John Sikes
11/21/2020

Modern Ethics in Cybersecurity and Artificial Intelligence

As technology trends further and further out of the realm of the control that its creators exist in, the figurative laws of ethics and morality become curious and worth investigating. A poignant example of such is that of artificial intelligence and its role in cybersecurity - both now and in the near future. Through their conception as advanced algorithms by an aspiring programming and into their infancy as applications running to monitor and deter threats, the operating capacity of artificial intelligence in the field of cybersecurity is a challenging topic to tackle due to the good natured intentions of such poorly mixing with existing legal definitions and moral expectations.

The current status of artificial intelligence is already highly advanced and well deployed across various systems. For the most part, these AI protocols operate to deter and disrupt bots operating on malicious intent, for which the vast majority of people would agree no moral issue can be found in, as no side of the argument is being harmed. However, outside of bots sparring with one another is the realm of actual attackers seeking to disrupt a system, secure private information, or wreck any other havoc that they so seek. In this realm, a real person is on the receiving end of whatever response an artificial intelligence protocol is set to return, which could in turn bleed into the legal pitfalls of hacking-back that exist for human beings themselves seeking to disrupt their attackers in real-time. To this extent, not only is a legal issue incurred by indirectly breaking the law through this protocol attacking another person as compared to a robotic attack, but a moral line is crossed in the sense that the artificial intelligence itself isn't to blame for its functionality. This, however, causes further confusion with the morality of the situation as the protocol could have been developed outside of the system through which the attack was landed, thus causing the true blame for the situation to be ambiguously spread between either the developer or the deployer. In a similar manner, this can also be expanded to indicate any sort of successful attack that may elude an artificial intelligence system as well, either presenting the legal and moral liability with the developer for not having accounted for more edge cases or the deployer for not having proper safety nets outside of the anticipated range of the protocol.

Through this murkiness, it is quite evident that both legality and morality are difficult to extrapolate. Unfortunately, further issues with the morality of artificial intelligence usage can further be found by way of attackers deliberately working to undermine the machine learning that makes up the framework of an AI scheme. An example of this can be found in the case of the bot Tay, developed by Microsoft and deployed to Twitter, that was fed inflammatory statements to the point that its learning set became filled with such concepts and led the bot to become inflammatory in turn. In a much less trivial manner, this can then be extrapolated to indicate the dangers an AI scheme could face if it were presented with enough significantly incorrect input to detect, therefore leading its original purpose to become much less clear and allow attackers to then capitalize on such weaknesses. Expectedly, this brings into question the morality of permitting such a protocol to have dominion over sensitive resources, on par with that of trained professionals. As a result, the concept

of an AI capable of manipulation from seasoned attackers could sway would-be deployers from scrapping ideas of using such in order to maintain confidence in their existing security infrastructure, despite the clear advantage one would have by means of automating attack detection and resolution.

In a final indication of the moral complexity of utilizing an artificial intelligence system in a security framework, the usage of such a system could quite clearly be drawn to a lack of availability or outright loss of jobs. Through the deployment of an artificial intelligence system capable of automating and resolving issues on the deployed system, the need for more widespread human monitoring of a system is rendered irrelevant or, at the minimum, less valuable. As a result of this, employer's will find themselves with the ethical decision of laying off employees, cutting salaries, or offering less positions moving into the future due to the lack of necessity for such, as automation has taken over.

As a result of advancements in the realm of machine learning, and artificial intelligence as a byproduct, the capabilities of systems to automate the detection and resolution of cyber-attacks is significantly increasing. However, alongside such are increasing moral and legal dangers by means of dataset manipulation by attackers, loss of employment by those phased out by technology, and potential hack-back attempts by the artificial intelligence system deployed specifically to protect a system. To this extent, developers and their legal teams must remain vigilant of the development of their protocols, taking into account the actions and weaknesses of such, and to remain transparent with their deployers to cover all technical blind spots to properly protect the systems they need to.

Works Cited

@dannybradbury, Danny Bradbury Contributing WriterFollow. “How Can the Law Keep Up With Cyber-Attacks on AI?” Infosecurity Magazine, 11 Mar. 2020, www.infosecurity-magazine.com/infosec/how-can-the-law-keep-up-cyber/.

Palmer, Danny. “AI Is Changing Everything about Cybersecurity, for Better and for Worse. Here's What You Need to Know.” ZDNet, ZDNet, 2 Mar. 2020, www.zdnet.com/article/ai-is-changing-everything-about-cybersecurity-for-better-and-for-worse-heres-what-you-need-to-know/.

“Use Cases for AI and ML in Cyber Security.” Information Age, 8 Sept. 2020, www.information-age.com/use-cases-for-ai-ml-cyber-security-123491392/.

“War of the AI Algorithms: the next Evolution of Cyber Attacks.” Information Age, 30 Sept. 2020, www.information-age.com/war-ai-algorithms-next-evolution-cyber-attacks-123491934/.

Wolff, Josephine. “How to Improve Cybersecurity for Artificial Intelligence.” Brookings, Brookings, 8 June 2020, www.brookings.edu/research/how-to-improve-cybersecurity-for-artificial-intelligence/.