney disease, one out of three Black patients (roughly over 700 black patients) were not placed into a more severe category due to an medical algorithm design that assigned Black people healthier kidney scores based on an inaccurate and racialized metric. As a result, these patients were automatically deemed ineligible to qualify for kidney transplants that were much needed. Instead, the algorithm metric overwhelmingly favored their white counterparts to qualify for this important procedure (Mendu et al., 2020).

Nevertheless, in the era of technological advancement, Artificial Intelligence (AI) has emerged as a powerful tool with the potential to revolutionize various industries, including healthcare. While the promises of AI in medicine are vast and transformative, it is crucial to critically examine the potential societal implications it may pose, and to further understand how AI applications in healthcare might disproportionately negatively impact Black patients. Nevertheless, I make the case that to accurately examine this impact, there needs to be a fundamental understanding of how this technology works, its development, and its continued evolution. Therefore, using a multitude of scholarly data, analytical reports, and news publications, this paper seeks to synthesize the three primary components of AI, which includes: algorithms, machine learning, and natural language processing. I suggest that it is incumbent upon us Black academics and researchers in the social sciences, to continue to foster a deeper understanding of the challenges posed by AI in perpetuating anti blackness. Only this will allow for the continued efforts in unravelling these complex layers of inherent, racialized biases that have been embedded in these AI systems.

Evolution of AI in the Medical Industry

The emergence of artificial intelligence (AI) in the United States healthcare industry has been a significant development that has revolutionized various aspects of healthcare delivery. AI technologies have proven the capacity to enhance efficiency, accuracy, and effectiveness in areas such as diagnosis, treatment planning, patient monitoring, and administrative tasks. According to the National Institute of Health, in the year 2020, approximately 86% of United States health care providers utilize at least one form of artificial intelligence within their medical practices (Kamensky, 2020). Various literature suggests that the development of AI grew simultaneously with the advancement of computers dating back to the 1940's and 1950's. However, it was in the late 1950's when the American medical industry began to take notice, with prominent medical practitioners such as Dr. Keeve Broadman predicting that "the making of correct diagnostic interpretations of symptoms can be a process in all aspects logical and so completely defined that it can be carried out by a machine." (Haug & Drazen, 2023, p.1201). Fast forward 11 years later, famed medical practitioner, Dr. William B. Schwartz published an article suggesting that "Computing science will probably exert its major effects by augmenting and, in some cases, largely replacing the intellectual functions of the physician." (Schwartz, 1970, p. 321). He further predicted that by the year 2000, computers would have an entirely new role in medicine, acting as a powerful extension of the physician's intellect (Haug & Drazen, 2023).

Today, one could make the case that Artificial Intelligence role in healthcare is largely predicated on what Swartz described years ago as the "physician's intellect". Nevertheless, this dynamic of intellect may bring into question how AI facilitates healthcare treatment and services among racial demographics. Many research institutions began to examine this question especially during the height of the COVID-19 pandemic, whereby telehealth and virtual medical services significantly increased. For example, in 2021 The National Institute for Health Care Management (NIHCM) Foundation published a study on racial bias in health care artificial intelligence. This study particularly revealed multiple disparities between black patients and white patients. Notably, the authors cited that Artificial intelligence (AI) tools such as algorithms are increasingly being used to determine who gets health care. Furthermore, the study indicated that AI tools can unintentionally increase the impact of existing racial biases in medicine through the explicit use of race to predict outcomes and risk. The authors also noted that using race as a factor has become a common practice when designing clinical algorithms (NICHM Foundation, 2021). Two years later researchers from School of Medicine, Stanford University, published a report that found AI powered ChatGPT and Google's Bard answer medical questions with racist, debunked theories that harm Black patients. The report found that all four models tested in the report—ChatGPT and the more advanced GPT-4, both from OpenAI; Google's Bard, and Anthropic's Claude—failed when asked to respond to medical questions about kidney function, lung capacity, and skin thickness. In some cases, they appeared to reinforce long-held false beliefs about biological differences between black and white people (Omiye, Lester, Spichak, Rotemberg, & Daneshjou, 2023). Despite evidence that race is not a reliable proxy for genetic differences, how to allocate clinical resources or treatment adherence, using race as a factor has become a common practice when designing clinical algorithms. Therefore, to clearly understand how AI may facilitate anti-blackness in health care, one must discern the following general AI components: algorithm, machine learning, and natural language processing. These digital frameworks are of great importance as they all play active roles in how AI operates.

What Exactly Is an Algorithm?

An algorithm is a step-by-step set of instructions or rules designed to perform a specific task or solve a particular problem. Algorithms are used in various fields, including computer science, mathematics, and everyday problem-solving. In computer science, algorithms are particularly important for writing computer programs. They provide a clear, unambiguous description of how to carry out a task or solve a problem, serving as a blueprint for the execution of a program. Algorithms can be expressed in natural language, flowcharts, pseudocode, or through a programming language.

The complexity of the term "Algorithm" dates to the Islamic Golden age, whereby many historians credit the famed mathematician Mohamed Al-Khuwariz as the architect of mathematical algebra. In their article titled "The Origin of the Algorithm", Baraka and Joseph suggest that it was in the 12th century when Khuwariz's algebra and other mathematical texts were translated into Latin, Spanish, and German and circulated throughout medieval Europe (Baraka, Salem, & Joseph, 1998). Among these mathematical teachings was the problem-solving tool "algorithm". However, after his death, the Germans modified the name of Al-Khawarizmus into "Algorismus" as the Latin/French equivalent, which soon became a derivative of the word known today as "algorithm" (Baraka et al., 1998). Fast forward centuries later, the study of algorithms became a focal point in the sciences from the 1960's-1980, as computer science was trying to establish itself as an independent academic discipline. While there's still no universally accepted definition of algorithm, a common one comes from a 1971 textbook written by computer scientist Harold Stone, who states: "An algorithm is a set of rules that precisely define a sequence of operations." (Stone, 1971, p. 321). To put this into simplistic terms, an algorithm functions as a set of instructions that a computer executes to learn from data. However, Gary (2008) suggests that coming up with the right answer at the end of a program is only the minimum requirement, as the best algorithms run fast, are sparing in their use of memory and other resources, and are easy to understand and modify. In his research on algorithms, Kitchin (2016) notes that algorithms can be conceived in several ways: technically, computationally, mathematically, culturally, economically, contextually, materially, philosophically, and ethically. This indicates that algorithms play a vital role in most activities taking place online. Furthermore, this suggests that algorithms heavily dictate what online users view on the internet, how interactions with others online users are facilitated, and how companies and organizations promote messaging to users. Algorithms provide a personalized, digital construct that tailors specific content to the individual user. For example, when a user makes an online purchase of an item, he or she will begin to see a multitude of advertisements (during internet surfing, social media scrolling, and gaming) that are directly related to their recent purchase. The same function applies to users who view video content on various online platforms. For example, on YouTube, when a user views a video about classic cars, the platform's algorithm will automatically begin to predict similar videos that will be recommended for future viewing. However, algorithm elements extend far beyond social media, online shopping, and gamming components of the internet, and have become a big factor in the healthcare sector as well.

Medical Algorithms

A medical algorithm is a set of step-by-step instructions or rules designed to assist healthcare professionals in making clinical decisions and solving problems. These algorithms are often based on evidence-based medicine, clinical guidelines, and expert consensus to provide a systematic approach to patient care. Such algorithms have existed even prior to the digitization of healthcare treatments and services. Initially, these algorithms were developed for the care of patients with acute minor illnesses, chronic disease, acute medical emergencies, and minor surgical problems, as well as to aid pediatric telephone triage, and to help detect disease in the worksite (Komaroff, 1982). Furthermore, medical algorithms were used to aid pharmacists to help physicians track patients returning for refills. In an early 1980's article titled "Algorithms and the Art of Medicine", famed physician Komaroff (1982) declared that over the past 30 years there have been increasing attempts to transform the art of medical decision making into a science, to supplement a spontaneous, informal, and implicit set of judgements with the conclusions of a predetermined, formal, and explicit scheme of logic. One could make the case that Komaroff's foresight seemed to manifest in the following decade as computerization technologies would become more prevalent.

In the early 90's healthcare providers began incorporating computer technologies within their practices. The Institute of Medicine (IOM) highlights this transition in their 1997 committee report. The report's focal point centered on improving the patient record in response to the increasing functional requirements and technological advances. In this case, the primary advancement was known as computer-based patient record or (CPR). The report indicated that CPR would play an increasingly important role in supplying data for computer-based population databases. It further explained that high-quality data will become essential to the management of care for individuals; and that such data are equally critical for research, to support public health activities, and to track the performance of health care providers—both individuals and institutions (Dick, Steen, & Detmer, 1997). Nevertheless, it is important to conceptualize how these algorithms are developed. For example, an algorithm for screening psychiatric patients for physical disease was empirically derived from a comprehensive assessment of 509 patients in California's mental health system. The first 343 patients were used to develop the algorithm, and the remaining 166 were used as a test group. Calculations were made for several versions of the algorithm, and the data were compared with the diagnoses listed in the patients' admission mental health record. The algorithmic procedure was more accurate and more cost-effective than the medical evaluation procedures used by the state mental health system. When applied to the test group, the algorithm detected up to 90 percent of patients who had an active, important physical disease at a cost of $156 per patient. The mental health system had detected 58 percent of test-group patients with a disease at a cost of $230 per patient (Jr Sox, Koran, Sox, Marton, & Dugger, 2006). Undoubtedly, this study yielded high accuracy in screening psychiatric patients for physical disease. However, the participation process itself exposes potential ethnic and racial biases that are inherently established through the racial makeup of such trial subjects. For example, various HIPPA laws and other medical research policies allow for test subjects to remain culturally and ethnically unidentified. Based on these premises, one could easily see how this enables pathways for medical algorithms to develop in a manner that may be potentially inherent racial biases that are favorable to one group over another.

Today, a large portion of the modern medical data is expressed as images or other types of digital signals, such as X-Rays, MRI, computer tomography (CT), positron emission tomography, single-photon emission computed tomography, electrical impedance tomography, and ultrasound. The acquisition of huge amounts of such sophisticated image data has given rise to the development of automatic processing and analysis of medical images. Digital image processing deals with the manipulation and analysis of images that are generated by discretizing the continuous signals. Segmentation of structures from 2-D and 3-D images is an important step for medical data analysis that can help in visualization, automatic feature detection, image-guided surgery, and for registration of different images. Such methods as digital image processing when combined with others like machine learning, fuzzy logic, and pattern recognition are so valuable in image techniques and can be grouped under a general framework (Fenster & Chiu, 2006). However, physical traits or phenotypes that are disproportionately present in subgroup populations (such as racial differences in breast or bone density)—can also introduce bias. To put this in historical context, there was a dramatic rise in medical imaging spanning from 1997 through 2006 whereby 377,048 patients in the U.S. underwent 4.9 million diagnostic imaging tests. The National Library of Medicine cited that diagnostic imaging essentially doubled during this time span, suggesting a growth from 260 to 478 examinations per thousand enrollees per year. Nevertheless, it is crucial to highlight that amidst this upswing, studies have indicated that AI models possess the capability to predict patient demographics, such as race, directly from medical images. This may occur even in the absence of discernible anatomical or physiological features recognized by human clinicians. Such was recently highlighted by Stanford University's James Zou, who took a deep dive into the hidden racial variables of algorithms associated with AI generated predictions. The study notes how medical images could improve or exacerbate health care disparities based on their algorithmic inputs along with machine learning techniques (Zou, Ho, & Obermeyer, 2023).

Machine Learning

Machine learning is a subfield of artificial intelligence (AI) that focuses on the development of algorithms and models that enable computers to learn from data and make predictions or decisions without being explicitly programmed (Brown, 2021). The objective of machine learning is to enable computers to automatically learn and improve from experience. For example, machine learning is behind chatbots and predictive text, and language translation apps. It is also behind the ability whereby the popular streaming platform Netflix suggests shows to an individual based on their selection history, and how one's social media feeds are presented to them in a customized, and targeted manner. However, in the U.S. healthcare sector, with a growing industry of smart watches, fit bits, and devices that constantly gather a plethora of health data, the prevalence of using machine learning to analyze this data has continued to gain momentum. Notably, in 2017, machine learning in healthcare was the leading topic at the Annual Computer Software and Applications Conference, whereabout software designers suggested that artificial intelligence in machine learning will prove to be the solution for both reducing the rising cost of healthcare and helping establish a better patient-doctor relationship. Big data solutions can be used for a plethora of health-related uses; some include helping doctors determine more personalized prescriptions and treatments for patients and helping patients determine when and if they should schedule (Bhardwaj, Nambiar, & Dutta, 2017). Habehh and Gohe (2021) suggests that current machine learning advancements in healthcare have primarily served as a supportive role in a physician or analyst's ability to fulfill their roles, identify healthcare trends, and develop disease prediction models. They further indicate that in large medical organizations, machine learning-based approaches have also been implemented to achieve increased efficiency in the organization of electronic health records identification of irregularities in the blood samples, organs, and bones using medical imaging and monitoring, as well as in robot-assisted surgeries (Habehh & Gohe, 2021). There are a multitude of machine learning methods, however, four primary techniques include: supervised, unsupervised, semi-supervised, and reinforcement learning.

Supervised Learning

Supervised learning utilizes a data set which includes both input features as well as the output class or target which are labeled at the start of training. In short, the machines are learning off the "labeled data" for the purpose of predicting. It is important to note that supervised learning is the most important methodology in machine learning, and it also has a central importance in the processing of multimedia data (Delany, Cord, & Cunningham, 2008). For example, let's say we want to teach a computer to recognize pictures of flowers. We can show it pictures of different breeds of flowers and tell it the name of each breed. In this case the label of the flower breed would be inputted and matched with the output data. In the context of machine learning in healthcare, this process is identical, therefore in the above example one could substitute pictures of breeds of flowers with those of pictures of healthy hearts vs. those plagued with heart disease. Another example would be the input of x-rays, CT scans, and genetic differences with predisposed labels attached.

Unsupervised Learning

Unsupervised machine learning uses algorithms to analyze and cluster unlabeled datasets. These algorithms discover hidden patterns or data groupings without the need for human supervision, guidance, or instructions. It is important to note that unsupervised learning algorithms do not take any feedback for the prediction. This learning finds the hidden patterns in data, two simple concepts i.e. principal component analysis (PCA) and cluster analysis are used in unsupervised ML (Chauhan, Rawat, Malik, & Singh, 2021). In the healthcare sector, this method enables the discovery of similarities and differences of various information (e.g. image recognitions, data analysis, patient segmentation, etc.) (IBM, n.d.). For example, computerized tomography scans in outcome unsupervised data sets may be comprised from random sample scans. Nevertheless, this unsupervised component does allow for algorithms to systematically associate, group, and label these samples based on similarities and opposites.

Semi-supervised Learning

Semi-supervised learning is a broad category of machine learning techniques that utilizes both labeled and unlabeled data. According to a 2023 review in the journal publication *Econometrics and Statistics*, there has been increasing attention to semi-supervised learning (SSL) approaches in machine learning to forming a classifier in situations where the training data for a classifier consist of a limited number of classified observations but a much larger number of unclassified observations (Ahfock & McLachlan, 2023). The authors suggest this is because the procurement of classified data can be quite costly due to high acquisition costs and subsequent financial, time, and ethical issues that can arise in attempts to provide the true class labels for the unclassified data that have been acquired. One of the simplest examples of semi-supervised learning is self-training. Self-training is the procedure in which you can take any supervised method for classification or regression and modify it to work in a semi-supervised manner, taking advantage of labeled and unlabeled data. This data can then be developed into distinct speech recognition models, image classifications, and internet content classifications.

Machine Learning and Anti-Black Racism

Simply put, ML systems in healthcare can project anti-blackness via the biased and discriminatory datasets they can be trained on. The result of this training then allows for the development of various medical algorithms. Even more problematic is that industry-standard medical algorithms have been developed and implemented from this ML process despite having been trained on inaccurate data. For example, last year, it came to light that six algorithms used on an estimated number of 60-100 million patients nationwide were prioritizing care coordination for white patients over black patients for the same level of illness (Nichols, 2020). The algorithm was revealed to be trained on costs in insurance claims data, predicting which patients would be expensive in the future based on who was expensive in the past. Historically, less is spent on black patients than white patients, so the algorithm ended up perpetuating existing bias in healthcare. To put this in context, only 18% of the patients identified by these algorithms as needing more care were black, compared to about 82% of white patients. However, if these algorithms were to reflect the true proportion of the sickest black and white patients, those figures should have been about 46% and 53%, respectively.

It is also important to note that the absence of black patient representation in medical datasets may prove equally harmful. A notable example of this was revealed in 2021 by the U.S. Preventive Services Task Force when they reported the agency was unable to make specific colorectal cancer screening guidelines for black patients despite this demographic having the highest incidence and mortality rates from colorectal cancer. The agency openly admitted this outcome was largely due to the absence of black people in clinical trials, which are often used as studies for innovative medicines (US Preventive Services Task Force, 2021). As ML continues to play a large role in the development of AI algorithms, the need to mitigate biased medical datasets, implement data set oversight agencies, and ensure equity in clinical trials, can be active components in combating anti-black dynamics in medical AI.

Natural Language Process

Natural Language Processing (NLP) in healthcare artificial intelligence (AI) involves the use of computational techniques to understand, interpret, and generate human language in the context of medical data. The abundance of healthcare data comes from various sources including electronic health records (EHRs), medical literature, patient-generated data (such as social media posts or wearable device data), and medical imaging reports. The objective of these NLP systems is to collect and aggregate this data for medical analysis. These systems can analyze patient notes and suggest potential diagnoses or treatment options based on the presented symptoms and medical history. However, when natural language processing (NLP) classifiers are developed with biased or imbalanced datasets, disparities across subgroups may be codified, perpetuated, and exacerbated if biases are not assessed, identified, mitigated, or eliminated.

It is important to note that healthcare physicians play a large role as to how language is utilized on behalf of patients when it comes to medical evaluations, recommendations, and overall analysis of treatment. For example, a recent study on the use of stigmatizing language by physicians highlighted ways in which physicians may often express negative feelings particularly toward black patients. To provide context, a total of 138 clinicians (attendings and residents) wrote 600 encounter notes with 507 patients. Most patients were identified in the medical record as female ($n = 350$ (69%)). Most patients were identified as black/African American ($n = 406$ (80%)), and 76 (15%) were identified as white. This finding indicated the use of stigmatizing language toward black patients fell into five following categories: (1) questioning patient credibility, (2) expressing disapproval of patient reasoning or self-care, (3) stereotyping by race or social class, (4) portraying the patient as difficult, and (5) emphasizing physician authority over the patient (Park, Saha, Chee, Taylor, & Beach, 2021). Understanding the ways in which bias might manifest in the language used in medical records, and developing interventions to eliminate biased language, could have a large impact on the reduction of disparities for black patients.

Other aspects of NLP systems have also revealed biases in people's namesakes, which are very important identifiers in healthcare. However, this aspect extends beyond the healthcare sector and has been explicitly shown throughout many facets of the internet. Many algorithms have developed racialized associations with sounding and text names, which may have different outcomes for different people. Think how one's name may be assigned a value within a system, and how that value may determine such things as employment opportunities, appointment availabilities, services offered, etc. Various studies have found that negative stigmas and associations in the form of language, images, and text have inherently been disproportionately ascribed to black people. One such study cited on by the Brookings Institute's website, revealed racial biases in widely used word "embeddings", trained (machine learning) on a corpus of 800 billion words collected from the web (data set), revealing that names of African Americans tend to co-occur with unpleasant words (Caliskan & Bryson, 2017). Findings also suggested that measuring the relative association of names of African Americans vs. names of white people with pleasant and unpleasant words shows that the word "embeddings" contains negative associations for the concept of an African American social group due to the overall biased depiction of this group on the internet. These types of associations that reflect negative attitudes toward one social group are considered harmful and prejudiced, which allows for the continued practice of anti-black racism in these digital spaces.

Conclusion

The utilization of artificial intelligence (AI) in healthcare presents a double-edged sword, demonstrating both promise and peril in addressing systemic issues such as Anti-Black Racism. While AI has the potential to streamline processes and improve healthcare delivery, its implementation must be approached with caution to avoid perpetuating existing biases and disparities. The evidence suggests that without deliberate efforts to mitigate algorithmic biases and ensure equitable access and treatment for all, AI systems can inadvertently exacerbate Anti-Black Racism in healthcare. Therefore, proactive measures, including diverse representation in AI development teams, transparent algorithms, and ongoing monitoring for bias, are essential to harnessing the full potential of AI while safeguarding against the reinforcement of discriminatory practices in healthcare.

With the lack of regulation and readily available bias auditing mechanisms, AI companies have not provided transparency in the everyday effects of the algorithms that they deploy in society, as these technologies for the most part are not standardized nor regulated. Furthermore, technology companies that develop cutting edge AI have become disproportionately powerful with the data they collect from billions of internet users, thus exercising full autonomy in how data are disseminated.

Recommendations

Combatting Anti-Black Racism in artificial intelligence (AI) in healthcare requires a multifaceted approach that addresses biases in data, algorithms, and decision-making processes. Here are some clear solutions:

- Establishing active and coordinated efforts put forth by black academics, STEM researchers, healthcare providers, social activists, to pressure and ensure that these technology companies are including a diverse representation in their AI developments.

- Policy and regulation: Black health care groups and associations need to align with legislators to actively advocate for policies and regulations that address Anti-Black Racism in AI in healthcare. This may include legislation requiring transparency and accountability in AI algorithms, as well as guidelines for the responsible use of AI in healthcare settings. However, these groups will constantly need to stay abreast on how AI is being introduced in their respective healthcare environments.
- Data collection and bias mitigation: Propose and implement a rigorous process for the collection and labeling of healthcare data to reduce racial biases. Advocate for a transparency mechanism that is publicly accessible. This will allow more patient stakeholders the power to ensure that datasets used to train AI models are representative of diverse populations and to detach negative and stigmatized associations facilitated by NLP systems.
- Community engagement and participation: Involve black communities and other marginalized groups in the design, development, and implementation of AI in healthcare. This ensures that AI systems are responsive to the needs and concerns of these communities, leading to more equitable outcomes. Furthermore, I would encourage these communities to work diligently with both academic and healthcare institutions, which will also ensure a level of transparency that builds trust between community members and these institutions.

References

Ahfock, D., & McLachlan, G. J. (2023). Semi-supervised learning of classifiers from a statistical perspective: A brief review. *Econometrics and Statistics, 26*, 124-138. doi:https://doi.org/10.1016/j.ecosta.2022.03.007

Baraka, A., Salem, R. M., & Joseph, N. J. (1998). The origin of the "algorithm". *Anesthesiology, 89,* 277. doi:https://doi.org/10.1097/00000542-199807000-00052

Bhardwaj, R., Nambiar, A. R., & Dutta, D. (2017). A study of machine learning in healthcare. In *2017 IEEE 41st annual computer software and applications conference (COMPSAC)* (pp. 236-241). Turnin, Italy: IEEE.

Brown, S. (2021, April 21). Machine learning, explained. Retrieved from https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

Caliskan, A., & Bryson, J. J. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science, 356*(6334), 183-186. doi: 10.1126/science.aal4230

Chauhan, T., Rawat, S., Malik, S., & Singh, P. (2021). Supervised and unsupervised machine learning based review on diabetes care. In *7th international conference on advanced computing and communication systems (ICACCS)* (pp. 581-585). India: Institute of Electrical and Electronics Engineers.

Delany, S. J., Cord, M., & Cunningham, P. (2008). Supervised learning. In *Machine learning techniques for multimedia* (pp. 21-49). Berlin: Springer Berlin.

Dick, R. S., Steen, E. E. B., & Detmer, D. E. (1997). *The computer-based patient record: An essential technology for health care*. Washington, D.C.: National Academy Press.

Fenster, A., & Chiu, B. (2006). Evaluation of segmentation algorithms for medical imaging. In *IEEE engineering in medicine and biology 27th annual conference*.

Fornell, D. (2023, February 6). FDA has now cleared more than 500 healthcare AI algorithms. Retrieved from https://healthexec.com/topics/artificial-intelligence/fda-has-now-cleared-more-500-healthcare-ai-algorithms

Gary, A. (2008). Back to basics: Algorithms; Computer scientists look for simplicity, speed and reliability. Sometimes they find elegance. *Computerworld, 42*(13), 30.

Gichoya, J. W., Banerjee, I., Bhimireddy, A. R., Burns, J. L., Celi, L. A., Cehn, L., … & Zhang, H. (2022). AI recognition of patient race in medical imaging: a modelling study. *Lancet Digit Health, 4*(6), 406-414. doi:10.1016/S2589-7500(22)00063-2

Habehh, H., & Gohe, S. (2021). Machine learning in healthcare. *Current Genomics, 22*(4), 291-300. doi:10.2174/1389202922666210705124359

Haug, C. J., & Drazen, J. M. (2023). Artificial intelligence and machine learning in clinical medicine, 2023. *The New England Journal of Medicine, 388*(13), 1201-1208. doi:https://doi.org/10.1056/NEJMra2302038

IBM. (2024, February 23). What is natural language processing (NLP)? Retrieved from https://www.ibm.com/topics/natural-language-processing

IBM. (n.d.). What is unsupervised learning? Retrieved January 3, 2024 from https://www.ibm.com/topics/unsupervised-learning

Jr Sox, H. C., Koran, L. M., Sox, C. H., Marton, K. I., & Dugger, F. (2006). A medical algorithm for detecting physical disease in psychiatric patients. *Hosp Community Psychiatry, 40*(12), 1270-1276. doi:https://doi.org/10.1176/ps.40.12.1270

Kamensky, S. (2020). Artificial intelligence and technology in health care: Overview and possible legal implications. *DePaul Journal of Health Care Law, 21*(3), 1-19.

Kitchin, R. (2016). Thinking critically about and researching algorithms. *Information Communication and Society, 20*(1), 1-16. doi:10.1080/1369118X.2016.1154087

Komaroff, A. L. (1982). Algorithms and the “art” of medicine. *American Journal of Public Health, 72*(1), 10-20. doi:10.2105/AJPH.72.1.10

Matheny, M. E., Whicher, D., & Israni, S. T. (2020). Artificial intelligence in health care: A report from the National Academy of Medicine. *JAMA, 6*, 509-510. doi:10.1001/jama.2019.21579

NICHM Foundation. (2021, September). Racial bias in health care artificial intelligence. Retrieved from https://nihcm.org/publications/artificial-intelligences-racial-bias-in-health-care

Omiye, J. A., Lester, J. C., Spichak, S., Rotemberg, V., & Daneshjou, R. (2023). Large language models propagate race-based medicine. *Nature partner Journals, 6*, 195. doi:https://doi.org/10.1038/s41746-023-00939

Park, J., Saha, S., Chee, B., Taylor, J., & Beach, M. C. (2021). Physician use of stigmatizing language in patient medical records. *JAMA Network Open, 4*(7), e2117052.

Rashidi, H. H., Tran, N., Albahra, S., & Dang, L. T. (2021). Machine learning in health care and laboratory medicine: General overview of supervised learning and Auto-ML. *International Journal of Laboratory Hematology, 43*(S1), 15-22. doi:https://doi.org/10.1111/ijlh.13537

Smith-Bindman, R., Miglioretti, D. L., & Larson, E. B. (2008). Rising use of diagnostic medical imaging in a large integrated health system. *Health Aff. (Millwood), 27*(6), 1401-1502. doi:10.1377/hlthaff.27.6.1491

Stone, H. S. (1971). *Introduction to computer organization and data structures*. New York: McGraw-Hill, Inc.

Zou, J., Ho, D. E., & Obermeyer, Z. (2023). Implications of predicting race variables from medical images. *Science, 381*(6654), 149-150. doi:10.1126/science.adh4260