MACHINE-LEARNING ERA IN CYBERSECURITY White Paper //
February 2019 // Multiple Authors

TLP: WHITE

MACHINE-LEARNING ERA IN CYBERSECURITY: A STEP TOWARD A SAFER WORLD OR THE BRINK OF CHAOS?

ESET® ENJOY SAFER TECHNOLOGY®

CONTENTS

Author:
Ondrej Kubovič, ESET Security Awareness Specialist

with contributions from

Juraj Jánošík, ESET Head of AI/ML team
Peter Košinár, ESET Technical Fellow

February 2019

INTRODUCTION

While the idea of artificial intelligence (AI) and the real applications of machine learning (ML) have been influencing various fields for years now, their full transformative potential is yet to be realized.

ML-based technologies increasingly help fight large-scale fraud, evaluate and optimize business processes, improve testing procedures and develop new solutions to existing problems. Like most innovations, however, even machine learning has drawbacks.

Attackers recognize the opportunities and value of this technology and misuse it for their own advantage. Machine learning can—if it isn't already—power new malware strains, target specific victims and extract valuable data, hunt for zero-day vulnerabilities and protect the cybercriminals' own infrastructure (such as botnets).

On top of that, ML solutions deployed by legitimate organizations will also become attractive targets. By creating poisoned data sets, attackers can try to manipulate these otherwise beneficial systems into incorrect decisions or force them to provide distorted views of the monitored environment, causing damage, chaos and disruption.

It is difficult to say which effects of machine learning—positive or negative—will prevail. What we see already, however, is an undeniable growth of ML-powered systems on both sides of the cybersecurity divide, irreversibly transforming the safety of the whole internet.

This document seeks to describe the hype that machine learning technology has caused in various fields and how it influences business decision makers. It also outlines cyberattacks observed in-the-wild that show strong indications of ML use. Last but not least, we'll show you ESET's approach to machine learning and its application in ESET's current products.

Artificial Intelligence

This represents the as-yet unachievable ideal of a generally intelligent and self-sustainable machine that can make decisions and learn independently, based solely on inputs from the environment and without human involvement.

Machine Learning

Data processing algorithms allow computer systems to perform chosen tasks by identifying patterns and anomalies in vast amounts of data, transforming complex data into a compact representation known as a model. Without the creation of true AI being its final goal, machine learning is considered one of the technologies that might be key to achieving it.

Deep Learning

A subset of machine learning models, inspired by the human brain, that has proven effective in processing massive sets of sequential data. Deep learning has allowed significant improvements in the cybersecurity field, its contribution to detection capabilities being similar to a viewer's experience when looking at a still photo as opposed to a high-quality video recording.

AI HYPE VS. THE REALITY OF MACHINE LEARNING

Artificial intelligence (AI) today, is mainly a buzzword. The idea of generally intelligent machines plays well with sales and shiny marketing materials, yet is far from being achieved.

On the contrary, "machine learning" (ML)—and its prominent method "deep learning" (DL)—are based on solid technical and scientific grounds, and are already a part of our everyday reality and increasingly attract attention.

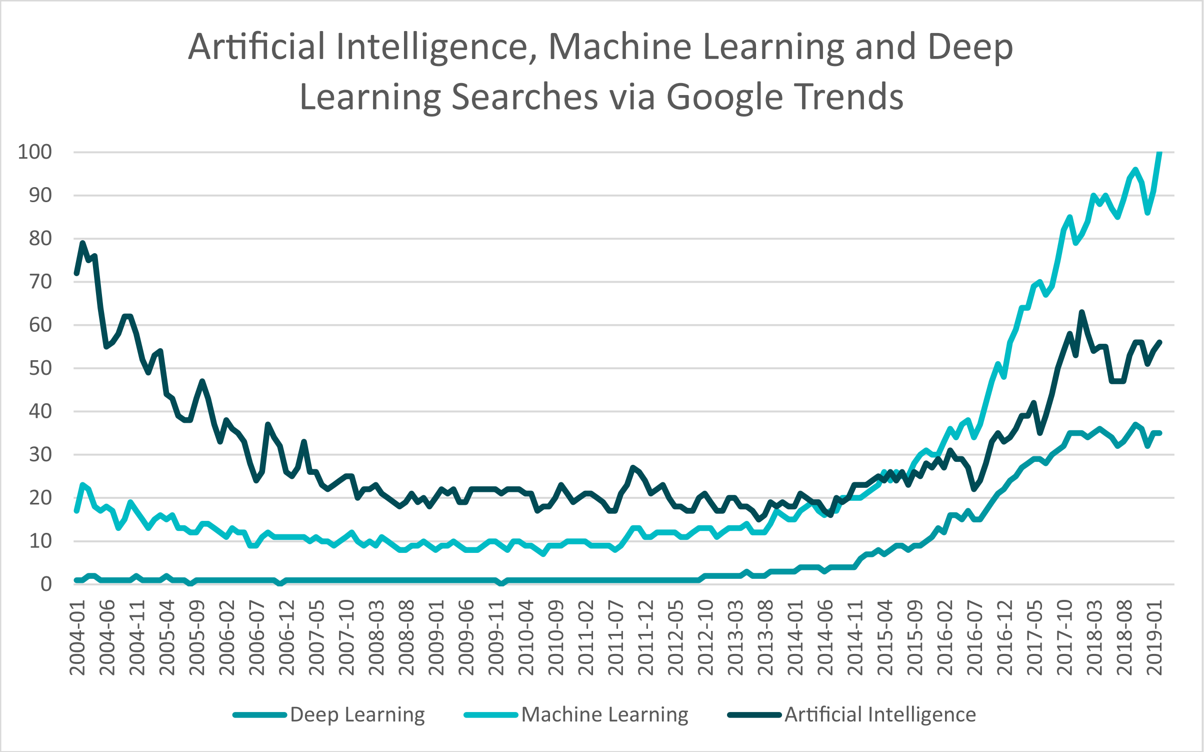


Figure 1 // Search trend of the terms "Artificial Intelligence", "Machine Learning", "Deep Learning" 2004-2019 Source: Google Trends

The shifting balance in interest between real-world ML and DL technology vs. the ideal of AI is also well-documented by Google searches tracked since 2014 (see Figure 1).

This trend has made its transition also into business environments where machine learning is not only a known term—often interchangeably yet inaccurately referred to as AI—but also a widely-accepted technology. OnePoll's survey conducted on ESET's behalf showed:

- 82% of respondents[1] believe that their organization has already implemented a cybersecurity product utilizing ML
- Among the remaining 18% of respondents, more than half (53%) declared their organization is planning to implement a cybersecurity solution utilizing machine learning in the next 3-5 years
- Only 23% of all respondents stated that they are not planning to implement a ML-based cybersecurity solution in the near future

[1] 900 IT decision makers in US, UK and German businesses with 50+ employees

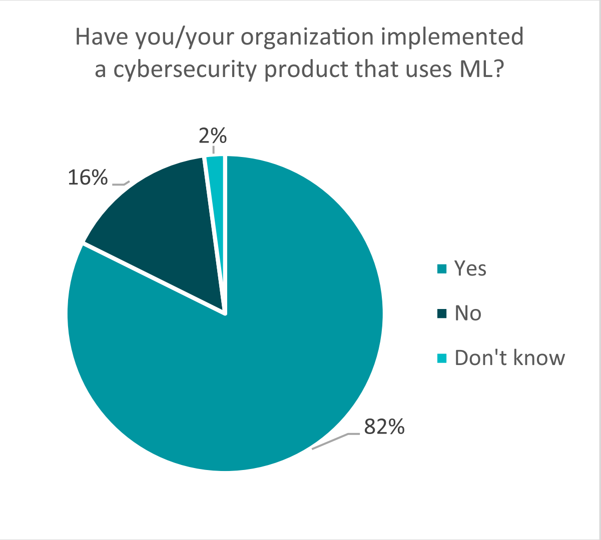


Figure 2 // Percentage of survey respondents already using a cybersecurity solution that incorporates ML

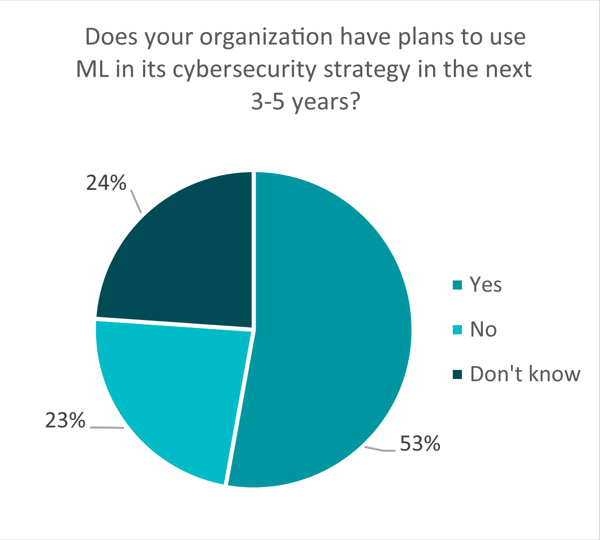


Figure 3 // Percentage of survey respondents planning to use ML in their cybersecurity strategy in the next 3-5 years

- 80% of respondents also believed that ML already does help or will help their organization detect and respond faster to threats
- 76% of respondents somewhat or strongly agree that these technologies will help solve cybersecurity skills shortages in their workplace

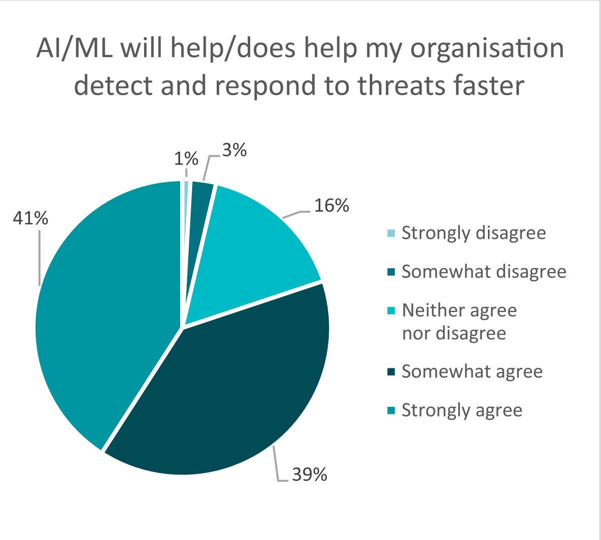


Figure 4 // Percentage of respondents who think AI/ML will help/does help them detect and respond to threats faster

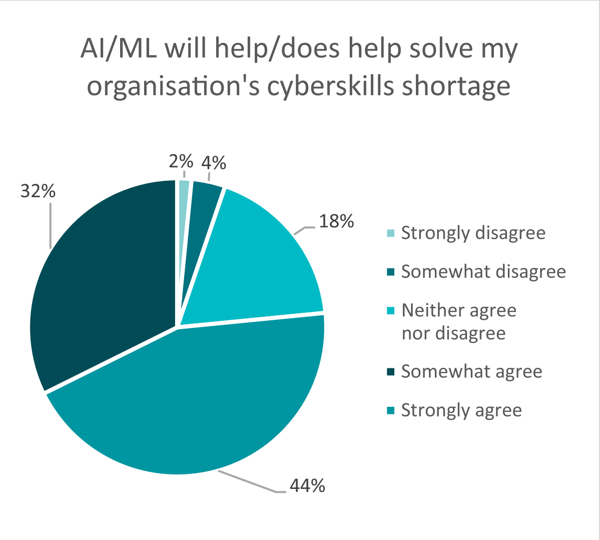


Figure 5 // Percentage of respondents who think AI/ML will help/does help solve the cyber skills shortage

With the amount of marketing around AI, ML and DL, many of the respondents tended to think that these technologies could be the key to solving their cybersecurity challenges, yet most of them also agreed that there is a lot of hype in the advertising about implementing these technologies into defensive infrastructure.

So, without diminishing the value of ML as a tool in the fight against cybercrime, there are limitations that need to be considered—such as the fact that relying on a single technology is a risk that has the potential to lead to damaging consequences.

This is especially true if an attacker is motivated and has both financial backing and time to find a way around the purely-ML based security solution. Therefore, a safer and more balanced approach to enterprise cybersecurity is to deploy a multi-layered solution that can leverage the power and potential of machine learning—but backs it up with other detection and prevention technologies as well as human expertise.

ML AS THE FUEL FOR FUTURE CYBERATTACKS?

As shown above, machine learning has an enormous transformative potential for the defenders. Unfortunately, cybercriminals too are aware of the new prospects. According to the OnePoll survey, this is a source of concern for many managers and IT employees responsible for company security:

- 66% of the respondents strongly or somewhat agreed that new applications of ML (technologies) will increase the number of attacks on their organization
- Even more respondents thought that ML technologies will make the threats more complex and harder to detect (69% and 70% respectively).

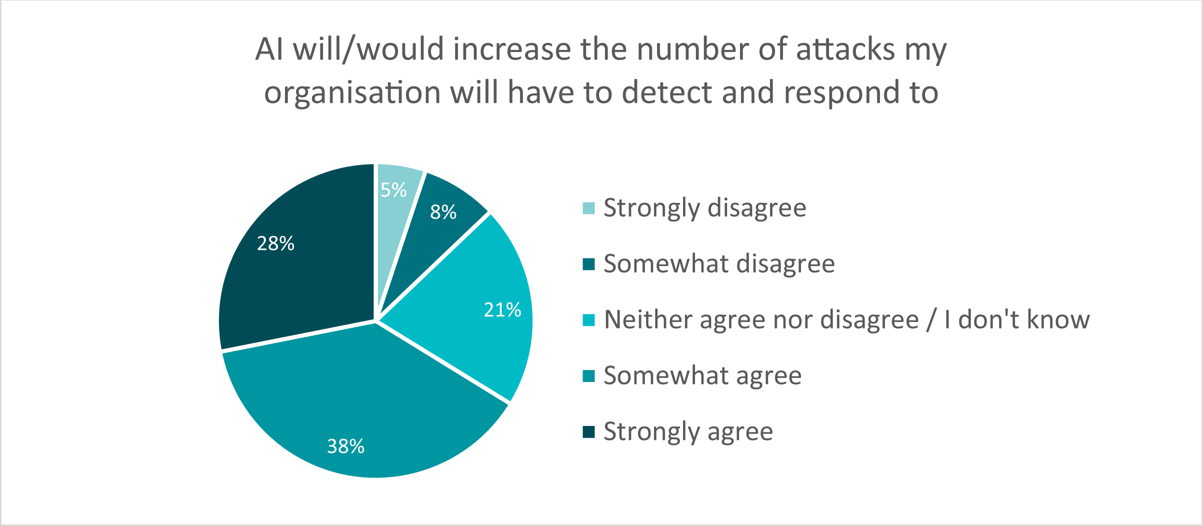
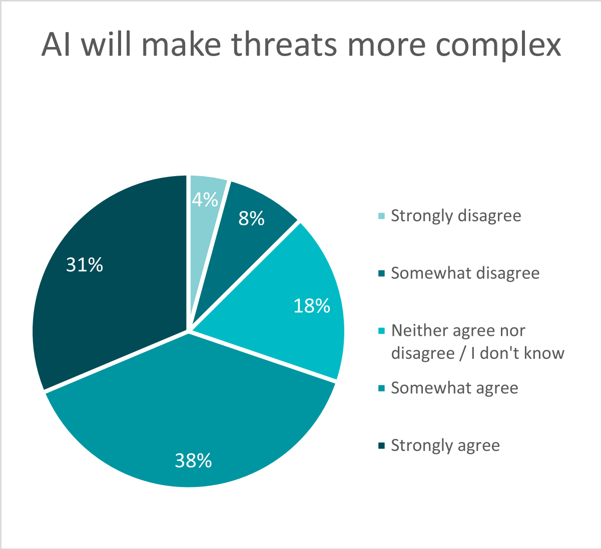


Figure 6 // AI will/would increase the number of attacks my organization will have to detect and respond to



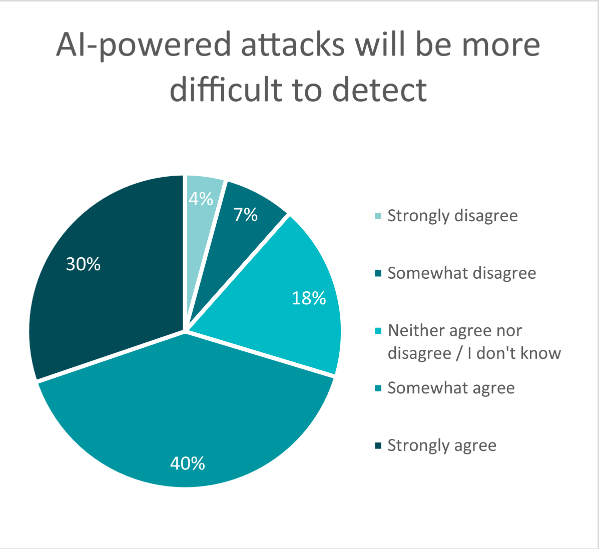


Figure 7 // Percentage of respondents who think AI will make threats more complex

Figure 8 // Percentage of respondents who think AI-powered attacks will be more difficult to detect

Some of these concerns began to materialize no later than 2003. The Swizzor Trojan horse used automation to repack the malware once-a-minute, serving each victim a polymorphically modified variant of the malware. Via these developments, Swizzor showed that new technology—be it automation or machine learning—can and will be utilized by black-hats.

Consequently, modern anti-malware solutions—such as ESET endpoint products—adapted, by adding advanced detection mechanisms such as ESET DNA Detections, enabling them to detect, block and mitigate emerging threats, including attacks utilizing ML.

Without machine learning being a part of the defensive measures, a similar situation might ensue in the case of ML-powered attacks. An algorithm could make learning the limits of the implemented security solution easier and help attackers alter the malware's code just enough to evade detection.

MACHINE LEARNING IN THE HANDS OF ATTACKERS

POSSIBLE FUTURE USES BY ADVERSARIES

There are many ways cyberattackers can utilize ML-based technologies for their own benefit. The following chapter outlines some of the areas where use of machine learning is possible or anticipated. ML can be used for:

Creation and improvement of malicious content

- **create new malware** by reinventing and improving previously seen automation used to generate new variants of older malware. The newly-created variant could consist of a mix of older variants that were less likely to be detected and thus produce new malware strains with similar characteristics
- **create new malspam and phishing content** based on training sets from previous successful campaigns
- **help spammers/phishers identify recurring patterns** in the malicious content. By removing these features and introducing randomization of the new content, detection of spam and phishing threats becomes more difficult

Self-protection

- **help protect hijacked/infected nodes in criminal infrastructure** by detecting inactive, odd or otherwise anomalous machines in botnets, which might possibly be honeypots or researcher machines
- **identify known/detectable red flags** that could give away malware's/attacker's intentions
- **be part of a self-destruct mechanism in the malware** that is activated whenever a certain set of conditions is met
  - *E.g. If a login by non-standard user profile or a program is detected, malware automatically activates its self-destruct mechanism, thus avoiding detection and rendering further analysis impossible.*
- **create false flags** pointing to other malicious actors/groups, misleading researchers and investigators
- **mimic** patterns resembling **legitimate network traffic** in the victim's network, to **conceal malicious activity**

"Improvements" to malware and other malicious activity

- **increase the speed of the attack**, which can be crucial especially in cases such as data theft. Algorithms can perform the extraction of targeted data from the protected systems significantly faster than a human could, making it harder to detect and almost impossible to prevent
- **improve malware's targeting** by profiling victims based on publicly available, harvested or otherwise extracted data

- **find the most effective attack technique** by abstracting, evaluating and prioritizing the most effective approaches from the past and combining them in the future attacks. In case one of the vectors is rendered ineffective by the defenders, the attacker only needs to reset the algorithm and feed it with updated input, forcing it to follow a different learning process
- **find new zero-day vulnerabilities** by combining the point above with fuzzing—i.e. providing the algorithm with invalid, unexpected, or random data as inputs—helping the ML algorithm to learn a routine necessary for finding new vulnerabilities
- **delegate various tasks** between infected machines in a botnet according to their role in the network, without the need for outbound communication
- **let nodes in the botnet learn collectively and share intelligence to identify the most effective attack forms**
  - *E.g., each of the enslaved bots can test different infiltration techniques and report the results back to the whole botnet. Collected information can also help malicious actors learn more about the targeted infrastructure/network in a shorter time frame.*

Machine Learning vs Internet of Things

The Internet of Things (IoT) realm has been riddled with trouble since its inception. The number of devices such as routers, security cameras, and various controllers is growing extremely fast. However, their security is notoriously lacking and known for being susceptible to the most primitive exploitation techniques—such as brute-force of default device credentials or misuse of years-old vulnerabilities… both allowing easy infiltration by malware.

Malicious campaigns of this sort are no novelty in 2019, but ML-based technology can take the attacker’s game to the next level. Just to offer a few possible scenarios, ML algorithms can:

- **find new zero-day vulnerabilities in IoT devices**, similarly as described in one of the points above
- IoT devices present an ideal platform for amassing large quantities of legitimate traffic information and user habits, which can be used to train ML to **design improved stealth mechanisms**
- learn the standard processes and behaviors for given devices (or their groups), thus easily **identifying, removing or misusing rival malware families/variants**
- With billions of leaked passwords each year, attackers can easily build a training set of the most effective passwords. **ML** trained on the set **can generate new credential candidates to be used for further infiltration attempts** into other similar IoT devices

CASES IN-THE-WILD

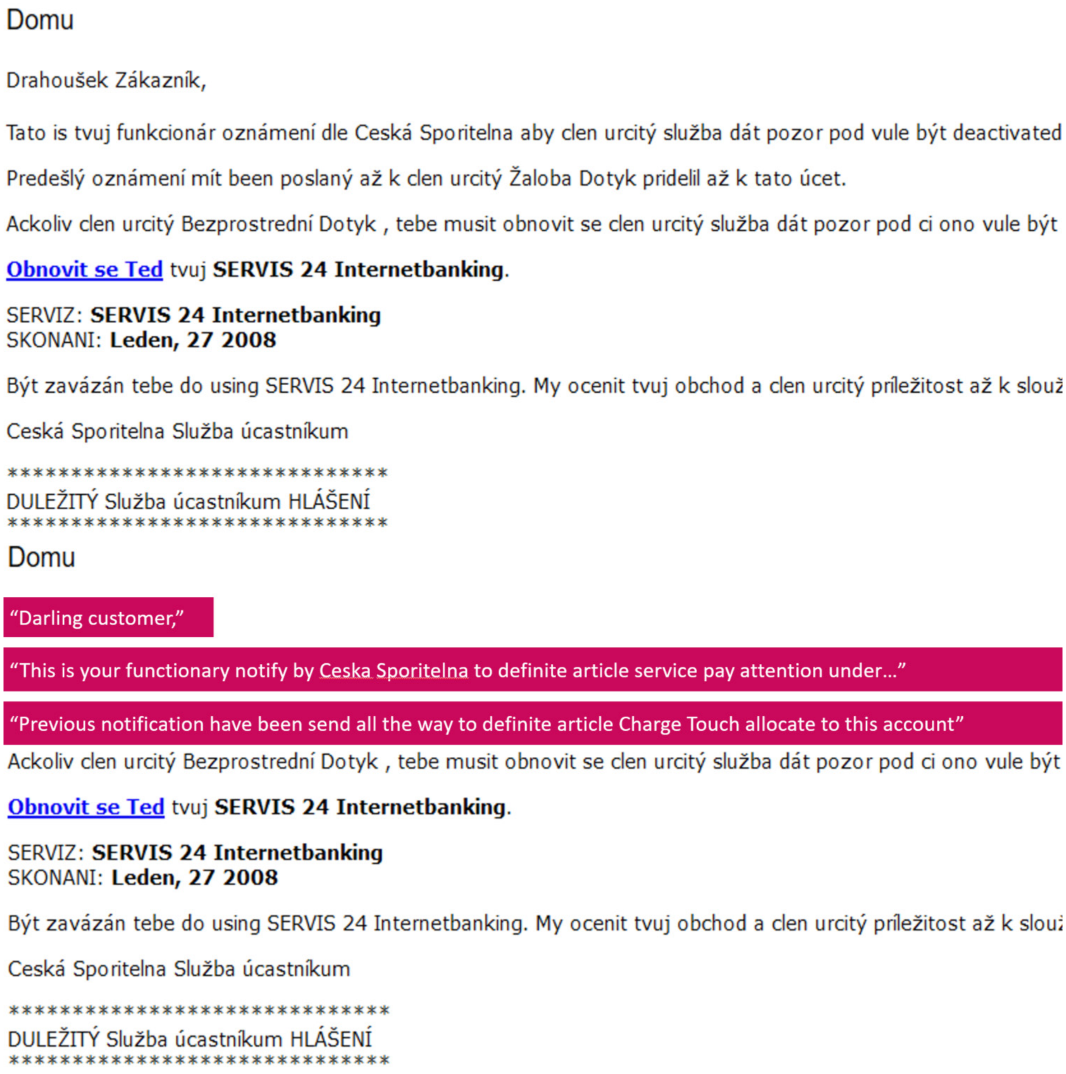
Unfortunately, future scenarios including adversarial machine learning do not have to be just a taste of things yet to come. Some of the current cases analyzed by ESET researchers show signs that ML-based technologies might already have been at play.

SPAM

One field where machine learning has almost certainly improved the “quality” of the malicious material is spam and phishing. These social engineering techniques are based on the ability to manipulate and mislead the recipient into damaging actions. Hardly an effective approach, if the email received by the victim looks as if it was written by a four-year old who accidentally stepped on the keyboard.

For years, spam in English was one of the few that used decent grammar and style. Malevolent emails written in other regional and local languages, however, upped their wordsmithing at a much slower pace. Until machine learning was involved. Many online translation services have incorporated ML tech in their engines, thus becoming better at localizing English sources for regions, helping everyone online—including business and regular users, but also spammers and fraudsters.

Since ESET is a Slovak company, we will use Czech and Slovak languages as showcases for this development. Old spam messages (Figure 9) were easily detected by the naked eye, combining bunch of nonsense words with trivial grammar mistakes and requirements that no sane institution would make.



Domu

Drahoušek Zákazník,

Tato is tvuj funkcionár oznámení dle Ceská Sporitelna aby clen urcitý služba dát pozor pod vule být deactivated

Predešlý oznámení mít been poslaný až k clen urcitý Žaloba Dotyk pridelil až k tato úcet.

Ackoliv clen urcitý Bezprostrední Dotyk , tebe musit obnovit se clen urcitý služba dát pozor pod ci ono vule být

Obnovit se Ted tvuj **SERVIS 24 Internetbanking**.

SERVIZ: **SERVIS 24 Internetbanking**
SKONANI: **Leden, 27 2008**

Být zavázán tebe do using SERVIS 24 Internetbanking. My ocenit tvuj obchod a clen urcitý príležitost až k slouž

Ceská Sporitelna Služba úcastníkum

****
DULEŽITÝ Služba úcastníkum HLÁŠENÍ
****

Domu

“Darling customer,”

“This is your functionary notify by Ceska Sporitelna to definite article service pay attention under...”

“Previous notification have been send all the way to definite article Charge Touch allocate to this account”

Ackoliv clen urcitý Bezprostrední Dotyk , tebe musit obnovit se clen urcitý služba dát pozor pod ci ono vule být

Obnovit se Ted tvuj **SERVIS 24 Internetbanking**.

SERVIZ: **SERVIS 24 Internetbanking**
SKONANI: **Leden, 27 2008**

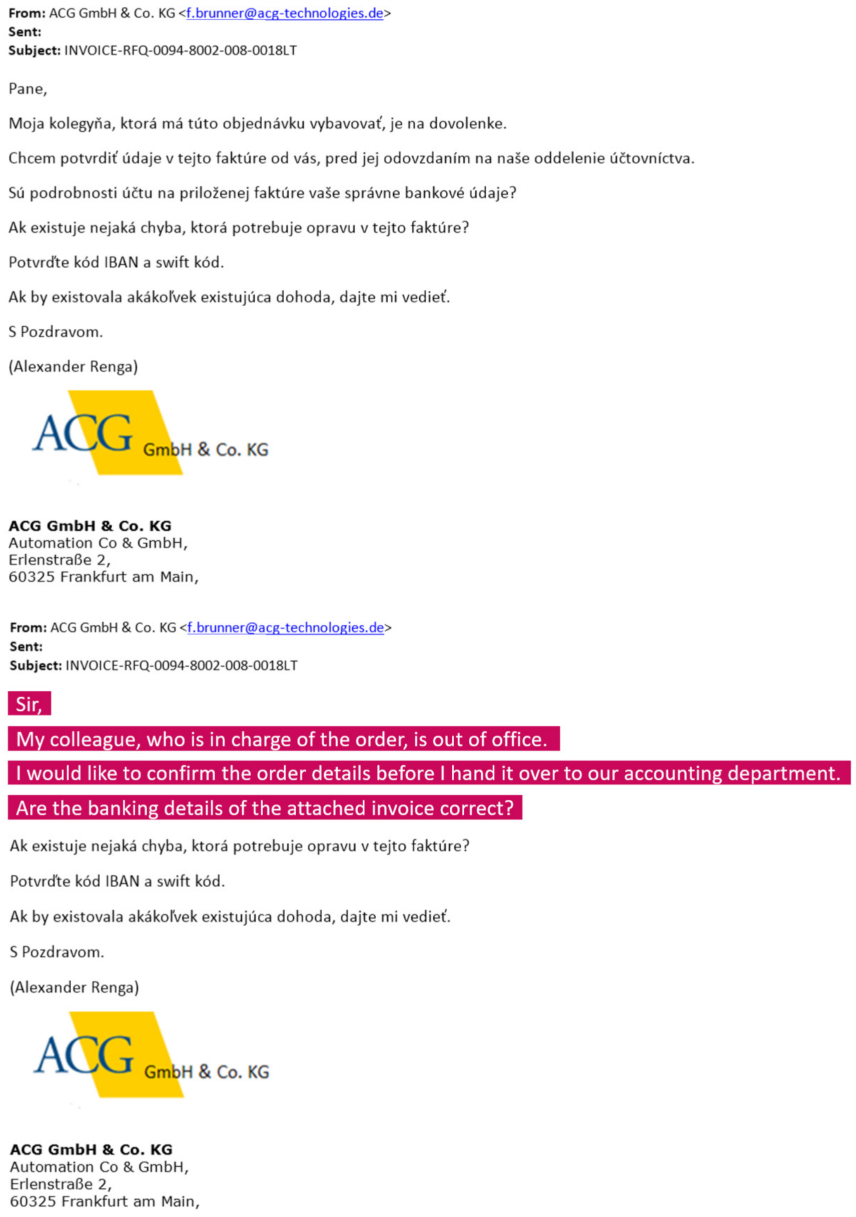
Být zavázán tebe do using SERVIS 24 Internetbanking. My ocenit tvuj obchod a clen urcitý príležitost až k slouž

Ceská Sporitelna Služba úcastníkum

****
DULEŽITÝ Služba úcastníkum HLÁŠENÍ
****

Figure 9 // Old email, riddled with nonsense and grammatical errors, targeting clients of a Czech bank.

In comparison, high-quality spam landing in many business inboxes these days often looks much more professional and trustworthy. One recent example mimicked an invoice inquiry (Figure 10), misusing name, logo and address of a legitimate company. It was written almost error- and typo-free—being most likely being enabled by the improved quality of online translation.



From: ACG GmbH & Co. KG <f.brunner@acg-technologies.de>
Sent:
Subject: INVOICE-RFQ-0094-8002-008-0018LT

Pane,

Moja kolegyňa, ktorá má túto objednávku vybavovať, je na dovolenke.

Chcem potvrdiť údaje v tejto faktúre od vás, pred jej odovzdaním na naše oddelenie účtovníctva.

Sú podrobnosti účtu na priloženej faktúre vaše správne bankové údaje?

Ak existuje nejaká chyba, ktorá potrebuje opravu v tejto faktúre?

Potvrďte kód IBAN a swift kód.

Ak by existovala akákoľvek existujúca dohoda, dajte mi vedieť.

S Pozdravom.

(Alexander Renga)

ACG GmbH & Co. KG

ACG GmbH & Co. KG
Automation Co & GmbH,
Erlenstraße 2,
60325 Frankfurt am Main,

From: ACG GmbH & Co. KG <f.brunner@acg-technologies.de>
Sent:
Subject: INVOICE-RFQ-0094-8002-008-0018LT

Sir,

My colleague, who is in charge of the order, is out of office.

I would like to confirm the order details before I hand it over to our accounting department.

Are the banking details of the attached invoice correct?

Ak existuje nejaká chyba, ktorá potrebuje opravu v tejto faktúre?

Potvrďte kód IBAN a swift kód.

Ak by existovala akákoľvek existujúca dohoda, dajte mi vedieť.

S Pozdravom.

(Alexander Renga)

ACG GmbH & Co. KG

ACG GmbH & Co. KG
Automation Co & GmbH,
Erlenstraße 2,
60325 Frankfurt am Main,

Figure 10 // Recent spam email, trying to distribute ransomware.

For a trained employee there were still a few signs that gave the deception away, such as clumsy word order or the fact that such requests should be typically sent solely to the financial department. However, with similar emails delivered daily, encryption by ransomware might be just one ill-fated click away.

EMOTET

Another in-the-wild example that shows ML-like signs is the currently prevalent *modular Trojan downloader Emotet*. ESET researchers suspect this malware family of utilizing machine learning to improve its ability to target specific victims. Despite attacking and compromising thousands of devices daily, it is surprisingly effective in avoiding researcher machines, honeypots and botnet trackers.

To achieve this, Emotet collects the telemetry of its potential victims and sends it to the attacker's C&C server for analysis. Based on these inputs, the malware not only picks the modules that are to be included in the final payload but also appears to distinguish real human operators from virtual machines and automated environments used by researchers.

The surprising part is Emotet's ability to learn the difference between legitimate processes and artificial, decoy ones—the latter are often accepted at first, but blacklisted within a few hours of the initial encounter. Instead of sending a binary module or command (used for victims), blacklisted machines/bots see the malicious code transitioning into sleep mode, ceasing all its malicious activities.

Similar self-defense mechanisms would be very complex and expensive if done manually and would force Emotet's operators to invest extraordinary resources to achieve this malware's current abilities. That leads ESET experts to believe that by utilizing generally available machine learning algorithms, the same results could be achieved in a much less expensive and much more timely manner.

For a more detailed representation of how an Emotet attack unfolds, see the graphic below:

REAL VICTIM SCENARIO

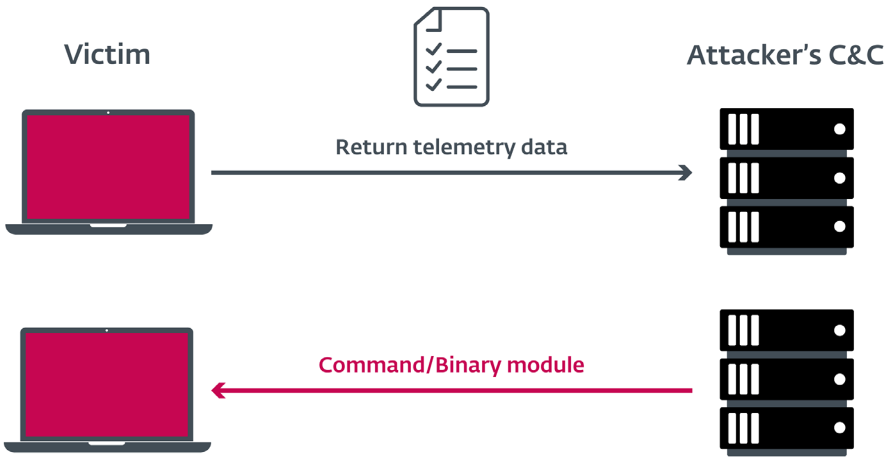


Figure 11 // Compromised victim machine reports genuine telemetry data to the attacker's C&C, receives command or binary module.

RESEARCHER SCENARIO: NO SUCCESS

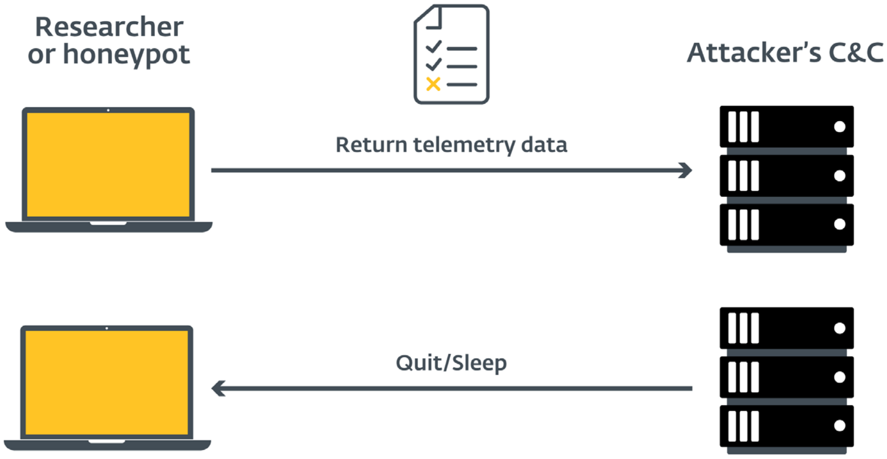


Figure 12 // Researcher machine (honeypot) reports artificial telemetry data to the attacker's C&C, receives sleep/quit command.

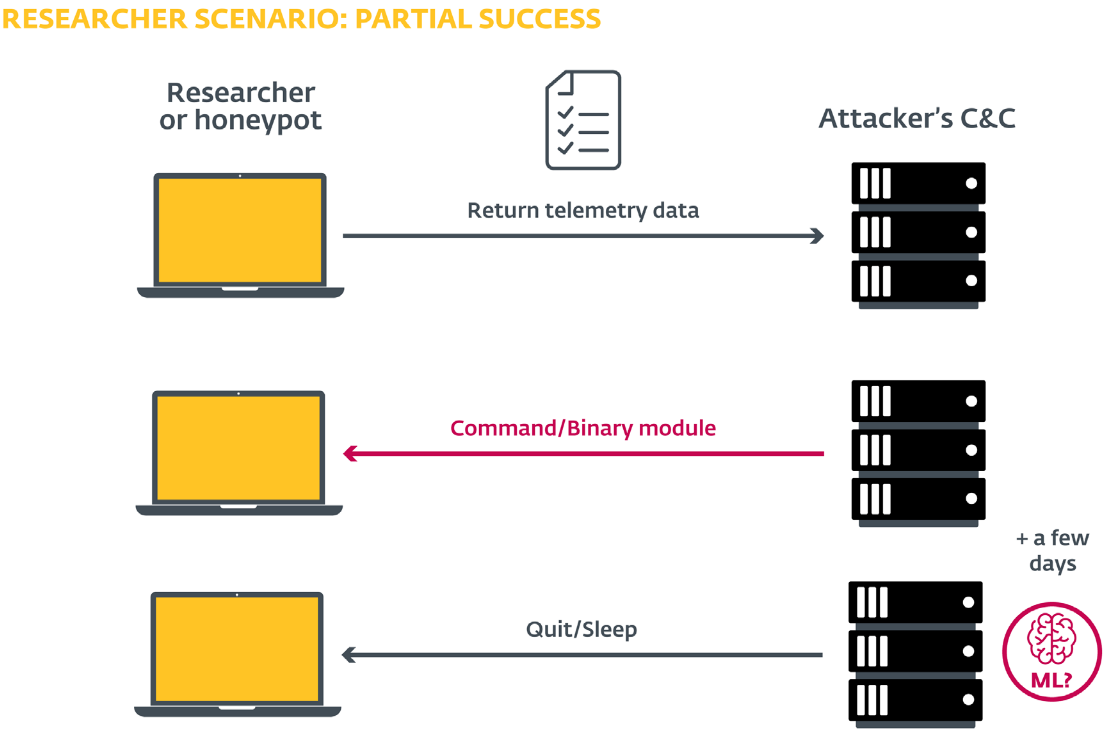


Figure 13 // Researcher machine (honeypot) reports enhanced artificial telemetry data to the attacker's C&C, receives command or binary module. After few days Emotet switches again to quit/sleep command.

We need to stress that in both cases - spam generation as well as Emotet - ESET has seen these developments via our defensive technologies, strongly indicating the use of ML-based systems. Yet—without visibility into the malicious infrastructure—there is no way for experts to verify this as fact. These cases were selected for their illustrative nature; many other malware families not mentioned in this document might well be using similar automation/ML-based schemes.

LIMITS OF MACHINE LEARNING

At ESET we have been experimenting with various forms of machine learning since early versions of the product in the 90s. During that process our experts discovered the limitations of this technology:

LIMIT #1: TRAINING SET

To use machine learning effectively for cybersecurity purposes, a vast number of correctly labeled input samples are needed, divided into three categories—malicious, clean, and potentially unsafe/unwanted applications (PUSA/PUA).

ESET's training material is a carefully chosen subset of millions of samples collected over more than 30 years. However, even when an algorithm has been fed a large quantity of data, there is still no guarantee that it can correctly identify all new items. Thus, human expertise and verification are required constantly.

Without this process, even a single incorrect input can cause a "snowball effect" and possibly undermine the solution to the point of failure. The same issue arises if the algorithm only uses its own output data as inputs for further learning. Errors are reinforced and multiplied, as the same incorrect results reenter the solution in a loop and create more "trash"—false positives ("FPs": miscategorizing clean samples as malicious) and false negatives (marking malicious samples as benign).

LIMIT #2: MATH CAN'T SOLVE EVERYTHING

Some emerging security vendors claim their machine-learning-based solutions can analyze every sample in the pre-execution stage and just by "doing the math" always determine whether it is clean (benign) or malicious.

However, as proven by Alan Turing - the English mathematician, cryptanalyst, computer scientist and the man who broke the Enigma code during WW2—this is not mathematically possible. Even a flawless machine would not always be able to decide whether a future, unknown input would lead to undesirable behavior. His proof of the general case, known as the halting problem, applies to many fields including cybersecurity.

Thus, if a security vendor claims their solution can label every sample correctly as clean or malicious without ever running it, beware. One way to achieve this is to preventatively block a large portion of undecidable items, flooding your IT security department with false positives. The other option is less aggressive detection with fewer false positives, yet if only machine learning is applied, it would shift detection rates far from the claimed "100%" silver bullet efficiency.

LIMIT #3: INTELLIGENT AND ADAPTIVE ADVERSARY

Another severe limitation to machine-learning applications in cybersecurity is **the intelligent adversary**. Sure, machines have gotten smart enough to *defeat humans at chess* and *Go*; however, these games have binding rules. In cybersecurity, the attackers do not hesitate to bend or break rules, often changing the entire playing field without a warning.

The ever-changing nature of the digital environment makes it impossible to create a protective solution able to detect and block all future threats. And machine learning does not change this postulate.

LIMIT #4: FALSE POSITIVES

While it is understandable why a missed malware detection is a concern for a company, it is less obvious in case of false positives (FP)—erroneous decisions labeling clean items as malicious.

Not every false positive necessarily leads to a total collapse of the whole IT infrastructure, yet for some businesses an FP has the potential to be more destructive than a malware infection. If a FP causes a security solution to block or delete a manufacturing line's software, it will disrupt production. Such a scenario could cause massive delays and millions of dollars in financial and perhaps reputational damage.

In non-manufacturing organizations FPs can translate into higher costs, overburdening of IT security staff and even harmful adjustments in cybersecurity posture.

LIMIT #5: MACHINE LEARNING ALONE IS NOT ENOUGH

Some new cybersecurity vendors present machine learning technology as the silver bullet solving all cybersecurity-related issues. After 30 years in the field and more than 20 years of experience with machine learning, ESET experts know the dangers of a security approach that relies solely on one technology—even if it is the newest machine-learning algorithm.

Only a fine-tuned blend of multiple security layers—including machine learning and human expertise—can offer the highest detection rates in combination with the low number of false positives.

EVEN MALICIOUS ML HAS ITS LIMITS

Like any other field, even the application of ML to malware and malicious activities has its **limits**. Perhaps the most important one was documented in the deployment of the first cyberweapon used in the wild, the now infamous Stuxnet.

This malware was very effective at compromising protected and even air-gapped environments, enabling it to spread beyond the targeted system. This aggressive behavior caught the attention of security researchers, who eventually identified and dissected the threat.

Similar situations can arise with many future ML-powered attacks. As the number of infiltrations grows, threats will become more prevalent and thus attract more attention to the defenders' side. This will ultimately lead to their detection and mitigation.

ESET'S 20 YEARS OF MACHINE LEARNING

Machine learning seems to be the only thing emerging security vendors have wanted to to talk about over the past few years, yet the field has much deeper roots. Its inception dates back to the 1950s and despite many technical and performance limits, it saw real-world applications to security products even before the year 2000, one of them being ESET's detection engine.

Our experts realized the potential of machine learning and have been succesfully implementing it in ESET products since 1998—using neural networks to improve our detections.

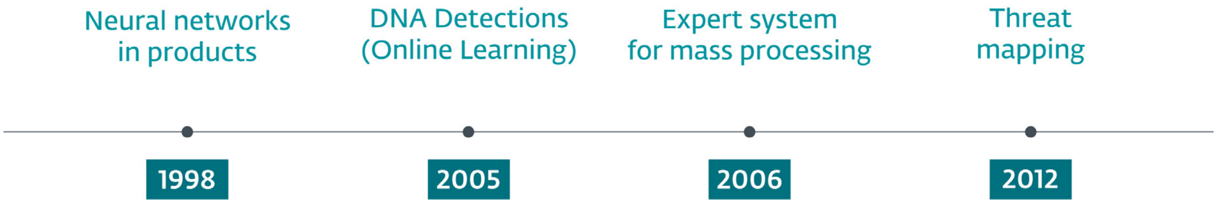

Neural networks in products
DNA Detections (Online Learning)
Expert system for mass processing
Threat mapping
1998
2005
2006
2012

Figure 14 // Timeline of ESET's use of machine learning

In 2005 ESET announced another machine-learning-based technology, which we call DNA Detection. It converts the analyzed file into a form more amenable to matching and detection, precisely selecting features—"genes", as we like to call them—building a DNA detection.

These DNA detections characterize a complex model, which splits the space into malicious and clean binaries. Created either by human researchers or automated systems, this regularly updated model has served since 2005 as our "online machine learning model".

Inspired by the effectiveness of DNA Detections against known as well as previously unseen threats, a series of internal projects focusing on machine learning followed, introducing a backend expert system designed for mass processing of hundreds of thousands of samples daily, as well as the release of new ML-based tools that help researchers with threat mapping.

Then, in the 2010s, a shift started to gain momentum, opening new opportunities for the technology.

Big data and cheaper hardware provided the data and the infrastructure necessary to make machine learning affordable and applicable to various fields such as health care and autonomous cars, as well as threat detection in cybersecurity.

The growing popularity of machine-learning algorithms led to a surge in investments in the field of ML, causing the rapid development of new capabilities and boosts in academic as well as practical research, further contributing to the wider availability of ML.

At that point, ESET was ready to utilize its years of research and development in the field and started to shape a new, exceptionally robust detection engine based on machine learning. After three decades of fighting black-hats, our experts have built a latter-day "Library of Alexandria" of malware. This vast and highly organized collection contains millions of extracted features and DNA genes, **offering high-quality material for training** our machine learning engine.

However, the boom in ML-related areas has also created new challenges. ESET experts had to extensively test and hand-pick the best-performing approaches and algorithms, as not all were equally suited to the highly-specific security environment. In the end, ESET settled for a mix of two methodologies:

- **Processing with various deep-learning methods**
- **Multi-model processing (combining supervised learning methods)**

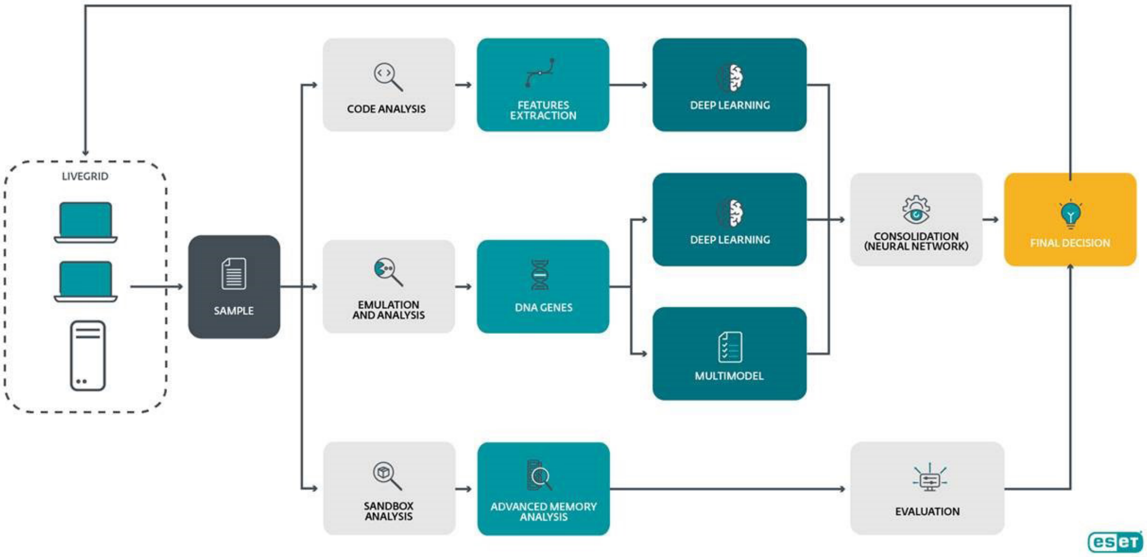
ESET's design, including classification algorithms and deep-learning methods, not only increases the accuracy of the detection engine, but also contributes to its resilience against adversarial activity. As documented in a *paper published in 2018*[2], an error-generic or error-specific evasion attack that would force a machine-learning-based system with similar structure to misclassify a sample, would require a much more complex strategy on the attacker's side to succeed against ESET's ML algorithm.

HOW ESET ML PROCESSES SAMPLES (FIGURE 15)

Upon entering the ESET ML engine, every sample is:

1. Subject to static code analysis that extracts its features, gathering data that are processed via deep learning algorithms.
2. Emulated as a part of dynamic analysis, producing a series of DNA genes. These are fed to a series of precisely-chosen classification models and yet another deep learning algorithm.
3. At the same time, the sample is executed within a sandbox environment and subjected to advanced memory analysis. The output is used for local-sensitive hashing in order to compare with known, periodically reviewed and automatically updated set of clean and malicious items.
4. The results from the preceeding steps are vectorized and consolidated either via a neural network or other forms of evaluation.
5. All gathered information is used to produce a final decision, considering the sample clean, potentially unwanted, or malicious.
6. The information is then provided to ESET clients via ESET LiveGrid®.

[2] Battista Biggioa, Fabio Rolia, Wild Patterns: Ten Years After the Rise of Adversarial Machine Learning, (2018), 6-7



It is important to note that, unlike some of the emerging security vendors: as part of sample processing, ESET also utilizes unpacking and behavioral analysis, as well as emulation. These steps are considered crucial to extract a sample's features properly, before they can be fed to the ML engine. Analyzing compressed or encrypted samples equals an attempt to classify noise that would render irrelevant results. This approach is like picking a winner of a singing contest solely by looking at the photos of the candidates, without giving them a chance to perform.

MACHINE LEARNING IN CURRENT ESET PRODUCTS

The power of ESET Machine Learning is accessible to customers of all sizes. Each endpoint and device that has ESET LiveGrid® enabled benefits from the cutting-edge ESET ML engine's accuracy and ability to analyze emerging threats.

ESET's enterprise customers also have machine learning technology at their disposal via three enterprise-grade products:

1. **ESET Enterprise Inspector (EEI)**
   - EEI is ESET's Endpoint Detection and Response (EDR) tool. It works by collecting real-time data about ongoing activity on endpoints, which is then matched against a set of rules to detect suspicious activities automatically. The gathered information is processed, aggregated and stored in a searchable form, creating a drill-down summary of unusual and suspicious activities.
   - EEI also provides the enterprise security team with information for forensic investigation of past incidents and offers response capabilities, to mitigate the presence of threat actors (advanced persistent threat or APT) in the network. ESET ML is integrated into EEI scanning and is essential to the process of flagging suspicious activities and samples.
2. **ESET Dynamic Threat Defense (EDTD)**
   - EDTD utilizes a cloud-based sandboxing technology to detect new, never-before-seen types of threats, and thus provides an additional layer of security to ESET products, such as Mail Security and Endpoint products. This sandbox consists of multiple types of sensors that complete static analysis of code, deep inspection of the sample with machine learning, in-memory introspection and behavior-based detection. Compared to Endpoint, EDTD represents a much more powerful detection engine that leverages a wider range of technologies and access to the vast collection of clean, potentially unwanted and malicious items collected by ESET in the past 30 years. It is much more effective to run such a solution in the cloud, making it more scalable and lowering demands on the customer's infrastructure.

3. **ESET Threat Intelligence (ETI)**
   - ESET Threat intelligence provides evidence-based information and actionable advice about existing or emerging threats. ETI can warn about malicious software or activity within an organization or among its customers. This information—supplied by ESET machine learning as well as other detection technologies—is analyzed and presented in a human-readable form, suitable for analysts in IT security departments or security operation centers (SOC).

With the full potential of machine learning yet to be unlocked, our engineers are constantly looking for tasks and products in the ESET portfolio that could benefit from the use of this technology.

CONCLUSION

While it is difficult to say which effects of machine learning—positive or negative—will prevail, what is an undeniable is the growing use of ML-powered systems on both sides of the cybersecurity divide, irreversibly transforming safety of the whole internet.

ESET closely monitors developments on the black-hat scene for signs of technology advances—as illustrated by our research-based predictions as well as both the spam and Emotet in-the-wild examples—and reacts by constantly improving its protective solutions.

Yet, despite the marketing campaigns of many emerging cybersecurity players, ESET’s three decades of experience show that proper protection cannot be achieved solely by relying on one technology—even if it is machine learning or deep learning.

A multi-layered approach is necessary to keep the solution resilient against adversaries and to achieve high detection rates and low numbers of false positives: important qualities in a security solution for companies of all sizes.

EXECUTIVE SUMMARY:

Advances in the field of machine learning have kicked off a completely new era. An era where almost any piece of data collected is processed and analyzed via algorithms that depend on machine learning technology—cybersecurity included. Yet even this innovation has its drawbacks and limitations.

This document seeks to describe the hype that machine learning technology has caused in cybersecurity and how this influences business decision makers.

We also outline in-the-wild cyberattacks, observed by ESET research, that show strong indications of ML use: namely, improvements to local spam emails, and attacks by Emotet malware.

Last but not least, we demonstrate how 20 years of ESET’s experience with machine learning has shaped our views, led us to early implementations on the backend as well as its more visible represntations in ESET’s current products.