SOPHOS
Security made simple.

Machine Learning for Cybersecurity, Demystified by Sophos

For a long time, defenders in cybersecurity have felt that the attacker had an asymmetric advantage. This is expressed in the saying "the defender has to be right every time, the attacker has to be right only once." Defenders have kept ahead of the wave, but it has been gaining on us. We seek to turn the tables, and recent advancements in machine learning will help us do that.

In February 2017 Sophos acquired machine learning security firm Invincea and has spent the months since then synthesizing their technology and brainpower into SophosLabs and our product line. Along the way, our data scientists have written several articles revealing the workings of machine learning—deep learning, in particular—how it will be brought to bear by Sophos, and how that will help secure our customers.

With that melding nearly complete, it's time to review where we've been and where we're going.

To date, our industry has mystified machine learning in secrecy and jargon. We prefer transparency, explanation, and information.

Please enjoy this guide – a collection of the articles we've shared throughout the year to tell the story.

There's more where this came from. We hope you find it useful.

--Joe Levy, Sophos CTO

Why machine learning?

When we first welcomed Invincea to Sophos, an explanation of what it would all mean for customers was in order. This was our opening message.

From there, we took our message to RSA Conference 2017:

Live from RSA Conference 2017: How machine-learning helps fight malware

Sophos product management director Russell Humphries talks about how machine learning will change the battle against malware.

How does it work?

Next, we wanted to go into as much depth as possible to explain how machine learning works and how we'd be applying it:

We're taking a quantum leap over traditional machine learning

Many security vendors already use machine learning. So, what makes our deep learning approach different and how much better does it perform?

5 questions to ask about machine learning

Machine learning isn't pixie dust to be spread on products. We look into the nuts, bolts and challenges involved, and how we approach it.

Demystifying deep learning: how Sophos builds machine learning models

An introduction to the process Sophos takes towards the development of a deep learning model.

Potential risks for machine learning

Next, we wanted to point out that it takes the right expertise to make machine learning work as intended. At Black Hat USA 2017 and BSidesLV, our data scientists gave talks about the risks and we wrote about them in Naked Security:

Garbage in, garbage out: a cautionary tale about machine learning

Here's the thing about machine learning: use the right datasets and it'll help you root out malware with great accuracy and efficiency. But the models are what they eat. Feed them a diet of questionable, biased data and it'll produce garbage. That's the message Sophos data scientist Hillary Sanders delivered in a talk called "Garbage in, Garbage Out: How Purportedly Great Machine Learning Models Can Be Screwed Up By Bad Data".

Where are the holes in machine learning – and can we fix them?

Like all good technological advances, the bad guys have the capacity to exploit machine learning algorithms. Sophos chief data scientist Joshua Saxe outlined the warnings and explained how Sophos would defend against such things.

For better machine-based malware analysis, add a slice of LIME

Sophos principal data scientist Richard Harang focused how Local Interpretable Model-Agnostic Explanations (LIME) can make machine learning results more accurate.

Defining the truth: how Sophos overcomes uncertain labels in machine learning

Uncertain labels are a challenge of applying machine learning to cybersecurity. We look at some of the ways Sophos overcomes this obstacle.

The death of us all?

We also sought to address a question that's been examined in both science and Hollywood: Will machine learning and artificial intelligence in general become too smart and someday turn on its creators?

Man vs machine: comparing artificial and biological neural networks

By comparing and contrasting biological learning to artificial intelligence, we can build a more secure infrastructure.

Why Artificial Intelligence isn't SkyNet in the making

Machine learning is a powerful tool, not a threat to our existence.

There will be many more articles to come, and this guide will be updated accordingly along the way.

United Kingdom and Worldwide Sales
Tel: +44 (0)8447 671131
Email: sales@sophos.com

North American Sales
Toll Free: 1-866-866-2802
Email: nasales@sophos.com

Australia and New Zealand Sales
Tel: +61 2 9409 9100
Email: sales@sophos.com.au

Asia Sales
Tel: +65 62244168
Email: salesasia@sophos.com

Oxford, UK