Development of a Novel Diagnostic Tool for Cardiovascular Disease
(Technical Paper)

Implicit Bias Effects of Artificial Intelligence Implementation in Cardiovascular Care
(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Jashanjeet Matharoo
Spring, 2021

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

Signature Jashanjeet Matharoo Date 05/14/2021
Jashanjeet Matharoo

Approved John Alexander Hossack 4 Digitally signed by John Alexander Hossack 4 Date: 2021.05.08 20:24:33 -04'00' Date ____________
Dr. John A. Hossack, Department of Biomedical Engineering

Approved H Rogers ________________________ Date _5/4/2021_____
Dr. Hannah Rogers, Department of Engineering and Society

Introduction:

There is a technology that, in the next decade, will contribute to over $15 trillion to the global economy and raise the United States' GDP by 14%. The entire autonomous vehicle market, expected to be $127 billion by 2027, rests on the shoulders of this technology. In fact, you probably interact with it every single day (Neiger, 2018). What, then, is this ubiquitous and influential technology that is projected to change the world as we know it? The answer: artificial intelligence.

Artificial intelligence (AI) is defined as the ability of a computer to perform tasks generally associated with intelligent beings (Artificial Intelligence | Definition, Examples, and Applications, n.d.). This definition includes learning, decision-making, and even thinking independently. Such an influential technology carries the ability to shape society in the years to come, and therefore its creators are burdened with the responsibility of ensuring that its influence is morally and ethically sound. Thus, the proposed sociotechnical analysis will explore the implementation of AI in the diagnosis of cardiovascular disease (CVD) and determine the effects of societal implicit biases on the artificial intelligence software, while seeking to predict this technological outcome's subsequent impact on societal biases.

The technical portion of this proposal will explore the use of pulse waveform data to determine arterial stiffness, an indicator of cardiovascular disease. Data from the pulse waveforms will be processed and used to initially train a machine learning (AI) algorithm, which will then be able to classify novel pulse waveform readings into different risk levels and return a risk score with clinical significance. The diagnostic device developed around this software will offer a novel and cheap diagnostic method for the early detection of cardiovascular disease that patients can even use in their own homes.

Artificial Intelligence Implementation in Cardiovascular Diagnostics:

Cardiovascular disease (CVD) refers to the group of diseases affecting the heart and blood vessels and are a major cause of death worldwide. CVD is responsible for 17.9 million deaths annually, with ~25% occurring prematurely in individuals under the age of 70. Approximately 31% of annual global deaths are due to CVD (Cardiovascular Diseases (CVDs), n.d.). The burden of CVD in the United States accounts for an estimated 655,000 annual deaths. Most of these deaths, however, can be prevented through early detection and risk factor modification (CDC, 2020). Arterial stiffness is a known biomarker of CVD, and pulse wave velocity is a reliable diagnostic indicator of arterial stiffness and predictor of CV events. There are, however, barriers to measuring pulse wave velocity in clinical settings; specifically, the protocol to collect measurements is cumbersome and time-consuming (Lee & Joo, 2019). Development of a user-friendly technique for measuring arterial stiffness represents a significant advancement in the early detection and prevention of CVD. This individual technical project aims to extend the work on measuring PVW using pulse plethysmography.

The previous year's project focused on developing protocol to collect patient data and compute pulse wave velocity through use of easily available clinical equipment – namely pulse oximeters and sphygmomanometers. The current work is focused on expanding the trials of the previous study and further streamline the data collection protocol for easier clinical implementation. Additionally, the current work looks to explore the analysis of pulse waveforms for CVD biomarkers and utilization of computational and data permutation techniques to implement this analysis in a commercialized diagnostic tool. The technical project will explore the use of pulse waveform data to determine arterial stiffness, an indicator of cardiovascular disease. This exploration will be done through computational analysis of raw pulse waveforms targeting the quantification of specific characteristics. These characteristics will then be stored in data structures and correlated with patient phenotype to then be used as training sets for a machine-learning algorithm focused on scoring and classifying new waveforms by level of risk for cardiovascular disease. This method will offer a cheap, quick, and noninvasive alternative to current clinical diagnostic and screening methods.

Through the development of a pulse waveform analysis algorithm and complementary device, I expect to simplify the evaluation of arterial stiffness in a clinical setting and provide a diagnostic tool that patients can use in their homes. The outcome of this project will assist in the early detection and prevention of CVD, leading to a decrease in the global burden of disease derived from the condition. Additionally, this novel diagnostic method will reduce the economic burden of CVD on both patients and providers through less frequent follow-up visits and lower likelihood of longitudinal care due to late detection.

Predictive Quantification of AI Implicit Bias in Cardiovascular Applications

Can a machine be racist, sexist, or homophobic? Early advances in artificial intelligence (AI) that are used widely today have highlighted potential issues that indicate implicit bias in machines a possibility. Criticisms of AI often include implicit biases that are "baked in" to systems by the developers that make them, and the data sets they are trained on. For example, image databases will return images of primarily white males when searching "CEO" and resume screening software will discriminate against names aligning with certain ethnicities despite identical applicant qualifications (Smith, 2020). Also, word recognition software has shown heavy bias to associating words with "male" or "female" following common gender stereotypes (Biased Bots, n.d.). These outcomes may be due to the data sets that AI systems are trained on, as these could reflect human decision-making and societal inequities (Lum & Isaac, 2016).

Regardless of the cause of implicit bias in AI, the impacts of this phenomenon are tangible across many sectors of society. Online advertisement delivery programs target advertisements for high-interest credit cards and criminal background checks to Black users (Sweeney, 2013). Facial recognition software ubiquitous in modern devices misidentifies Black women between a quarter and a third of the time (Study Finds Gender and Skin-Type Bias in Commercial Artificial-Intelligence Systems, n.d.). Software used to determine whether or not defendants can be granted parole consistently report higher risk rates for Black individuals than White individuals, even if the Black individual has no prior record and the White individual does (Mattu, n.d.). Even household names like Amazon use hiring software that discriminates against racial minorities and women, as reports found that the AI used by the online retail giant perpetuated the 60-75% male composition of the workforce by citing a lack of "fit" as a reason to hire less women (No Surprise Amazon's AI Was Biased against Women, Says Sandra Wachter - Business Insider, n.d.).

Implicit bias in AI systems will have a particularly devastating impact in the field of medicine. Specifically, training healthcare AI technologies on past diagnoses that are influenced by provider biases will scale up the effects of harmful stereotypes prevalent in medicine today (Hall et al., 2015). For example, this technology will further exacerbate the unacceptable statistic that Black women are 2-3 times more likely to die from pregnancy related complications (Racial and Ethnic Disparities Continue in Pregnancy-Related Deaths | CDC Online Newsroom | CDC, 2019). Additionally, bias detected by patients reduces physician trust and leads to worsened outcomes long-term (How Does Implicit Bias by Physicians Affect Patients' Health Care?, n.d.). Thus, it is vital to understand the potential implications of implementing AI in various healthcare settings.

Although they are the prevailing attitudes of a subset of society that influence the actions of AI, the technology itself has an influence on the attitudes of the entire society. With AI specifically, the developers that create these programs are largely White or Asian males and therefore have a collective set of overlapping biases that are translated onto the technology they are developing (Manyika et al., 2019). Additionally, the technology they develop is trained on historical data that reflects often antiquated status quos in society. However, AI's subsequent effect on society is often overlooked. It is perceived by those without a technical background as completely objective and incapable of prejudice. Thus, when inequity as described above is a result of the implementation of AI systems, it is instead prescribed to fundamental differences or deficiencies in groups of people rather than systemic effects and intersectionality, diverting societal attention and resources away from combatting these pressing issues.

Given this positive feedback loop solidifying implicit bias and prejudice prevalent in today's society resultant of AI, the technology is best viewed through the lens of technological momentum. This framework is receptive to the social influences on technological artifacts but is also able to prospectively view the technology's counterinfluence on society (Hughes, 1994). Critiques of this approach come from proponents of social constructivism and technological determinism, with the former stating that technological artifacts are given meaning and agency by the society that adopts them, and the latter arguing that technology shapes society independent from the actions of that society's members (Klein & Kleinman, 2002; Smith, 1994). Additionally, critics cite technological momentum's lack of originality and label the framework derivative of its aforementioned predecessors (Colarossi, n.d.). However, technological momentum lies between social constructivism and technological determinism on the same continuum, and therefore affords the ability to determine how society has shaped artificial intelligence and predict how society will be influenced by artificial intelligence.

Research Question and Methods:

Ultimately, this analysis will provide a response to the following: How will the implementation of AI in cardiovascular medicine directly and indirectly impact the prevalence of medical bias as measured by outcome inequity?

The method used to approach the above question will primarily involve historical case studies of industries that have already implemented AI systems. Data from these studies will be grouped largely into three categories: implicit bias before AI implementation, due to AI implementation, and accompanying AI implementation. Open-source data on implicit bias tests and surveys has been aggregated and will be used for this analysis (Xu et al., 2013). Comparing the first two categories will allow the exploration of societal impact on technology, which contrasting the first and third categories will permit quantification of technological impact on society. Specifically, these aims will be accomplished by examining statistics on implicit bias aggregated through surveys and prior studies in all three of the aforementioned avenues across various industries. Then, a correlation factor will be computed between implicit bias before AI and due to AI as well as between implicit bias before AI and accompanying AI. These factors will then be applied to historical medical bias data to quantitatively determine future medical bias stemming directly and indirectly from the implementation of artificial intelligence in medicine, specifically in cardiovascular settings. Preliminary data available from the currently implemented AI systems in medicine will serve as benchmarks to validate this analysis method (Panch et al., n.d.).

Conclusion:

This prospectus presents a robust exploration of artificial intelligence's (AI's) implementation in cardiovascular medicine and the potential drawbacks of artificial intelligence systems used in healthcare settings. The technical component of this report offers a novel diagnostic tool for cardiovascular disease (CVD) that can be used by patients in their own homes, reducing costs for both patients and providers in addition to accelerating lifesaving diagnoses. The STS component of this report provides a predictive quantitative analysis of implicit bias influence on adverse patient outcomes through AI implementation in cardiovascular settings by conducting historical case studies across adjacent industries previously using AI. Overall, this prospectus aims to address an urgent clinical need in a responsible manner through understanding societal implications of technological advancement.

References

Artificial intelligence | Definition, Examples, and Applications. (n.d.). Encyclopedia Britannica. Retrieved October 16, 2020, from https://www.britannica.com/technology/artificial-intelligence

Biased bots: Artificial-intelligence systems echo human prejudices. (n.d.). Princeton University. Retrieved October 16, 2020, from https://www.princeton.edu/news/2017/04/18/biased-bots-artificial-intelligence-systems-echo-human-prejudices

Cardiovascular diseases (CVDs). (n.d.). Retrieved October 2, 2020, from https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

CDC. (2020, September 8). *Heart Disease Facts | cdc.gov*. Centers for Disease Control and Prevention. https://www.cdc.gov/heartdisease/facts.htm

Colarossi, A. (n.d.). *Summary and critique of Technological Momentum by Thomas P. Hughes*. Retrieved November 2, 2020, from https://www.academia.edu/1881558/Summary_and_critique_of_Technological_Momentum_by_Thomas_P_Hughes

Hall, W. J., Chapman, M. V., Lee, K. M., Merino, Y. M., Thomas, T. W., Payne, B. K., Eng, E., Day, S. H., & Coyne-Beasley, T. (2015). Implicit Racial/Ethnic Bias Among Health Care Professionals and Its Influence on Health Care Outcomes: A Systematic Review. *American Journal of Public Health*, *105*(12), e60–e76. https://doi.org/10.2105/AJPH.2015.302903

How does implicit bias by physicians affect patients' health care? (n.d.). Https://Www.Apa.Org. Retrieved October 16, 2020, from https://www.apa.org/monitor/2019/03/ce-corner

Hughes, T. P. *Technological Momentum*. (1994). Cambridge, Massachusetts. London, England. The MIT Press.

Klein, H. K., & Kleinman, D. L. (2002). The Social Construction of Technology: Structural Considerations. *Science, Technology, & Human Values,* 27(1), 28-52. https://doi.org/10.1177/016224390202700102

Lee, J.-G., & Joo, S.-J. (2019). Arterial stiffness and cardiovascular risk. *The Korean Journal of Internal Medicine*, *34*(3), 504–506. https://doi.org/10.3904/kjim.2019.110

Lum, K., & Isaac, W. (2016). To predict and serve? *Significance*, *13*(5), 14–19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Manyika, J., Silberg, J., & Presten, B. (2019, October 25). What Do We Do About the Biases in AI? *Harvard Business Review*. https://hbr.org/2019/10/what-do-we-do-about-the-biases-in-ai

Mattu, J. A., Jeff Larson,Lauren Kirchner,Surya. (n.d.). *Machine Bias*. ProPublica. Retrieved October 16, 2020, from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing?token=nponEfVIhbu_Yb4_j1EeltF4h5FW_Uc4

Neiger, C. (2018, July 28). *5 Reasons You Should Own at Least 1 Artificial Intelligence Stock*. The Motley Fool. https://www.fool.com/investing/2018/07/28/5-reasons-you-should-own-at-least-1-artificial-int.aspx

No surprise Amazon's AI was biased against women, says Sandra Wachter—Business Insider. (n.d.). Retrieved October 16, 2020, from https://www.businessinsider.com/amazon-ai-biased-against-women-no-surprise-sandra-wachter-2018-10

Panch, T., Mattie, H., & Atun, R. (n.d.). Artificial intelligence and algorithmic bias: Implications for health systems. *Journal of Global Health*, *9*(2). https://doi.org/10.7189/jogh.09.020318

Racial and Ethnic Disparities Continue in Pregnancy-Related Deaths | CDC Online Newsroom | CDC. (2019, September 6). https://www.cdc.gov/media/releases/2019/p0905-racial-ethnic-disparities-pregnancy-deaths.html

Smith, C. S. (2020, January 2). Dealing With Bias in Artificial Intelligence. *The New York Times*. https://www.nytimes.com/2019/11/19/technology/artificial-intelligence-bias.html

Smith, M.R. (1994). Technological Determinism in American Culture. *Does Technology Drive History?: The Dilemma of Technological Determinism*. (pp. 1-17). Cambridge, Massachusetts. London, England. The MIT Press.

Study finds gender and skin-type bias in commercial artificial-intelligence systems. (n.d.). MIT News | Massachusetts Institute of Technology. Retrieved October 16, 2020, from https://news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212

Sweeney, L. (2013). *Discrimination in Online Ad Delivery* (SSRN Scholarly Paper ID 2208240). Social Science Research Network. https://doi.org/10.2139/ssrn.2208240

Xu, F. K., Nosek, B. A., Greenwald, A. G., Ratliff, K., Bar-Anan, Y., Umansky, E., Banaji, M. R., Lofaro, N., Smith, C., & Axt, J. (2013). *Project Implicit Demo Website Datasets*. https://doi.org/10.17605/OSF.IO/Y9HIQ