CRYPTOGRAPHICALLY SECURE ENCRYPTION USING ADVERSARIAL NEURAL NETWORKS

THE PRESENCE OF UNINTENDED BIAS IN ARTIFICAL LEARNING USED IN THE HIRING PROCESS

An Undergraduate Thesis Portfolio
Presented to the Faculty of the
School of Engineering and Applied Science
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By

Nicholas Winans

May 9, 2022

SOCIOTECHNICAL SYNTHESIS

Machine learning is a big buzzword in computer science, with people seeing it as the solution to all problems or an overhyped paradigm destined to solve nothing significant. The computing promises of machine learning have led to its experimentation and implementation all across our digital lives, and made the understanding of the technology more important than ever. One such emerging application of artificial intelligence is in the cryptography industry as a solution to the risk that quantum computing poses to encryption. However, not every application of machine learning is theoretical. To deal with the increasing number of applications per open job position, many companies seeking employees have turned to machine learning models to help make decisions. But, these models can introduce and reinforce biases present in human decisions. By examining both the positive potential and possible shortcomings of machine learning models, we can begin to understand the complete picture of machine learning, which lies somewhere in the middle.

The computational power provided by quantum computing poses a threat to modern encryption, which relies on the difficulty for computers to solve certain mathematical problems. As the scale of quantum computing increases, researchers are looking into alternative encryption strategies, including looking in the field of artificial intelligence. The technical report looked into alternative methods that researchers have used to introduce machine learning into cryptography. It is a survey of the attempts and successes of these researchers, detailing their approaches.

The technical report focuses on two categories of applications, traditional encryption concepts and alternative methods (steganography). In the case of traditional cryptography, researchers applied the ideas of symmetric and asymmetric key encryption to machine learning, but importantly did not tell the models how to use these tools. The resulting models were able to successfully encrypt information so that it was hidden from an adversary, but recoverable by the intended recipient. In the steganography experiments, experimenters used neural networks to hide information in photographs. They were able to realize promising results, increasing the deception/hiding capabilities of a steganographic system as described in the technical report.

The STS Report focuses on the presence of bias in machine learning models used in the hiring industry, asking how it got there and why. To answer the questions of how bias enters models, a case study of Amazon's foray into using models to screen candidates is explained and analyzed. Law & Callon's Actor Network Theory is used to structure the analysis, including both human and non-human actors, as well as highlighting different motives and agency levels. A case study allows for a deeper understanding of one particular issue, and lends well to being combined with Actor Network Theory, which clearly lays out complex issues.

To analyze how biases were introduced into Amazon's model, the data pipeline was analyzed first. How Amazon obtained their data was a large part of where biases came from and is something future models can learn from. Additionally, there are other companies using machine learning in the industry, and Pymetrics is introduced as a comparison in how they conducted third-party bias investigations to create trust in their system. Government regulations in the hiring industry are brought up, specifically the Civil Rights Acts, and how the models fall in an area not explicitly illegal, but clearly immoral.

Machine learning is an exciting technology that promises to revolutionize how we use computers. It can free us from flaws of human thinking and create innovative solutions to existing problems, but it can also reinforce how humans behave without careful considerations. Before introducing models that interact with our livelihoods, it is paramount that people understand how artificial intelligence works to understand and account for its limitations.

TABLE OF CONTENTS