Design and Implementation of Test Automation Frameworks for Loan Delivery Assets

(Technical Project)

Predictive Analytics on Physicians' Decision-Making in Pediatric Healthcare

(STS 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
Nikita Jeyasingh
December 15th, 2023

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.

ADVISORS

MC Forelle, Department of Engineering and Society

Wajih Ul Hassan, Department of Computer Science

Introduction

In the ever-changing world of technological advancements, integrating analytics and data-driven technologies has transformed outcomes across various industries. "PA [predictive analysis] has been proliferating over the last decade or so in a range of sectors, with direct or indirect impact on children, including health care" (Rahman & Keseru, 2021). This integration has had a profound impact on pediatric healthcare, specifically the implementation of predictive analytics, which has shaped the way healthcare professionals diagnose and treat their young patients. The considerable amount of data generated from medical devices, electronic health records, and patient monitoring systems, with advancements in computing technologies and artificial intelligence, has led to a rapid shift towards data-driven solutions.

The implementation of predictive analytics (PA) for pediatric healthcare has received a considerable amount of attention for its potential capabilities to enhance patient outcomes and to aid healthcare professionals' decision-making for treatment. Data-driven solutions can significantly impact the pediatric healthcare industry. "The world is heavily evolving regarding healthcare, with an increasing reliance on AI-driven solutions" (Kwok et al., 2022). However, the integration of such technologies into this social system comes with its own set of challenges. Ethical accountability ensures that the advancements of AI-driven solutions will not compromise moral standards and patient well-being (Farhud et al., 2021). Overreliance on PA could lead to misdiagnosis or incorrect treatment decisions and undermine the role of the healthcare professionals' expertise, particularly if the models are not thoroughly tested accurately.

The ripple effect of integrating PA into the existing healthcare structures presents a crucial aspect to consider regarding its potential to impact decisions made by healthcare providers. Exploring the impact of predictive analytics on healthcare provider decisions

recognizes the importance of acknowledging how integrating technology can reshape the social and cultural fabric of the healthcare environment. (Selbst et al., 2019).

Automation testing is pivotal in ensuring the reliability and accuracy of the data and algorithms that support predictive analytics. Interning at Fannie Mae this past summer, I was able to take part in a project that designed a test automation framework to support loan delivery assets in the context of the finance industry. Test automation allows for the rapid and repeatable execution of test cases, thus significantly reducing the time and effort required for testing. The process of automating tests not only streamlines testing procedures but significantly reduces the chances of human error, which can compromise the accuracy of the data. This project facilitated the development of a test automation framework, which worked to eliminate the need for manual testing. Manual testing, especially for complex applications, is time-consuming and error prone. Automation tests ensure more accurate data in the realm of predictive analytics for pediatric healthcare.

Inconsistencies with manual testing inadvertently led to data discrepancies, making it difficult to rely on results for making informed decisions for pediatric healthcare. By implementing automation, it ensures tests are executed consistently, thus minimizing the possibility of data errors and inaccuracies. By further examining fairness and abstraction (Selbst et al., 2019), it will analyze how the use of PA impacts decisions made by healthcare providers in pediatric healthcare.

Technical Project

The summer of 2023, I interned at Fannie Mae, a mortgage financing company, as a data engineer for the Chief Data Office. Throughout my time working at Fannie Mae, I was placed on

the LSDI ADS Team (Loan Sourcing Data Integration Acquisition Data Store). The team worked to consume data from upstream services, process that data, and apply it to business transformations to vend it for downstream systems. For ten weeks, I had the privilege of taking part in daily standups, spring planning, grooming sessions, and live demos. During this time, I was tasked to aid with the design and implementation of the test automation framework for loan delivery assets. Data from loan delivery assets includes data from the process of delivering loans in the context of the financial industry. These assets include loan documents, mortgage-backed securities, servicing rights, and various data sources that are crucial to financial institutions. This data is crucial as it allows Fannie Mae to make critical business decisions and develop risk management strategies. The Chief Data Office (CDO) set a goal for their assets to be primarily covered under test automation. Automated testing is a software technique that automates the process of validating the functionality of software and ensures its requirements are being tested before it's released into production. Some benefits of automation include accelerated results, improved coverage, earlier detection of bugs, and accurate data validation.

The ultimate goal of this project was to design a custom-built test automation framework for LSDI (Loan Sourcing Data Integration) processes and all of their source systems. This framework was created to move along with the ongoing cloud migration and is structured in a manner that could be used for multiple streams and different types of testing for the acquisition of data store assets. This framework was built using AWS Cloud services, AWS Lambda, DBeaver, and Python. Each source system required a unique test automation framework to successfully test data across multiple streams accurately. It was critical that the data from upstream sources for each source system was tested thoroughly against test data to ensure that during the downstream vending process, the packaged data was accurate.

Working closely with the members of my team at Fannie Mae, we created a design process to ensure that the implemented framework accurately tested incoming data to ensure consistency and accuracy. For this project, the automation testing enhances testing processes to ensure data is validated and generated accurately. Human errors, time-consuming efforts, and limited test coverage often plague manual testing. Such errors can introduce inaccuracies and provide inaccurate results, thus affecting decisions made from data results. This vulnerability of human error emphasized the need for automation testing, which offers a systematic approach to ensure reliable data.

STS Topic

Pediatric healthcare represents a unique and sensitive branch within the healthcare ecosystem. The implementation of predictive analytics has the potential to disrupt and influence healthcare provider decision-making in ways that may not seem immediately apparent. The application of PA in this sector requires a comprehensive examination that acknowledges the consequences of data-driven decision-making. (Rahman et al., 2021). This socio-technical study seeks to examine the following: How does the implementation of predictive analytics in pediatric healthcare impact the decision-making process of healthcare providers for patient care and treatment?

The use of PA in pediatric healthcare is defined by its reliance on past and real-time patient data, which can contain biases, and its use of algorithms to make predictions that affect patient care decisions. "An estimate of the National Academy of Sciences states that around 5% of US adults seeking healthcare advice are subject to diagnostic error. Furthermore, relevant diagnostic errors are assumed to contribute to approximately 10% of all patient deaths" (Grote &

Berens, 2020). While diagnostic errors are occasionally inevitable as part of human error, when the use of PA is involved, it requires a decision-making process by the healthcare provider to assess all evidence and experience available. "Healthcare professionals typically specify and enforce obligations for their members, thereby seeking to ensure that persons who enter into relationships with these professionals will find them competent and trustworthy" (Beauchamp et al., 2001). Patients and families put their lives in the hands of healthcare professionals, which highlights the critical importance of their ability to trust them with their decisions for care and treatment plans. Ethical considerations in this context relate to the fairness in algorithmic design and testing process for biases to affect the quality of care and patient outcomes; it is the provider's decision how they will treat the patient with the understanding of the risks of the use of PA to aid their decision for treatment and care.

The concept of abstraction in computer science is used to ensure that machine learning applications are producing algorithms that produce 'fair' outcomes. By determining that technical interventions, predictive analytics in this case, are deemed ineffective, it can be dangerous when these implementations have direct effects on decision-making systems. PA increasingly makes critical decisions in youth patients care, understanding how such technologies impact healthcare providers' decision-making is vital to ensure ethical, equitable, and responsible use of these tools; it can be done by examining each of the abstraction traps proposed by Selbst et. al. (2019) including: the framing trap, the portability trap, the formalism trap, the ripple effect trap, and the solutionism trap will allow for thorough research and analysis. "Failure to understand how the insertion of technology into an existing social system changes the behaviors and embedded values of the pre-existing system" (Selbst et al., 2019). Understanding how these socio-technical changes are essential for addressing concerns for data-driven interventions and decisions, it is

important for ensuring that the implementation of PA in the pediatric healthcare industry aligns with the best interests of both the healthcare providers and the pediatric patients.

Tackling these issues requires sufficient research and comprehensive reviews from reviewing existing literature on pediatric case studies, healthcare guidelines on the use of technology in patient care, and physician testimonies regarding AI implementation in healthcare. Therefore, I will examine STS literature regarding fairness, abstraction traps, and ethical debates in order to synthesize data-driven decision-making processes of healthcare providers.

In the context of predictive analytics for pediatric healthcare decisions, the use of automated testing assumes critical importance. Predictive analytics in pediatric healthcare relies on historical and real-time data from pediatric patients for healthcare professionals to make informed decisions regarding patient care and treatment. Automated testing ensures that the data used in the decision-making process is error-free, consistent, and dependable. It safeguards against inaccuracies and discrepancies that could result in poor decision-making by healthcare providers using the outcomes of PA to assume diagnosis and treatment plans for patients. "Cognitive automation, empowered with A.I. Machine Learning, enables new predictive analysis methods to transform the way development teams develop, test, and deploy powerful products that meet users' needs." (Nadira). Automated testing for the use of predictive analytics will empower healthcare providers to make data-driven decisions that will positively impact the health and well-being of pediatric patients in order to ensure that each treatment plan is built on a solid foundation of accurate and high-quality data.

Decisions made in pediatric healthcare carry a profound significance, as they directly influence the well-being of the most vulnerable patients - children. Every decision regarding diagnosis, treatment, or care made by a healthcare professional carries the life-altering

responsibility of a child's future health, development, and quality of life. The decisions impact not only the patients but also families, communities, and the overarching healthcare system; therefore, the examination of the influence of PA on these decisions is critical.

Conclusion

Analyzing the impact of predictive analytics on data-driven decision-making by healthcare providers holds the potential for transformative and positive change in the realm of pediatric healthcare. Healthcare providers, pediatric patients, and their families stand to gain from the increased accuracy and reliability of these decisions that are enabled by thorough automation testing for these models. Healthcare providers will have a more solid foundation of accurate and high-quality data, leading to improved patient care, diagnosis, and treatment plans. In addition, the reduction of errors and efficiency in financial institutions, as a result of the technical project, will indirectly support the stability of healthcare providers and their decisions for pediatric patient care. By creating more accurate data sets for predictive analytics and fostering a deeper understanding of socio-technical implications, this research will contribute to resolving challenges introduced by the integration of predictive analytics in pediatric healthcare, which ultimately benefits the well-being of children and their families who are relying on the healthcare system as a whole.

References

Baer, G. R., & Nelson, R. M. (2007). *A review of ethical issues involved in premature birth*. National Library of Medicine. https://www.ncbi.nlm.nih.gov/books/NBK11389/

Beauchamp, T., & Childress, J. (2001). *Principles of Biomedical Ethics*. Oxford University Press Inc. https://books.google.com/books?hl=en&lr=&id=_14H7MOw1o4C&oi=fnd&pg=PR9&ots=1x_n0JzmYw&sig=ijr3esuVK8CBCL15htWTX-y3k6g#v=onepage&q&f=false

Daboval, T., & Shidler, S. (2014, June). *Ethical framework for shared decision making in the neonatal intensive care unit: Communicative ethics*. Paediatrics & child health. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4173956/Dabokin doi: 10.1093/pch/19.6.302. PMID: 25332659; PMCID: PMC4173956.

Drummond, M. (2023, October 12). *Using machine learning & AI to improve NICU outcomes*. Physician Resource. https://www.uvaphysicianresource.com/ai-machine-learning-improve-nicu/

Farhud, D. D., & Zokaei, S. (2021, November). *Ethical issues of Artificial Intelligence in medicine and Healthcare*. Iranian journal of public health. doi: 10.18502/ijph.v50i11.7600. PMID: 35223619; PMCID: PMC8826344.

Grote, T., & Berens, P. (2020, March 1). *On the ethics of algorithmic decision-making in Healthcare*. Journal of Medical Ethics. https://jme.bmj.com/content/46/3/205

Kwok, T., Boardman, J., Boylan, G., Saffaran, S., & Henry, C. (2022, April 18). *Application and potential of artificial intelligence in neonatal ...* Seminars in Fetal and Neonatal Medicine. https://www.sfnmjournal.com/article/S1744-165X(22)00025-7/fulltextDOI:https://doi.org/10.1016/j.siny.2022.101346

Nadira, K. (2022, February 24). *Software testing optimization with predictive analysis*. Gleematic A.I. https://gleematic.com/optimizing-and-improving-software-testing-with-predictive-

analytics/#:~:text=Predictive%20analytics%20in%20software%20testing,data%20then%20understand%20the%20pattern.

Pai, K.-C., Wang, M.-S., Chen, Y.-F., Tseng, C.-H., Liu, P.-Y., Chen, L.-C., Sheu, R.-K., & Wu, C.-L. (2021, June 29). *An artificial intelligence approach to bloodstream infections prediction*. MDPI. https://www.mdpi.com/2077-0383/10/13/2901

Rahman, Z., & Keseru, J. (2021, November). *Predictive analytics for children: An assessment of ethical considerations, risks, and Benefits*. UNICEF. https://www.unicef-irc.org/publications/1275-predictive-analytics-for-children-an-assessment-of-ethical-considerations-risks-and-benefits.html

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019, January 1). *Fairness and abstraction in sociotechnical systems: Proceedings of the conference on fairness, accountability, and transparency*. ACM Conferences. https://dl.acm.org/doi/10.1145/3287560.3287598

Siegal, E. (2020, October 23). When does predictive technology become unethical? https://hbr.org/2020/10/when-does-predictive-technology-become-unethical

Spaeder, M. C., Moorman, J. R., Tran, C. A., Keim-Malpass, J., Zschaebitz, J. V., Lake, D. E., & Clark, M. T. (2019, July 31). *Predictive analytics in the Pediatric Intensive Care Unit for early identification of sepsis: Capturing the context of age*. Nature News. https://www.nature.com/articles/s41390-019-0518-1

Stanak, M. (2018, September 7). *Professional ethics: The case of neonatology - medicine, health care and philosophy*. SpringerLink. https://doi.org/10.1007/s11019-018-9863-9

Sullivan, A., & Cummings, C. (2020, April). *Historical perspectives: Shared decision making in the NICU*. NeoReviews. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8049458/