



www.endida.com

info@endida.com



How AI and Machine Learning Improve Enterprise Cybersecurity

Connecting all of the Dots
in a Complex Threat Landscape

A recent study¹ by Information Risk Management points out that 86 percent of enterprises will be impacted by AI (Artificial Intelligence) in the next five years. One of those impacts will be in cybersecurity defenses. As the volume of cyberattacks grows, security analysts have become overwhelmed. With more security-related data collected and more tools in the market, more alerts are generated every day, which is known as the “alert fatigue” problem. On the other hand, the industry lacks enough trained security analysts to process such alerts. AI and its component, Machine Learning (ML) can help automate laborious tasks like threat hunting, threat evaluation, and threat response, thereby relieving the burden on analysts who are awash in alerts.

AI uses algorithms to make decisions about what to do in given situations. The algorithms are a plan for what the computer should do given certain specific inputs, and the inputs come from data collected in the environment in which the AI system works. It tries to mimic human cognitive functions, where a computer is given the environment and the problem setting, and instructed to find a solution to the problem. ML uses data and prior algorithm responses to learn what to do in similar situations in the future. In autonomous cars, for example, the algorithms get trained by the information regarding the car’s surroundings and what maneuvering decisions are corrected, and then the algorithms know how to handle future situations that mirror those dealt with previously. Cybersecurity is a fertile field for applying AI and ML because there is a lot of data about the security environment, and many attacks are similar. With ML, the algorithms can learn how to respond to similar attacks, achieving the automation with AI. ML is thus a fundamental concept that drives AI.

Researchers have tried to implement AI and ML in cybersecurity solutions since the late 1980s, but progress has been slow. Today, AI and ML are showing increasing promise with the advent of Big Data because the quality of information that

informs AI or from which ML can learn is improving. However, there is much more to be done.

Anomaly Detection – The Early Days

When we talk about security, we want a system that can separate good from bad, normal from abnormal. Therefore, it is quite natural to apply anomaly detection to security. We can trace the beginning of anomaly detection back to 1987², when researchers started building intrusion detection systems (IDS). Around 1998-1999, DARPA (the government agency that created the Internet), created benchmark sets and called for research on ML methods in security³. Unfortunately, few of the results were practical enough and even fewer products got to the operational stage.

Anomaly detection is based on unsupervised learning, which is a type of self-organized learning that helps find previously unknown patterns in a data set without the use of pre-existing labels. In essence, a system based on unsupervised learning knows what is normal, and identifies anything abnormal as an anomaly. For example, an IDS might know what ‘normal’ traffic looks like, and it will alert on any traffic variants that don’t match that learned knowledge, such as a vulnerability scanner. Usually, the ML scientist needs to design ways to model the normal situation using ML models, and then train the model with a large amount of normal data. After that, the model is applied to label new cases; if the cases have a very small chance to fit with the normal model, the cases are reported as anomalies. In short, anomaly detection systems based on unsupervised learning make a decision (normal/abnormal). Some refer to unsupervised learning applications as ‘one-class problems.’

As you might imagine, systems based on unsupervised learning can generate a lot of false positives, because a situation deemed abnormal can be perfectly innocuous (think vulnerability scanner again). Anomaly Detection is good at detecting things that are different, but not

necessarily always malicious. This is a problem that security analysts still struggle with today.

The Rise of Big Data

After 2000, developers and researchers began creating anti-spam, anti-phishing, and URL filtering systems based on supervised learning. In supervised learning, decisions are learned from a set of labelled data. The labels specify the normal case and perceived threats. One such example is a URL classifier. First, large datasets of benign URLs and malicious URLs are used to train a URL classifier. Then, the URL classifier is deployed and matches the URLs extracted from incoming e-mails, and for each input URL, the URL classifier will produce a label to tell whether it is benign or malicious. A supervised learning algorithm analyzes the data and produces an inferred function (i.e., this traffic behavior matches this input data, therefore it is bad), which can be used for mapping new examples.

Early filtering systems using supervised learning were based on relatively small datasets, but datasets have grown in size and sophistication with the advent of Big Data. For example, Gmail employs an Internet-scale dataset to train its classifier to improve accuracy. Such large scale, using Big Data and huge computation resources, enables Google to train more sophisticated models.

Big models (in terms of number of parameters) based on Big Data, such as deep learning models, have gradually become more popular. For example, supervised ML has been successfully used in anti-virus signature generation for years, and in 2012, Cylance began offering next-generation anti-virus systems based on datasets other than signatures, such as anomalous traffic behavior. The advantage of ML-based methods compared with simple signature matching is that the ML-based method might find similar but not exactly identical cases, while signature matching looks for exact the same pattern.

Combining Supervised and Unsupervised Learning

Supervised learning has shown more success in security applications, but it requires easy access to large sets of labelled data, which are very difficult to generate for cyberattacks such as APT (advanced persistent threats) and zero-day attacks targeted at enterprises. Therefore, we cannot easily apply supervised ML to solve all cyberattacks.

This is where unsupervised learning comes back into the situation. We need to develop more advanced AI/ML that can be unsupervised or semi-supervised (e.g., through adaptive learning) to solve the additional cybersecurity challenges. Adaptive learning (human guided analysis) coupled with unsupervised learning improves your ability to detect those APTs and zero-day exploits. Unlike in supervised learning, where we need to label a large number of attack cases, adaptive learning leverages limited guidance from humans to drive the direction or preference of the anomaly detection to produce more accurate results.

A New Direction: Connecting the Dots

One of the big problems with simple anomaly detection is the volume of false positives. One way to address this issue is by correlating multiple events (dots) and then evaluating whether or not the correlation indicates a strong signal for a cyberattack. For example, one 'dot' might be an executive logging into the network at 2AM, and while this alone might be seen as a false positive, it wouldn't be enough to trigger an alert. However, if the executive is seen logging in at 2AM from an IP address in Russia or China, it would trigger an alert, because the executive seldom logs in at 2AM, and is also unlikely travel to Russia or China based on his schedule.

Developers and researchers are just now combining supervised and unsupervised learning into cybersecurity products. For example, Stellar Cyber's Starlight product correlates multiple events and evaluates whether, when looked at together, they constitute a threat. This approach significantly reduces false positives and helps analysts

identify APTs or zero-day attacks more quickly. The next frontier will be to make AI and ML self-teaching. This is called transfer learning in ML literature. The goal of transfer learning is to extract the more abstract knowledge in one domain, and apply it to address the other new domain better. Humans are good at transfer learning. For example, a guy who has mastered one kind of math can learn other kinds of math faster than other people. With a similar capability, past experiences in detecting and responding to threats will be factored into new evaluations of potential threats. The system would thus grow more accurate over time. Some security systems are beginning to implement self-teaching technology today, but the history

of AI and ML in cybersecurity is relatively short, and large-scale improvements will emerge in the future.

While security analysts will always be needed to make the ultimate decision about whether to kill a threat, AI and ML can make their jobs much easier if applied properly. For example, AI and ML can weed out false positive alerts by comparing data from previous false positives, and they can correlate seemingly isolated incidents to rapidly identify complex threats that might otherwise take days or weeks to discover.

[1] Information Risk Management, "Risky Business," 2019.

[2] D. E. Denning, "An Intrusion-Detection Model," IEEE Transactions on Software Engineering, vol. 13, no. 2, pp. 222–232, 1987.

[3] R. Lippmann, R. K. Cunningham, D. J. Fried, I. Graf, K. R. Kendall, S. E. Webster, and M. A. Zissman, "Results of the 1998 DARPA Offline Intrusion Detection Evaluation," in Proc. Recent Advances in Intrusion Detection, 1999.