Full-Stack Development: Involving Relevant Stakeholders in Design
(Technical Project)

The Ethical and Legal Implications of Algorithmic Diagnosis in Medical Imaging
(Technical Project)

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

By
Tillman Dean

November 1, 2022

Technical Team Members: Tillman Dean

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.

Signed: ___________Tillman Dean______________________

ADVISORS

Approved: ____
MC Forelle, Department of Engineering and Society

Approved: ____
Brianna Morrison, Department of Computer Science

Introduction

Over the past couple decades, machine learning has become an almost universal tool to be applied with success across many different fields. Algorithms' capacity to predict, analyze, and optimize has dramatically reshaped a lot of jobs and businesses. One industry in which it has begun to be applied is that of healthcare (Bhardwaj et al., 2017). For instance, some deep learning algorithms have been proven to be more accurate in analyzing medical imaging, such as MRI, than human doctors (Nishida et al., 2022). Early detection of a disease or tumor from a scan can dramatically improve a patient's chances of survival and recovery. The potential for dramatically increased accuracy and efficiency would be welcome in the US, a country where "healthcare costs are rising at rates close to double of our economic growth rate" (Bhardwaj et al., 2017).

Despite some documented success, these algorithms must also deal with the fact that the healthcare industry is one with uniquely extreme ethical concerns. Misdiagnosis by a computer does not allow for an easy assignation of blame. A doctor not properly understanding the way the algorithm works could lead to misuse or lack of trust. In many image recognition algorithms, accuracy has been shown to differ based on race or gender, creating inequity and bias in healthcare (Chen et al., 2021). Fundamentally, as Karmakar (2021) notes, the nature of algorithms and medicine itself are at odds:

> Whilst any AI can be said to operate through objective millisecond binary decisions, with outcomes ascribed fixed values by stakeholders… medicine works by implementing long-term decisions optimised by complex overlapping ontologies, such as guides of best practice, that call for subjective value decisions, utilising a plurality of objective data with discretionary nuance.

Due to the widespread implementation of the technology and the documentation of these problems, a variety of solutions have been proposed.

One proposed solution to eliminate some of these concerns is implementing participatory design, a theory in which all stakeholders are involved in the design process (de Boer & Kudina, 2021). In the technical portion of my project, I will be discussing my internship from last summer at Capital One, focusing in on the importance of the participatory design strategy we used when developing our application for the company. By involving the users of our new application in the design process early on and throughout the demo phase, we were able to save a lot of time and develop a product that was more easily understood and more useful to them. In my STS portion, I will be exploring how the implementation of medical image analysis algorithms manage the ethical risks with the benefits, focusing on the participatory design framework.

Technical Topic

Over the summer before my fourth year at the University of Virginia, I had the opportunity to intern at Capital One, a large financial company based in McLean, VA, as a software engineer. Although I had taken part in group projects in software development classes at school, this internship was my first professional experience as a developer, and as a result, I was able take away a lot of important lessons from it.

On the job, I worked on a team of nine: four interns, a manager, a product owner, an agile development lead, and two full-time developers. Our team was tasked with developing an enterprise-used software application based off a previous one that was no longer being used and had not ever been fully developed to solve the problem at hand. Our application was intended to streamline the process of adding a new data model to the enterprise document management platform, which at the time required extensive meeting time with a developer team and ended with them hard coding the solution into the database. The application was also supposed to allow easy viewing of the database so that users could explore and even build off existing models.

Our team worked within an Agile framework called Kanban. We had daily stand-ups where we met to review our progress on our currently assigned tasks and to bring up any issues or help we needed. We used JIRA and Slack to communicate and keep our progress up to date. I worked on a variety of tasks including developing new UI using Vue.js, creating API endpoints to allow our front-end to communicate with our DynamoDB database, and create files to deploy our application through the enterprise Jenkins pipeline.

My goal over the course of the internship was to learn as much as I could about what being a professional developer is like as well as acquire valuable skills that I might use within the industry. I got a lot of exposure to AWS, Vue.js, and Jenkins, all of which I was very unfamiliar with before the summer, so I was forced to ask a lot of questions to complete my tasks. I also got valuable experience applying an Agile framework in development from start to finish on a project and experiencing daily check-ins on the status of everyone on the team.

One of the lasting things that stuck with me the most was the amount of time and thought we spent on design before we started coding and after we had developed our earliest demos. Our team was very deliberate in demoing the software for the teams that would be using it the most and taking feedback on how to make it more intuitive and useful to them. Up until this point, I had mainly just briefly sketched out a basic design or dove straight into coding my solution. However, as this was the first time that I was developing something that would make a quantifiable impact on many people's lives, it made sense that we were more methodical and thorough with our planning. Being able take the frameworks and techniques learned in a class like Advanced Software Development and apply them in a work environment where the results were tangible and important to people around me was a great framing for my last year of college before the professional world. I had not learned about the framework of participatory design at this point, but now that I have, I plan on looking back at our design process through that lens.

STS Topic

Medical imaging such as MRIs, x-rays, and ultrasound are crucial tools used by doctors to make diagnoses on patients. Traditionally these images are examined by doctors to detect problems, such as signs of a stroke or tumor, but in recent years, machine learning image recognition algorithms have been used in tandem with human doctors to analyze these medical images. These algorithms have gotten to the point where they are often performing on or above the level of human analysis (Nishida et al., 2022; O'Connell et al., 2022). Even in other areas such as stroke diagnosis, where they are limited by the amount of data available, algorithms are still a useful tool to be paired with experts when examining medical imaging (Mainali et al., 2021). The benefits brought to fields like radiology make this technology impossible to ignore as the future of the field (Wang & Summers, 2012).

Use of a technology to make important medical decisions inevitably leads to a variety of ethical and legal concerns (Grote & Berens, 2020). For example, if a mistake is made, placing the responsibility on a doctor is not possible. If the doctors or patients, do not fully understand the technology, a lack of trust can develop even if studies show it to be more accurate. A group of scientists completed several studies showing that patients trusted AI diagnosis less than human doctors, even when they are shown that the AI is more accurate (Juravle et al., 2020). Dai and Tayur (2022) argue that for the technology to succeed trust and buy-in from both patients and physicians are necessary. Some of the ethical issues arise from purely technical problems such as insufficient data and weak labels in the algorithms (de Bruijne, 2016). In response to these issues, frameworks and guidelines have been proposed on how to implement it while avoiding these concerns (Leimanis & Palkova, 2021). One popular solution is that of participatory design, the concept that all users and stakeholders should have equal input into design (de Boer & Kudina, 2021). The solution argues that a diverse group of doctors and patients should be included in the design process to mitigate bias and lack of trust.

Unfortunately, machine learning algorithms are notoriously black box, meaning that they are hard to understand from an outside view. Because they are not fully coded by a human and instead trained on a large dataset, changing automatically to become more accurate, their decision-making and true design process cannot even be explained by the developer. Komura and Ishikawa note this complication: "Typical DL processes for histopathological image class identification (such as cancer versus normal tissue) only provides the result of the classification and its score; there is no way for the pathologists to know the rationale for classification and the features captured by the DL network. (2019)." Due to this opaque design process traditional participatory design frameworks are difficult to apply and may require change to work on this problem (Donia & Shaw, 2021). The algorithms themselves often give no explicit reasoning behind identifying a tumor in an MRI image, just a yes or no output. To a patient or a doctor attempting to explain the diagnosis, this lack of reasoning can be both frustrating and concerning. Other frameworks have been suggested to repair this damage to the physician and patient relationship, including Guobin and Xiaoxi's combination of traditional Chinese medical theory with AI (2019).

In my STS project, I will explore these problems and the framework of participatory design in detail, attempting to understand how this framework could make algorithmic medical imaging analysis a viable tool that physicians can use with confidence and good conscience.

Research question and methods

How can participatory design be used to balance ethical concerns with potentially increased accuracy in machine learning algorithms analyzing medical imaging?

This question is vital due to the concerns outlined in the STS section above. Without addressing these concerns, the healthcare industry, one that is built on the universal agreement to pursue the goal of personal health, could become inequitable, legally liable, and full of distrust with disastrous consequences. I will research this question first by reviewing data from existing studies on the accuracy of machine learning algorithms in medical imaging. Doing so will help establish the potential benefits that algorithmic medical image analysis offers and the provide some insight into the usage trends. I will also spend time studying how popular image recognition algorithms were designed and compare different types of these algorithms' accuracy and bias over the same datasets (Zhang & Sejdic, 2019).

On the ethical and legal side, I will examine several pieces written by medical professionals and STS academics on the potential challenges created by introducing this technology to make diagnoses with regards to bias, accountability, and data privacy (Jaremko et al, 2019; Ngiam & Khor, 2019; Sollini et al., 2020). Then I will examine studies that explore human trust in AI diagnosis compared to a diagnosis from doctors.

Finally, I will examine many of the proposed solutions to these challenges, focusing on participatory design and the potential challenges of implementing this framework to this specific problem. I will also examine guidance from experts and healthcare professionals on how to practically implement algorithmic diagnosis on medical imaging.

Conclusion

Researching this question will be extremely valuable and important in making sure that patients receive equal and quality care in the future when these algorithms inevitably take on an even larger role in healthcare. I anticipate finding that biases can be incredibly hard to identify or fix when looking directly at the code for the algorithm itself, as much of the way these algorithms work is very black box, such that the user is only seeing the inputs and outputs to the code. I also anticipate there being a much greater distrust of algorithmic diagnosis as a result that must be overcome for them to take on a larger role in medicine. I expect to find that while participatory design can help alleviate the second of these two concerns, it will be tough to do anything for the first concern other than ensure the data collection process is equitable. This project will impact each key stakeholder in the technology including developers, medical practitioners, and patients receiving care, and I hope that it will give each of them a better understanding of the larger issue and an approach to solve it.

Bibliography

Bhardwaj, R., Nambiar, A. R., & Dutta, D. (2017). A Study of Machine Learning in Healthcare. In S. Reisman, S. I. Ahamed, C. Demartini, T. Conte, L. Liu, W. Claycomb, M. Nakamura, E. Tovar, S. Cimato, C. H. Lung, H. Takakura, J. J. Yang, T. Akiyama, Z. Zhang, & K. Hasan (Eds.), *2017 Ieee 41st Annual Computer Software and Applications Conference (compsac), Vol 2* (pp. 236–241). Ieee. https://doi.org/10.1109/COMPSAC.2017.164

Chen, I. Y., Pierson, E., Rose, S., Joshi, S., Ferryman, K., & Ghassemi, M. (2021). Ethical Machine Learning in Healthcare. In R. B. Altman (Ed.), *Annual Review of Biomedical Data Science, Vol 4* (Vol. 4, pp. 123–144). Annual Reviews. https://doi.org/10.1146/annurev-biodatasci-092820-114757

Dai, T., & Tayur, S. (2022). Designing AI-augmented healthcare delivery systems for physician buy-in and patient acceptance. *Production and Operations Management*. https://doi.org/10.1111/poms.13850

de Boer, B., & Kudina, O. (2021). What is morally at stake when using algorithms to make medical diagnoses? Expanding the discussion beyond risks and harms. *Theoretical Medicine and Bioethics*, *42*(5–6), 245–266. https://doi.org/10.1007/s11017-021-09553-0

de Bruijne, M. (2016). Machine learning approaches in medical image analysis: From detection to diagnosis. *Medical Image Analysis*, *33*, 94–97. https://doi.org/10.1016/j.media.2016.06.032

Donia, J., & Shaw, J. (2021). Co-design and Ethical Artificial Intelligence for Health: Myths and Misconceptions. *Aies '21: Proceedings of the 2021 Aaai/Acm Conference on Ai, Ethics, and Society*, 77–77. https://doi.org/10.1145/3461702.3462537

Grote, T., & Berens, P. (2020). On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics*, *46*(3), 205–211. https://doi.org/10.1136/medethics-2019-105586

Guobin, C., & Xiaoxi, W. (2019). Between AI Doctor and Patient: The Logic of AI Diagnosis and its Impact on Patient Subjectivity. *International Journal of Chinese & Comparative Philosophy of Medicine*, *17*(2), 11–36.

Jaremko, J. L., Azar, M., Bromwich, R., Lum, A., Cheong, L. H. A., Gibert, M., Laviolette, F., Gray, B., Reinhold, C., Cicero, M., Chong, J., Shaw, J., Rybicki, F. J., Hurrell, C., Lee, E., & Tang, A. (2019). Canadian Association of Radiologists White Paper on Ethical and Legal Issues Related to Artificial Intelligence in Radiology. *Canadian Association of Radiologists Journal-Journal De L Association Canadienne Des Radiologistes*, *70*(2), 107–118. https://doi.org/10.1016/j.carj.2019.03.001

Juravle, G., Boudouraki, A., Terziyska, M., & Rezlescu, C. (2020). Trust in artificial intelligence for medical diagnoses. In B. L. Parkin (Ed.), *Real-World Applications in Cognitive Neuroscience* (Vol. 253, pp. 263–282). Elsevier. https://doi.org/10.1016/bs.pbr.2020.06.006

Karmakar, S. (2022). Artificial Intelligence: The future of medicine, or an overhyped and dangerous idea? *Irish Journal of Medical Science*, *191*(5), 1991–1994. https://doi.org/10.1007/s11845-021-02853-3

Komura, D., & Ishikawa, S. (2019). Machine learning approaches for pathologic diagnosis. *Virchows Archiv*, *475*(2), 131–138. https://doi.org/10.1007/s00428-019-02594-w

Leimanis, A., & Palkova, K. (2021). Ethical Guidelines for Artificial Intelligence in Healthcare from the Sustainable Development Perspective. *European Journal of Sustainable Development*, *10*(1), 90–102. https://doi.org/10.14207/ejsd.2021.v10n1p90

Mainali, S., Darsie, M. E., & Smetana, K. S. (2021). Machine Learning in Action: Stroke Diagnosis and Outcome Prediction. *Frontiers in Neurology*, *12*, 734345. https://doi.org/10.3389/fneur.2021.734345

Ngiam, K. Y., & Khor, I. W. (2019). Big data and machine learning algorithms for health-care delivery. *Lancet Oncology*, *20*(5), E262–E273. https://doi.org/10.1016/S1470-2045(19)30149-4

Nishida, N., Yamakawa, M., Shiina, T., Mekada, Y., Nishida, M., Sakamoto, N., Nishimura, T., Iijima, H., Hirai, T., Takahashi, K., Sato, M., Tateishi, R., Ogawa, M., Mori, H., Kitano, M., Toyoda, H., Ogawa, C., & Kudo, M. (2022). Artificial intelligence (AI) models for the ultrasonographic diagnosis of liver tumors and comparison of diagnostic accuracies between AI and human experts. *Journal of Gastroenterology*, *57*(4), 309–321. https://doi.org/10.1007/s00535-022-01849-9

O'Connell, A. M., Bartolotta, T. V., Orlando, A., Jung, S.-H., Baek, J., & Parker, K. J. (2022). Diagnostic Performance of An Artificial Intelligence System in Breast Ultrasound. *Journal of Ultrasound in Medicine*, *41*(1), 97–105. https://doi.org/10.1002/jum.15684

Sollini, M., Bartoli, F., Marciano, A., Zanca, R., Slart, R. H. J. A., & Erba, P. A. (2020). Artificial intelligence and hybrid imaging: The best match for personalized medicine in oncology. *European Journal of Hybrid Imaging*, *4*(1), 24. https://doi.org/10.1186/s41824-020-00094-8

Wang, S., & Summers, R. M. (2012). Machine learning and radiology. *Medical Image Analysis*, *16*(5), 933–951. https://doi.org/10.1016/j.media.2012.02.005

Zhang, Z., & Sejdic, E. (2019). Radiological images and machine learning: Trends, perspectives, and prospects. *Computers in Biology and Medicine*, *108*, 354–370. https://doi.org/10.1016/j.compbiomed.2019.02.017