Thesis Portfolio

Extracting Machine Learning Features to Detect Malicious HTTPS Traffic

(Technical Report)

Maintaining Accountability in a Criminal Justice System that Uses Machine Learning

(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

Adam Klein
Spring, 2020

Department of Computer Science

Table of Contents

Sociotechnical Synthesis

Machine learning algorithms are part of a methodology that allows computers to extract insights from data without being programmed to follow a specific model. The lack of a hardcoded model for data analysis allows the computer to unlock insights in data that may have otherwise been invisible to the human eye. However, machine learning has been adopted into many realms of society without a clear understanding of its limitations. This portfolio describes two use cases for machine learning. The first is in the realm of cybersecurity. The second is in the criminal justice system. In both of these cases, machine learning offers the potential to assess threats to society. However, the use of machine learning in cybersecurity involves the analysis of network traffic, while the use of machine learning in the criminal justice system involves the analysis of human behavior. The contrast between these two uses serves as a case study in how society should temper its expectations of what can be accomplished with machine learning.

The technical portion of this portfolio discusses how machine learning can be used to detect encrypted malware communications. Encryption allows the public to securely send data across the Internet. However, encryption also allows actors with harmful intentions to coordinate cyberattacks in secrecy. This paper offers an approach that would allow for the detection of threats by inspecting metadata that is available prior to decrypting traffic. By manipulating this metadata, it is possible to compute many different features that could be indicative of malicious behavior. This paper describes how a software package can be developed to extract these features from the vast quantities of data that are produced each day on an enterprise-scale network, such as that at the University of Virginia. These results provide a new methodology to detect cyberattacks on enterprise networks.

The STS portion of this portfolio discusses how software that helps humans make decisions impacts a society's ability to hold decision-makers accountable for their actions. Specifically, it answers the question of how law enforcement agencies and courts can be held accountable when they make decisions that are informed by an algorithmic black box. This question is investigated by first looking at how machine learning technologies have promoted discrimination in the criminal justice system. Establishing machine learning's failure to create equitable outcomes demonstrates that mathematical processes do not necessarily yield less biased decisions. Case law is then used to determine how accountability has traditionally been viewed in the criminal justice system. This policy analysis is organized within the framework of actor-network theory. The paper builds a network to describe the relationships between actors such as the public, law enforcement agencies, and the social values of accountability and transparency. Technological actors, such as software and data sets, are then introduced into the network to establish how these new actors change the relationship between the public and the value of accountability. Based on these findings, a framework is proposed to help uphold accountability in systems that incorporate software into decision-making processes. This framework will serve as a guide for how a society can ethically incorporate more intelligent software into key areas of life. This guide is essential to both STS and engineering, as it will provide principles for the safe development and use of machine learning technologies in many parts of society.

The work on my technical paper has proven to be essential in informing my STS research. Building a model to detect threats on a computer network taught me how a computer uses machine learning to solve classification problems. Understanding that these "decisions" are based on a completely indecipherable mathematical process inspired research into how this could

create conflicts in the criminal justice system. Given that the math inside a machine learning algorithm cannot be translated into an articulable decision process, it seems that there is an inherent conflict with the judgment of a person's guilt in a court of law, which is a process that is inextricably attached to deliberation. By having a technical understanding of machine learning models, I was able to more effectively analyze how they would serve as actors in the actor-network of the criminal justice system. As machine learning is integrated into more parts of society, it is essential that engineers explain how this technology works and take part in a dialogue on how it can be used safely.