Thesis Project Portfolio

Demystifying Black Boxes: Analysis of Neural Network Transparency Research

(Technical Report)

Analyzing Societal Transitions in a Computer Science Curriculum Revamp

(STS Research Paper)

An Undergraduate Thesis

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Parker Hutchinson

Fall, 2023

Department of Computer Science

Table of Contents

Sociotechnical Synthesis

(Executive Summary)

An Analysis of AI Explainability and Computer Science Education Revamping

As a computer science student at the University of Virginia, I have noticed gained valuable knowledge in programming skills and in other disciplines. Through software engineering internships, however, I have noticed which of the skills I was taught were readily applicable to the outside world of engineering. These skills include familiarity with collaborative tools like GitHub, the ability to follow a rigorous testing protocol while programming, and more. I have also realized that some of these concepts are not emphasized in the UVA computer science curriculum. I sought to analyze this disparity, although it did not directly relate to my machine-learning focused research project, because I think that addressing it will bring great value to the next generation of UVA computer science students.

First, in my STS research, I analyzed how the UVA computer science curriculum could undergo a transition to a new system which incorporated the values of stakeholders while maintaining its well-known reputation. I utilized Geel's Multi-Level Perspective (MLP) to note the various stakeholders at play and to thoughtfully consider how changes could percolate throughout the socio-technical context. I utilized my knowledge of the system present at UVA and did further research into what elements seem to work to the benefit of students and stakeholders at other universities while doing my research.

Next, in the technical portion of my STS research, I produced a literature review on the state of artificial intelligence (AI) explainability. I covered research papers on various different types of models and how they can be explained in different ways; for example, object-detection

models can be made to produce “saliency maps” showing how relevant each area of an image is to its output, whereas a text-based model would likely use a different model for producing insights about its reasoning. I synthesized the sources to arrive at a summary of modern tools utilized, and also noted how certain areas, especially the medical sector, stand to gain enormous benefits from continued AI research because of the importance of understanding in those areas of work. My project helped me to better understand ethical considerations as they pertain to engineers because the researchers building these models must keep in mind the duty they owe to the end user, and how they can accomplish that by making their tools as predictable and explainable as possible.

By considering technical, organizational, and cultural elements simultaneously, one can arrive at a holistic view of a technology and its impact. By only considering one of these aspects at a time, one may arrive at some useful insights, but will likely forego important impacts to stakeholders who should be considered. Utilizing these STS perspectives supports ethical responsibility in engineering because it forces an engineer to consider views and ideas that are not their own, broadening the scope of their knowledge and increasing the likelihood that the values of all people involved in the process and its effects will be respected. I found that using this broadened scope to analyze the computer science curriculum at UVA was particularly beneficial for considering how educational ideas might be tested, take root, and progress upwards into real departmental change that benefits students, employers, and other stakeholders involved.