WiNG

For Cybersecurity, Machine Learning Offers Hope Beyond The Hype

DECEMBER 15, 2015 MARTIN GILES SECURITY

The following entry first appeared in edited form as a blog post (http://gigaom.com/2015/12/11/for-cyber-security-machine-learning-offers-hope-beyond-the-hype/) *on GigaOm*

As businesses wind down for the holiday period, they'll need to keep their cyber defenses up. While executives are tucking into their dinners, hackers will be trying to tuck into their businesses' data. High profile breaches this year at organizations ranging from Anthem Healthcare to Ashley Madison and the US government's Office of Personnel Management are a stark reminder of the threats that lurk online. And they raise the question of whether the cyber security industry can come up with a powerful new tool to frustrate the bad guys.

There's been plenty of discussion at security conferences about the impact that machine learning will have on the cyber landscape. A subset of artificial intelligence, it involves the use of powerful algorithms that spot patterns and relationships from historical data, and get better over time at making predictions about brand new data sets based on this experience. Companies such as Amazon and Netflix use machine learning to help drive their recommendation engines, and banks and other financial institutions have long used it to tackle credit card fraud.

Now we are starting to see some cyber security firms offering solutions that involve a machine-learning component. Huntsman Security, which counts intelligence agencies amongst its clients, recently announced what it claims is the security industry's "first machine-based threat verification technology" that uses machine-learning algorithms to help analysts spot serious threats swiftly and take corrective action. Startups such as Cylance, Palerra and Darktrace are also employing machine-learning techniques in their services. (Disclosure: Wing Venture Capital is an investor in Palerra) (http://palerra.com/) .

A bullet point

It’s tempting to portray machine learning as a silver bullet that can be used not just to wipe out hackers, but also to wipe out jobs too by automating tasks performed by expensive personnel. This has provoked a backlash from some commentators, who have warned companies not to waste money on an unproven technology, and encouraged them to invest more in security teams and other tools instead.

However, that critique is based on a false claim about the technology’s potential—and a false dichotomy between human and machine.

Let’s take the issue of efficacy first. Machine-learning models work best when they can “train” on large volumes of data. Thanks to the rise of big data and extremely cheap storage, it’s now possible to feed vast amounts of information into models, which greatly improves their ability to detect suspicious activity. The goal is to distinguish anomalous behavior in things such as network traffic that might indicate a breach while minimizing false alerts (or “false positives” to use the industry’s terminology).

There are certainly challenges to be overcome. Algorithms are only as good as the quality and quantity of the data they are trained on, and data sets on the most sophisticated kinds of attacks mounted by nation-state actors (or their proxies) are still relatively thin. Sophisticated hackers can also try to fool models by employing tactics that seek to convince them that malicious activity is in fact legitimate.

Defensive play

In spite of such caveats, the machine-learning approach is still a great asset in a defensive arsenal. Given the volumes of data that security teams now have to deal with, adopting a more automated approach to querying network traffic and looking for anomalies that are not detected by traditional, signature-based systems makes sense. For instance, an analyst who has threat intelligence which suggests a network may be subject to a particular kind of data exfiltration attack could task a machine-learning model to look for telltale signs of this. Models can also provide analysts with other valuable insights, such as correlations between suspicious events.

To minimize false positives, many models rely not just on “unsupervised learning”, which involves crunching data to spot patterns themselves, but also on customer-driven, “supervised” learning.

This can take the form of specific security policies, such as one that requires an alert to be issued if a bunch of sensitive files are suddenly sent to a new location. It can also involve analysts to give a digital thumbs-up or thumbs-down to alerts issued. Over time, this training can help a model to identify what really matters to an organization and reduce the risk of false alerts.

Will human trainers ultimately be displaced by the “machines” they teach? Some companies may use machine-learning as an excuse to downsize, but I think they’ll be the exception rather than the rule. When I speak to chief information security officers, I often hear that they are concerned about a worrying shortage of skilled cyber personnel. By putting machine-learning models to work in support of existing staff, security leaders can boost productivity and free up their teams to work on the most pressing and strategic issues.

There is another consideration that might resonate at this time of year. Algorithms don’t need to take a holiday, so they can keep on working while some of their human masters are taking a well-deserved break!

---

http://wing.vc/blog/for-cyber-security-machine-learning-offers-hope-beyond-the-hype
© 1990 — 2018 Wing, Menlo Park, CA