2024

seceon

CYBERSECURITY:

Machine Learning + Artificial Intelligence = Actionable Intelligence

by Smit Kadakia, Co-founder, Seceon Inc.

OVERVIEW:

The goal of AI (artificial intelligence) is to enable computers to do things normally done by people -- in particular, things associated with people acting intelligently. In the case of cybersecurity, its most practical application has been automating human-intensive tasks to keep pace with attackers! It is important to note that attackers are also using AI to be more effective in their attacks. Progressive organizations have begun using AI in cybersecurity applications to defend against attackers. However, on its own, AI is best designed to identify “what is wrong.” For today’s enterprise, that’s only half the challenge of defending against attackers. What today’s enterprise needs is to know not only “what is wrong” in the face of a breach, but to understand “why it’s wrong” and “how to fix it!.” By combining AI with ML (machine learning) it is now possible to analyze and interpret actions and behaviors to predict attack patterns and recommend actions to stop threats in motion; technologists have devised a means to generate actionable intelligence. This article will seek to explain how AI and ML can be used to a practical effect in defending the enterprise in real time. The age of AI/ML powered cybersecurity platforms is here.

Introduction:

Machine Learning (ML) and Artificial Intelligence (AI) have been around in one form or another since the latter half of last century, but in 2024 are rapidly advancing and growing the ability to change our day to day life. The data science which now includes both ML and AI, among other things, has taken off on its own to become a major discipline in the educational world. Business stakeholders have also recognized the importance of this discipline and have been leveraging contemporary data science methods to mine valuable information, make smarter business decisions and explore new opportunities using existing or newly harvested information.

The internet and mobile revolution of the last few decades have helped generate volumes of valuable information with potential to derive critical behavioral and structural insights from its collection. Naturally, the information gathering was vastly helped by the highly connected world of people, the devices that they use and the mechanisms that facilitate collaboration on both a professional and social level. However, along with such benefits comes the dark side of exposing this information, leaving it open to bad elements of society for unintended use, such as financial exploitation through ransomware and malware, two of the most prevalent types of cybersecurity attacks.

In the face of these attacks, technologists have begun developing a combination of cybersecurity defense techniques that rely on the collection of large volumes of real-time network, application and user interaction and behavioral data. This mix of data science techniques is the crux of how ML and AI disciplines can be leveraged in cybersecurity for proactively thwarting such attacks.

Andy Veluswami of Change.org expresses a visionary insight as, "*We're going to have a day, and I hope it's soon, where machines aren't just smart, but they're also wise – and they have a context. Once we start getting there, and we already are, we're going to start making a lot more progress.*" We all intuitively know that this change is happening all around us, however, the practical aspect of this development, such as translating learning into "*Actionable Intelligence,*" is a key requirement that today's cybersecurity practitioner must have.

So, how do we define what is and is not actionable intelligence in cybersecurity defense?

Machine Learning can be broadly defined as a focused approach of math and statistics-based algorithms that are designed to improve the performance of specific tasks through experience or learning that may or may not be easy to do by humans. On the other hand, Artificial Intelligence can be defined as a focused engineering approach for computing machines to do the tasks that we as people can do quite naturally, but conduct them without mistakes and, sometimes, much faster.

So, how do we define what is and is not actionable intelligence in cybersecurity defense?

In the study of Machine Learning, the focus is on supervised and unsupervised learning. (We will not be considering deep learning in this paper.) Supervised learning, and many aspects of unsupervised learning require the known anomalies to be available to learn from and then predict anomalies in test data using the trained models and then fine-tune them through techniques such as cross-validation. In cybersecurity, one is usually looking for an anomaly amid a huge amount of normal traffic or behavior. Such a characteristic makes anomaly detection a very difficult problem—like finding a needle in a haystack.

In April 2016, researchers from MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) demonstrated an artificial intelligence platform called AI2 that predicts cyber-attacks significantly better than existing systems by continuously incorporating input from human experts. The premise behind the finding is that it only needs an unsupervised suite of algorithms with feedback from expert security analysts to develop an AI-based algorithm that will detect the threats accurately. This too is a good approach, but these expert security analysts' services come at a significant cost, both in terms of time and money.

Furthermore, it is unrealistic to expect that training with anomalous data points in one industry, say eCommerce, is applicable to another one such as a healthcare data center. Additionally, modern attacks are more sophisticated and they hide themselves among many false attacks to defeat threat detection systems. Such complexity makes identifying anomalous training data points for all target industries a huge uphill battle. Lack of or difficulties in obtaining training data points make unsupervised learning a necessity in the world of cyber defense.

Given that unsupervised learning is required for such environments, one has to think through its pitfalls, recognizing that the prediction of the anomalies does not increase false positives or compromise accuracy. Various measures are used to assess the confusion matrix for accuracy, sensitivity, etc., such as the Matthews Correlation Coefficient, however, it is not easy to consistently get good measures from the matrix in practice, demonstrating that more than just ML is needed to get to the desired results. There are various approaches that one can take, but the end result has to be the actionable outcome from the algorithms with minimal noise. This is where AI comes into play.

Furthermore, many of the ML algorithms see the threats long after they have been introduced and the attacker has taken advantage of the vulnerability. For example, a clustering algorithm may detect a threat after analyzing a history of patterns and indicate that the anomaly that occurred has been introduced in the network or a subnet sometime back in history. These threat findings are useful once you deploy an army of security operations staff to hone in on the root cause for the anomaly and then address it. This sounds well and good, but the anomaly may already have spread by the time the security ops staff identified the root cause, requiring a much wider investigation with an increasing budget and delayed response time.

Genevieve Bell, Senior Fellow Vice President, Corporate Strategy Office, Corporate Sensing and Insights of Intel said in one of her recent presentations, "*AI is the next big wave in computing. Like major transformations before it, AI is poised to usher in a better world.*" The signs are all around to take us there.

Clearly, it is desirable to identify the threat that occurred in real-time as soon as it happens, as well as provide the specificity about where it occurred. Furthermore, the method of arriving at this conclusion should also be provided for the added benefit of understanding and quick action to address the root cause. Such actionable intelligence must reduce the skill set required to address the threat. Moreover, in a highly developed system, it must eliminate the need to engage a human, offering the intelligence just in time to prevent any further damage, reducing the time it takes for a corrective action or a good combination of all to minimize the operational cost while setting the organization ahead of the attacker's plans.

Inherently, such actionable intelligence must instill confidence in the user to prevent future attacks by learning from the attack and the response behavior. Real-time actionable intelligence should do more than simply help perform a quick analysis. It should also help the organization learn from the intelligence much more rapidly and thoroughly to develop a better defense against not-yet-seen attacks.

We operate in an era where such systems are now in development with enterprising startups, and some platforms like Seceon are leading the way with production deployments. The advent of big-data platforms and related technologies are making this all possible. These systems are expected to dominate the cyber defense efforts of many of the elite organizations worldwide and will be writing the next chapter in the forefront of cybersecurity.

About the Author

Smit Kadakia

Co-founder, Seceon Inc.

Smit leads Seceon's data science and machine learning team, focused on developing a state-of-the-art behavior anomaly detection solution. Smit holds a B.S. from VJTI, Mumbai, an MS in Computer Science from Indian Statistical Institute, Kolkata, and an MBA from Southern New Hampshire University, Manchester. Smit and the team at Seceon have built the industry's first and only fully automated threat detection and remediation system using a combination of machine learning and artificial intelligence techniques. Seceon's approach includes analysis of all traffic, flows, and processes in and out of the network and correlates them near-simultaneously with behavioral analytics, recognized and zero-day exploits, and policies to surface threats and proposed responses in near-record real-time.

Further Reading:

Seceon Resources:

Read our Buyer's Guide to SIEM platforms and learn how to determine how AI/ML should be a part of your selection criteria (SIEM Buyer's Guide).

Seceon Blog:

Read how MSPs and MSSPs are evolving their strategy in the age of AI/ML

https://www.seceon.com/msp-vs-mssp-is-there-a-distinction-anymoreread-the-blog/

About Seceon

About Seceon Seceon enables MSPs, MSSPs, and IT teams to reduce cyber threat risks and their security stack complexity while greatly improving their ability to detect and block threats, and breaches at scale. Seceon augments and automates MSP and MSSP security services with an AI and ML-powered aiSIEM and aiXDR platform. It delivers gapless coverage by collecting telemetry from logs, identity management, networks, endpoints, clouds, and applications. It's all enriched and analyzed in real-time with threat intelligence, AI and ML models built on behavioral analysis, and correlation engines to create reliable, transparent detections and alerts. Over 430 partners are reselling and/or running high-margin, efficient security services with automated cyber threat remediation and continuous compliance for over 7,800 clients.

Learn more about Seceon and

Schedule a Demo www.seceon.com/contact/