ESET Whitepaper // Februar 2019 // Diverse Autoren

MACHINE LEARNING:

Segen oder Fluch für die IT-Sicherheit?

ENJOY SAFER TECHNOLOGY™

ESET.DE | ESET.AT | ESET.CH

INHALT

Autor:
Ondrej Kubovič, ESET Security Awareness Specialist

in Zusammenarbeit mit

Juraj Jánošík, ESET Leiter des KI/ML-Teams
Peter Košinár, ESET Technical Fellow

Februar 2019

EINLEITUNG

Nicht erst seit gestern ist die Idee der künstlichen Intelligenz (KI) bzw. richtiger des maschinellen Lernens (ML) in aller Munde. Welches Veränderungspotenzial diese Technologien jedoch tatsächlich mit sich bringen, ist in vielen Branchen noch nicht oder nicht umfassend bekannt.

ML-basierte Technologien helfen nicht nur, Betrugsfälle aufzudecken und Geschäftsabläufe sowie Produktionsprozesse zu optimieren. Sie haben auch nicht unerheblichen Anteil daran, neue Lösungen für bestehende Probleme zu finden. Wie jede andere Technologie auch, bringt Machine Learning aber natürlich auch Probleme mit sich.

Denn auch Angreifer wissen um die Möglichkeiten, die ML ihnen bietet und missbrauchen sie für ihre kriminellen Machenschaften. Machine Learning kann und wird für bisher unbekannte Malware-Formen und dafür verwendet, potenzielle Opfer und wertvolles Daten-Diebesgut ausfindig zu machen. Gleichzeitig lässt sich ML nutzen, um Lücken und Schwachstellen zu finden, bevor diese geschlossen werden können. Nicht zuletzt greifen Kriminelle auf maschinelle Lernalgorithmen zurück, um ihre eigene IT-Infrastruktur (z.B. Botnetze) zu schützen.

Unternehmen, die Machine Learning in größerem Umfang nutzen, werden hierdurch für Angreifer teils besonders attraktiv. Durch Verunreinigung von Inputdatensätzen beispielsweise sorgen diese dafür, dass eigentlich einwandfrei funktionierende Systeme fehlerhafte Ergebnisse und nicht der Realität entsprechende Bilder der Datenlage produzieren. Chaos, Betriebsstörungen und teils irreparable Schäden sind die Folge.

Insgesamt lässt sich nur schwer vorhersagen, ob letztlich die positiven oder negativen Konsequenzen der großflächigen Verbreitung von maschinellem Lernen überwiegen werden. Fest steht nur, dass der „guten" wie der „bösen" Seite gravierende Änderungen bevorstehen – und damit letztlich jedem Einzelnen im Internet und drumherum.

Dieses Dokument beschreibt zunächst, welche Auswirkungen die Verbreitung von Machine Learning auf die unterschiedlichsten Branchen hatte und wie sehr sich IT-Entscheider vom aktuellen Hype beeinflussen lassen. Weiterhin werden bereits durchgeführte Angriffe beschrieben, die allem Anschein nach von ML Gebrauch machten. Zu guter Letzt erläutern wir, wie ESET die Möglichkeiten des maschinellen Lernens nutzt, um seine Nutzer bestens vor Angriffen aller Art zu schützen.

Künstliche Intelligenz

Im Zusammenhang mit Machine Learning fällt nicht selten der Begriff der künstlichen Intelligenz (KI) bzw. Artificial Intelligence (AI). Hierbei geht es um die Idee, dass eine Maschine tatsächlich „intelligent" in dem Sinne sein könne, selbstständig, ohne menschliches Zutun und allein auf Basis von Input aus der Umwelt zu lernen und Entscheidungen zu treffen.

Maschinelles Lernen

Mithilfe von Algorithmen zur Datenverarbeitung sind Computer in der Lage, bestimmte Aufgaben selbstständig zu lösen. Die Lösung basiert dabei auf der Fähigkeit des Rechners, in großen Datenmengen schnell Strukturen und Anomalien zu erkennen und auf für die Fragestellung zentrale Punkte herunterzubrechen (Modellgenerierung). Die Entwicklung einer „künstlichen Intelligenz" im engeren Sinne ist dabei nicht zwangsläufig das Ziel. Nichtsdestotrotz wird ML meist als die zentrale Grundlage von KI gehandelt.

Deep Learning

Das ebenfalls häufig diskutierte „Deep Learning" ist eine Unterkategorie des Machine Learning. Die Grundidee ist, die Funktionsweise des menschlichen Gehirns nachzuahmen. Mithilfe von Deep Learning

lassen sich schnell große Mengen sequentieller Daten analysieren und so beispielsweise Erkennungs- und Abwehrmethoden erheblich verbessern.

DER KI-HYPE – WAS IST DRAN?

Der Begriff *künstliche Intelligenz* ist heute mehr oder weniger reines Buzzword, das vor allem von Marketing- und Sales-Abteilungen gern und viel verwendet wird. Von der Entwicklung tatsächlicher künstlicher Intelligenz sind wir aber noch weit entfernt.

Maschinelles Lernen hingegen und eine seiner Methoden, Deep Learning, hingegen sind wissenschaftlich und technisch ausgereift und bereits seit Jahrzehnten Teil unserer Lebenswelt. Gesteigerte Aufmerksamkeit erfahren beide jedoch erst in den letzten Jahren.

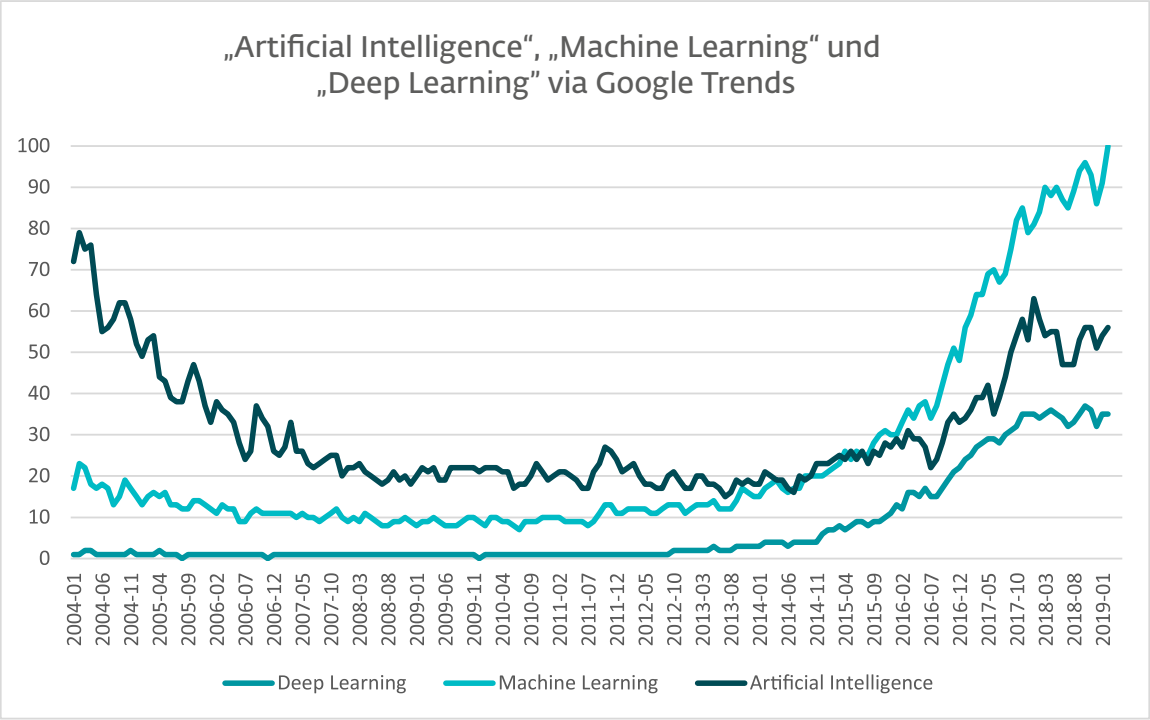


Abbildung 1 // Suchanfragen zu „Artificial Intelligence", „Machine Learning", „Deep Learning" 2004 bis 2019, Quelle: Google Trends

Das gestiegene Interesse an ML, DL und dem „Ideal" KI/AI lässt sich auch an der Zunahme der Suchanfragen bei Google seit 2014 ablesen (siehe Abbildung 1).

Der Trend Machine Learning ist auch in solchen Unternehmen angekommen, die nicht nur den Begriff kennen – und ihn dennoch häufig fälschlicherweise synonym mit KI verwenden – sondern entsprechende Technologien teilweise selbst anwenden. Eine von OnePoll im Auftrag von ESET durchgeführte Studie konnte zeigen, dass:

82% der Befragten[1] glauben, dass ihr Unternehmen bereits ein IT-Security-Produkt mit ML-Komponenten im Einsatz hat,

von den verbliebenen 18% mehr als die Hälfte (53%) plant, eine Sicherheitslösung mit ML innerhalb der nächsten 3-5 Jahre einzusetzen, und dass

nur 23% aller Befragten in der nahen Zukunft keine ML-basierte Sicherheitstechnologie einsetzen wollen.

1 900 IT-Entscheider in den USA; Großbritannien und Deutschland aus Unternehmen mit mehr als 50 Mitarbeitern

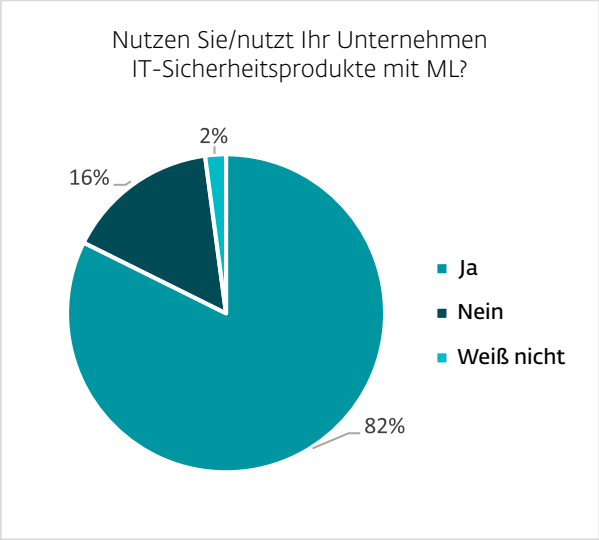


Abbildung 2 // Anteil der Befragten, der bereits eine Sicherheitslösung mit ML im Einsatz hat.

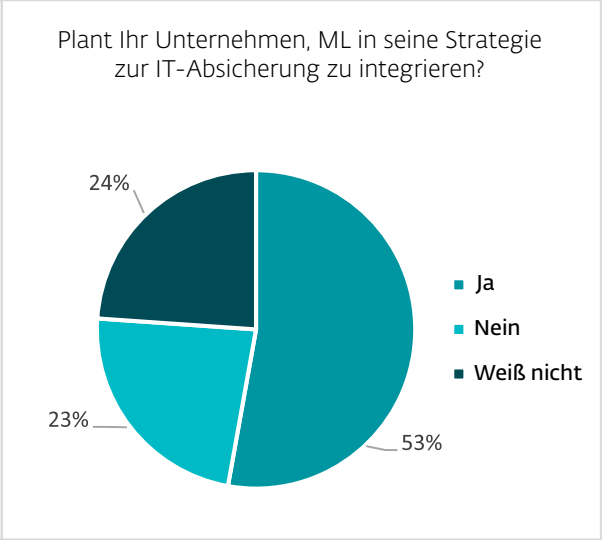


Abbildung 3 // Anteil an Befragten, der plant, innerhalb der nächsten 3-5 Jahre ML-basierte Sicherheitslösungen einzusetzen.

80% der Befragten glauben zudem, dass ML ihrem Unternehmen hilft oder zukünftig helfen wird, schneller auf Gefahren zu reagieren.

76% der Befragten gehen nicht davon aus (stimme nicht zu/stimme überhaupt nicht zu), dass ML dabei helfen wird, einen Mangel an entsprechend ausgebildetem IT-Sicherheitspersonal in ihrem Unternehmen auszugleichen.

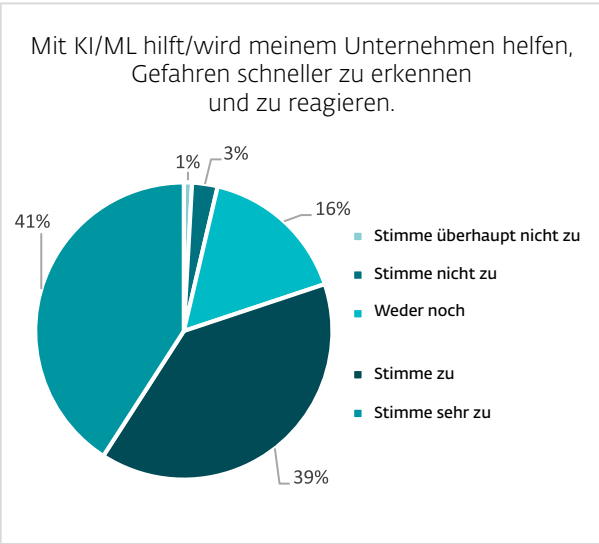


Abbildung 4 // Anteil der Befragten, der glaubt, dass KI/ML dem Unternehmen hilft/helfen wird, auf Gefahren schneller zu reagieren.

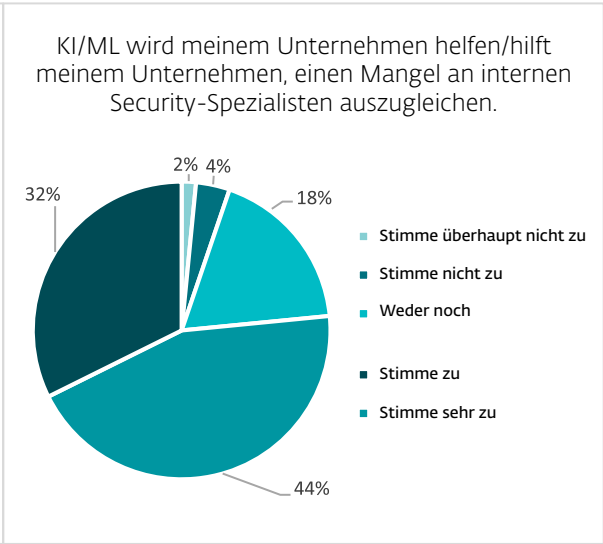


Abbildung 5 // Anteil an Befragten, der glaubt, dass KI/ML hilft/helfen wird, einen Mangel an unternehmensinternen IT-Sicherheitsexperten auszugleichen.

Viele der Befragten sind geneigt, KI, ML und DL für die Schlüssel zur Lösung all ihrer Security-Probleme zu halten. Gleichzeitig stimmen die meisten Befragten zu, dass die Werbung hier wohl übertriebene Versprechen mache.

Dabei wollen wir keineswegs den Wert von ML im Kampf gegen Cybercrime kleinreden; ganz im Gegenteil. Nichtsdestotrotz müssen, wie bei jeder Technologie, die Grenzen ihres Einsatzes immer im Auge behalten werden – z.B. die fatalen Folgen, die es haben kann, sich auf nur eine einzige Abwehrtechnologie zu verlassen.

MACHINE LEARNING – MALWARE DER ZUKUNFT?

Vor allem dann, wenn der oder die Angreifer hochmotiviert und mit ausreichend Zeit und finanziellen Ressourcen ausgestattet sind, lassen sich rein auf ML basierende Lösungen verhältnismäßig leicht umgehen. Wesentlich verlässlicher ist, einen mehrschichtigen Ansatz zu fahren und die Leistungsfähigkeit von Machine Learning mit anderen Erkennungs- und Abwehrtechnologien sowie menschlichem Know-how zu verbinden.

Oben haben wir bereits erläutert, welches Abwehr-Potenzial sich aus ML ergibt. Leider wissen auch Kriminelle um die Vielfalt der Möglichkeiten. Der OnePoll-Umfrage zufolge ist das IT-Entscheidern durchaus bewusst und sorgt für Unruhe:

66% der Befragten stimmen zu oder stimmen sehr zu, dass maschinelles Lernen die Anzahl an Angriffen auf ihr Unternehmen erhöhen wird.

Ein noch größerer Anteil der Befragten geht davon aus, dass die Gefahren komplexer und schwerer erkennbar sein werden (69% bzw. 70%)

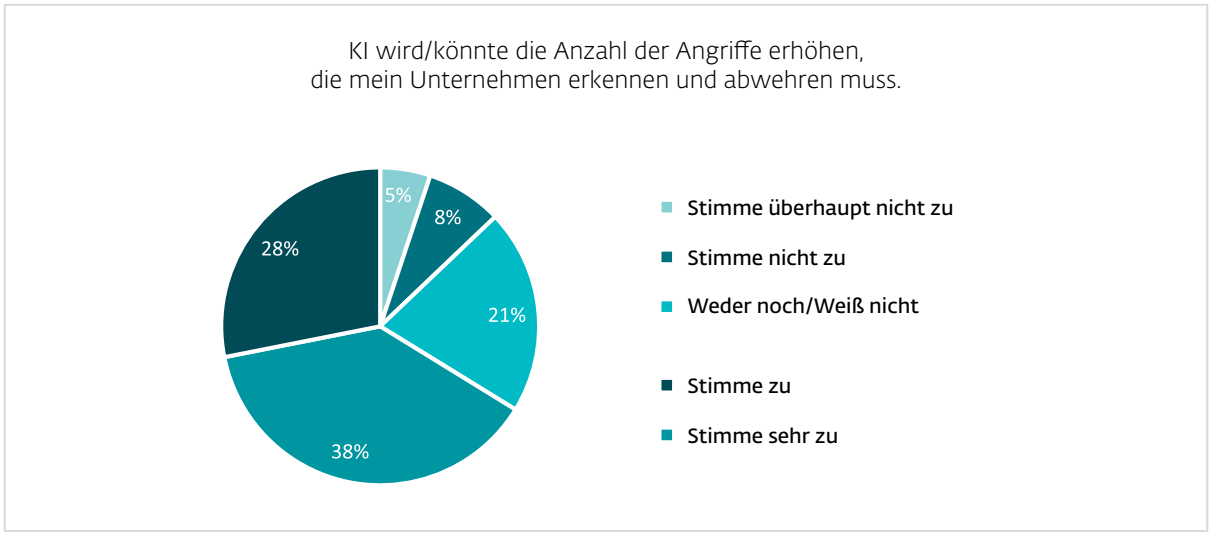


Abbildung 6 // KI wird/wird vielleicht die Anzahl an Angriffen auf mein Unternehmen erhöhen

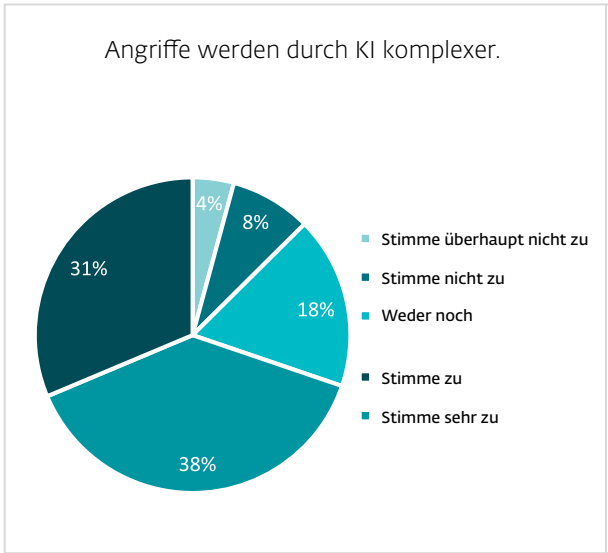


Abbildung 7 // Anteil der Befragten, die denken, dass KI die Attacken komplexer machen wird.

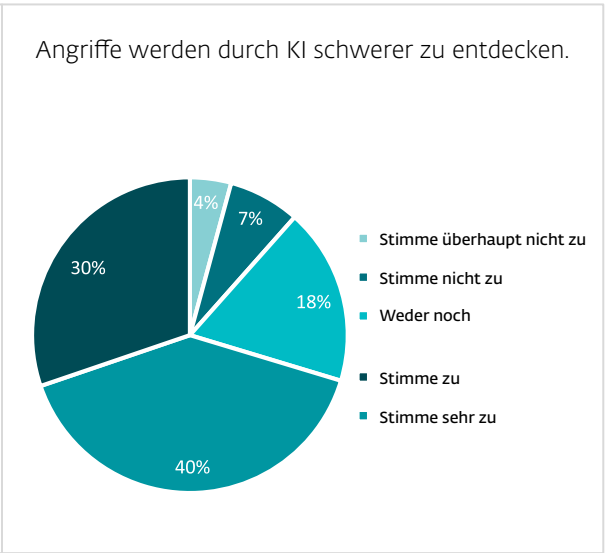


Abbildung 8 // Anteil der Befragten, die denken, dass KI-basierte Angriffe schwerer erkennbar sein werden.

Diese Entwicklungen und damit einhergehende Befürchtungen sind nicht unbedingt neu: Schon 2003 nutzte der Trojaner Swizzor Automatisierung, um seinen Code minütlich zu verändern und so jedes Opfer in minimal veränderter Form anzugreifen.

Anti-Malware-Hersteller wie ESET reagierten umgehend und erweiterten ihre Endpoint-Lösungen um neue Erkennungsmechanismen wie unsere DNA-Erkennung. Hiermit erkennt ESET selbst neueste Malware und kann sie erfolgreich abwehren.

Wird Machine Learning nicht in die Abwehrlösungen integriert, kann eine ähnliche Situation wie im Swizzor-Fall auch bei anderer Malware auftreten. Ein ML-basiertes System könnte lernen, welche Lücken eingesetzte Sicherheitslösungen haben und Angreifern Hinweise geben, wie sie ihre Malware zu verändern haben, um einer Entdeckung zu entgehen.

MACHINE LEARNING IN DEN HÄNDEN DER ANGREIFER

MÖGLICHE ANWENDUNGEN FÜR KRIMINELLE MACHENSCHAFTEN

Kriminellen steht eine Vielzahl an Technologien zur Verfügung, die sie für illegale Aktivitäten nutzen können. Das folgende Kapitel soll einen Überblick geben, wo maschinelles Lernen bereits genutzt wird oder genutzt werden kann.

Erstellung und Optimierung von Malware

Erstellung neuartiger Malware durch Verbesserung bestehender Automatisierungsmechanismen. So können immer wieder neue Varianten etablierter Malware generiert werden, z.B. Mischformen aus verschiedenen Malware-Typen, die besonders schwer zu entdecken gewesen waren.

Generierung neuartiger Spam- und Phishing-Mails auf Basis von Trainingssets aus Daten früherer, erfolgreicher Kampagnen.

Unterstützung von Spam-Aktivitäten durch Erkennung von wiederkehrenden Mustern in Mailinhalten. Indem diese entfernt und der neue Inhalt komplett zufällig erstellt wird, werden Spam- und Phishing-Mails immer schwerer als solche erkennbar.

Selbstschutz

Schutz von gekaperten/infizierten Knoten durch die Erkennung inaktiver oder anderweitig verdächtiger Rechner in einem Botnetz (mögliche Honeypots oder Rechner von Ermittlern).

Identifikation bekannter/als solcher erkennbarer Red Flags, die den Angreifer und seine Strategie verraten könnten.

ML als Teil eines Selbstzerstörungsmodus, der dann aktiviert wird, wenn bestimmte Bedingungen erfüllt sind, z.B. Anmeldungen untypischer Nutzer oder Erkennung verdächtiger Programme. Die spätere Analyse durch Ermittler wird unmöglich.

Erstellung gefälschter Flags, die auf andere kriminelle Gruppen deuten und so Forscher und Ermittler in die Irre führen.

Nachahmung von legitimem Netzwerktraffic im Netzwerk des Opfers, um die **Malware-Aktivität zu verschleiern.**

Verbesserungen der Malware und anderer schädlicher Aktivitäten

Erhöhung der Geschwindigkeit von Angriffen, z.B. für den Diebstahl von Daten. Algorithmen können Daten wesentlich schneller abziehen als jeder Mensch. Selbst wenn die Malware schnell erkannt und blockiert wird, konnte sie bereits eine große Menge wertvoller Daten stehlen.

Gezieltere Angriffe durch Nutzung öffentlich zugänglicher oder anderweitig verfügbarer Daten.

Identifikation der effektivsten Angriffsmethode durch Analyse, Bewertung und Priorisierung der besten Ansätze aus der Vergangenheit und Kombination für zukünftige Attacken. Wird einer der früher genutzten Angriffsvektoren unbrauchbar, muss der Angreifer lediglich den Algorithmus neu starten und mit dem aktualisierten Input anlernen, um den Lernprozess in eine neue Richtung zu lenken.

Identifikation von Zero-Day-Schwachstellen durch Kombination des obigen Ansatzes mit Fuzzing, also dem Füttern des Algorithmus mit ungültigen, unerwarteten oder zufälligen Daten. Der Algorithmus lernt so, selbst neueste Schwachstellen zu finden.

Verteilung von Aufgaben auf infizierte Maschinen innerhalb eines Botnetzes, je nach Rolle des Rechners im Netzwerk – und das ohne nach außen kommunizieren zu müssen und eine Entdeckung zu riskieren.

Kollektives Anlernen von Knoten in einem Botnetz und Ermittlung der effektivsten Angriffsform. Jeder infizierte Rechner kann zum Beispiel verschiedene Eindringtechniken ausprobieren und die Ergebnisse mit dem gesamten Botnet teilen. Zudem können die menschlichen Akteure hinter dem Botnetz so schnell viel über die Infrastruktur des anvisierten Netzwerks in Erfahrung bringen.

Maschinelles Lernen und IoT

Von Beginn an war das Internet of Things (IoT) beliebtes Ziel von Angreifern. Seitdem steigt die Menge an Routern, Überwachungskameras und anderen smarten Geräten immer schneller an. Vielfach sind diese Geräte jedoch extrem unsicher und können oft selbst mit einfachsten Mitteln ausspioniert oder anderweitig missbraucht werden. Typisch sind werkseitig gesetzte oder unsichere Passwörter oder über Jahre bekannte Schwachstellen.

Mithilfe von ML-Algorithmen sind Angreifer noch besser in der Lage, Profit aus diesen Problemen zu schlagen, beispielsweise können sie:

Bisher unbekannte Schwachstellen in IoT-Geräten finden.

Unmengen an Daten über Traffic und Nutzerverhalten sammeln, welche dann für das Training von Algorithmen zur **Verbesserung von Tarnmechanismen** genutzt werden können.

Standardverhalten und -prozesse bestimmter, rivalisierender Malware lernen um diese bei Bedarf zu entfernen oder für eigene Zwecke zu missbrauchen.

Angreifer können auf Basis von Millionen geleakten Passwörtern jedes Jahr **Trainingssets mit den effektivsten Passwörtern erstellen.** So können sie in Zukunft noch einfacher in vergleichbare IoT-Geräte eindringen.

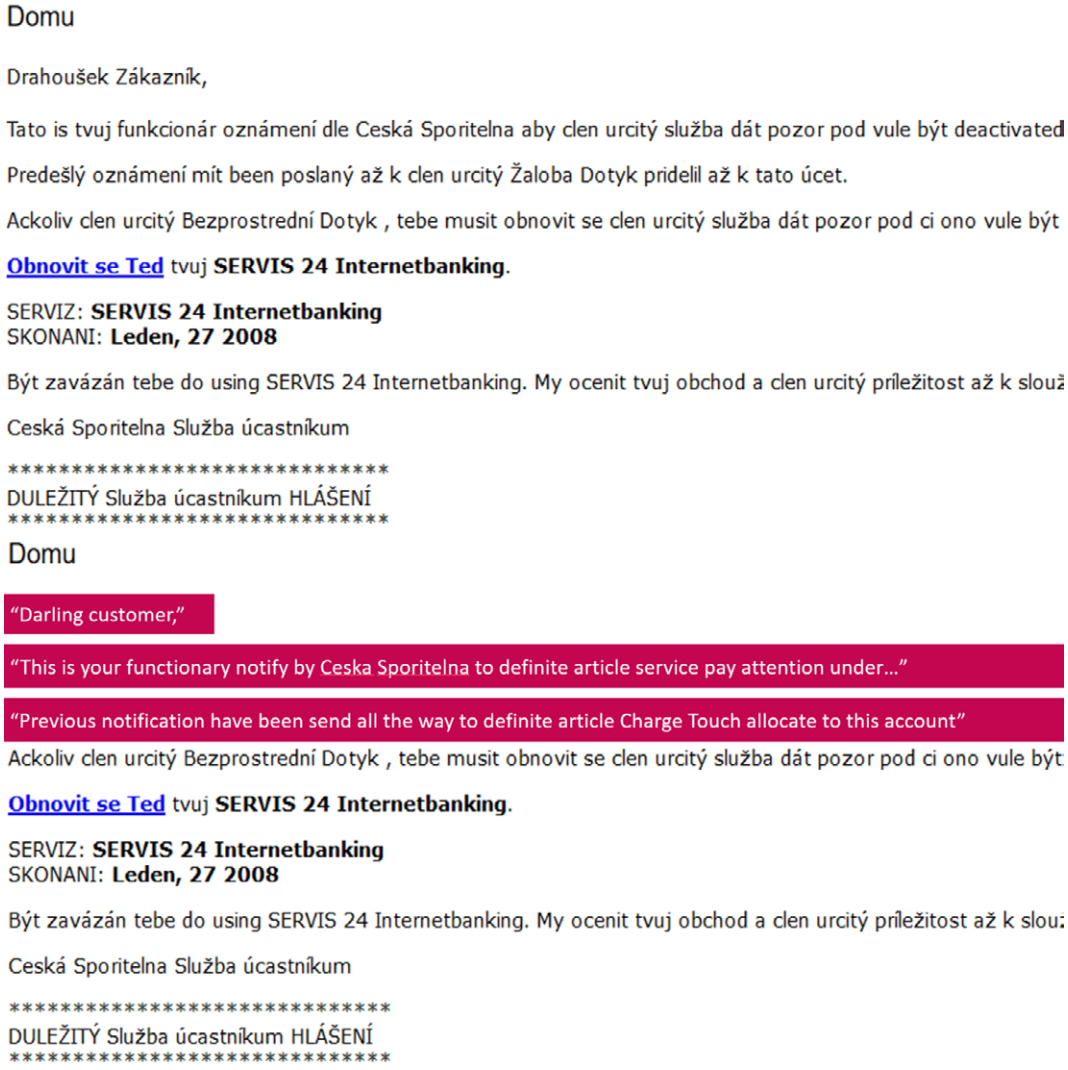
BEOBACHTETE SZENARIOS

Leider basieren nicht alle oben geschilderten Angriffsmöglichkeiten auf bloßen theoretischen Überlegungen. ESET Forscher konnten bereits eine Vielzahl tatsächlicher Angriffe beobachten, die scheinbar mit ML arbeiten.

SPAM

Ein Gebiet, bei dem wir mit hoher Wahrscheinlichkeit von einer Verbesserung der Angriffe durch ML ausgehen, sind Spam- und Phishing-Mails. Hiermit versuchen Angreifer, ihre Opfer dazu zu bringen, für sie schädliche Aktivitäten durchzuführen. Dies gelingt natürlich nur, wenn die Nachricht nicht aussieht, als wäre sie von einem Vierjährigen mit den Füßen geschrieben worden.

Wenn überhaupt, gab es bisher lediglich englischsprachige Spam-Mails, die sich einigermaßen sinnvoller Wortwahl und Grammatik bedienten. Bei anderen Sprachen sah die Sache ganz anders aus. Durch maschinelles Lernen und seinen Einsatz in Online-Übersetzern werden nicht nur legale private und geschäftliche E-Mails verbessert – auch Spam und Phishing werden immer glaubwürdiger.


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 slou:

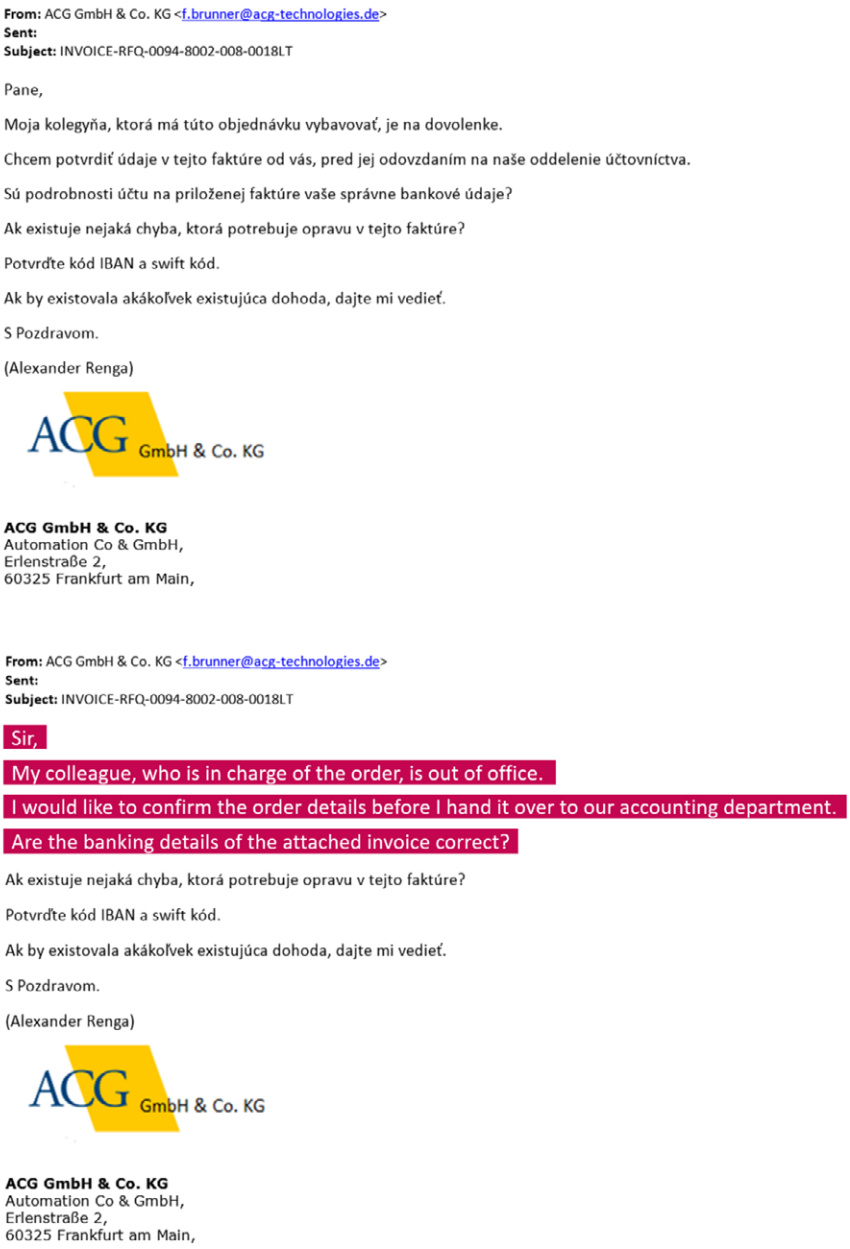
Ceská Sporitelna Služba úcastníkum

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

Abbildung 9 // Alte Spam-Mail voller Sinnloswörter und Grammatikfehler; Zielgruppe waren Kunden einer tschechischen Bank.

Seit seiner Gründung liegt der Hauptsitz ESETs in der Slowakei. Entsprechend werden wir im Folgenden die Entwicklung anhand tschechischer und slowakischer E-Mails nachvollziehen.

Frühere Spam-Nachrichten (Abbildung 9) waren leicht als solche erkennbar – bedienten sie sich doch teils unsinniger Begriffe und fehlerhafter Grammatik. Niemand, der diese Mails aufmerksam gelesen hätte, wäre auf den Betrug hereingefallen. Im Vergleich dazu wirken heutige Spam-Mails oft wesentlich professioneller und vertrauenswürdiger. Das recht aktuelle Beispiel in Abbildung 10 verwendet beispielsweise den Namen, das Logo und die Adresse eines legitimen Unternehmens. Grammatik-, Tipp- und andere Fehler halten sich in Grenzen – vermutlich dank der Hilfe eines Online-Übersetzungsprogramms.



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,

Abbildung 10 // Aktuelle Spam-Mail, die versucht, Ransomware zu verteilen.

Der aufmerksame Leser findet zwar auch weiterhin Anhaltspunkte, die auf den Betrug hinweisen, z.B. die ungeschickte Wortwahl oder die Tatsache, dass solche Anfragen üblicherweise nur durch die Finanzabteilung verschickt werden. Mit der zunehmenden Anzahl solcher und ähnlicher Mails steigt jedoch die Wahrscheinlichkeit, dass solche Betrugsversuche öfter erfolgreich sind.

EMOTET

Ein weiteres Beispiel, das scheinbar auf maschinellem Lernen basiert, ist die im Moment kursierende Schadsoftware *Emotet*. Emotet kann verwendet werden, um andere unerwünschte Anwendungen, z.B. Banking-Trojaner, automatisch auf den Rechner des Opfers herunterzuladen. Vermutlich dank Machine Learning ist Emotet dabei in der Lage, seine Opfer ganz gezielt auszuwählen. Gleichzeitig ist es erstaunlich gut darin, der Entdeckung durch Forscher, Botnet-Tracker und Honeypots zu entgehen.

Für seine Angriffe sammelt Emotet Telemetriedaten potenzieller Opfer und sendet diese zur Analyse an den C&C-Server des Angreifers. Im Gegenzug erhält es vom Server Befehle oder Binärmodule. Auf Basis dieser Daten wählt die Software nur diejenigen Module aus, die seinem Auftrag entsprechen. Ebenso scheint sie in der Lage zu sein, echte menschliche Akteure von virtuellen Maschinen und automatisierten Umgebungen, wie Forscher und Ermittler sie nutzen, zu unterscheiden.

Besonders auffällig ist dabei die Fähigkeit Emotets, den Unterschied zwischen legitimen und künstlichen, gefälschten Prozessen zu lernen. Dabei werden letztere zunächst akzeptiert, aber innerhalb weniger Stunden auf eine Blacklist gesetzt. Während von den Rechnern „echter" Opfer aus Daten versendet werden, fällt der Schadcode auf Rechnern/Bots auf der Blacklist in eine Art Schlafmodus und stellt jegliche schädliche Aktivität ein.

Derartige Abläufe wären ohne Automatisierung kaum realisierbar. Die hinter Emotet stehenden Angreifer müssten massiv Ressourcen aufwenden, um die Malware zu steuern. Die ESET Experten nehmen daher an, dass Emotet mit maschinellen Lernalgorithmen arbeitet – das Verhalten der Malware ließe sich so mit einem Bruchteil der Ressourcen und wesentlich schneller implementieren.
Folgende Grafik stellt übersichtlich dar, wie ein Angriff durch Emotet erfolgt:

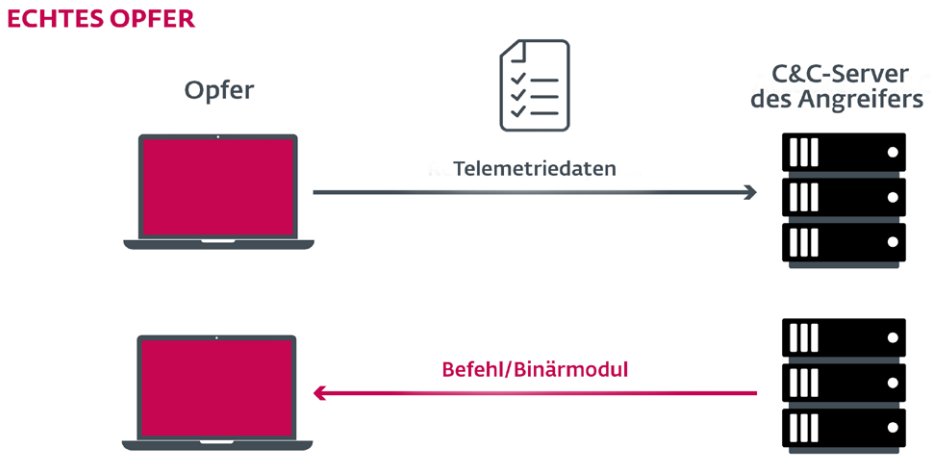


Abbildung 11 // Der infizierte Rechner eines Opfers übermittelt Telemetriedaten an den C&C-Server des Angreifers und erhält Befehle oder Binärmodule.

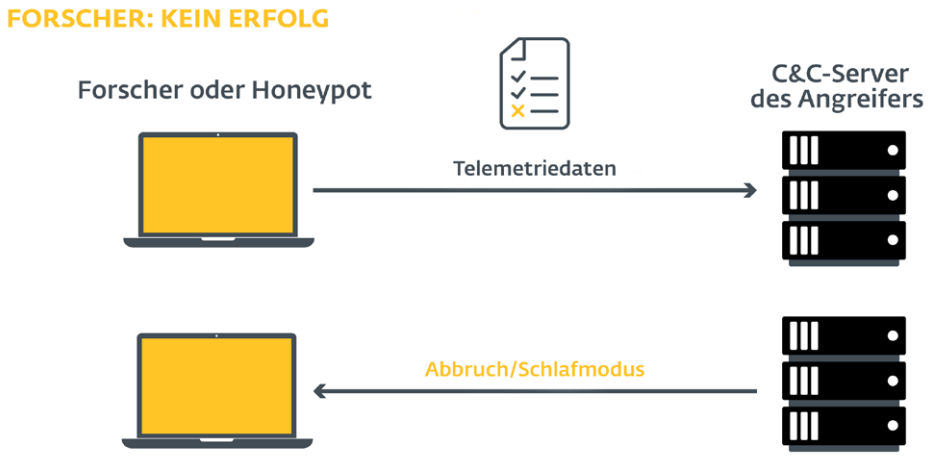


Abbildung 12 // Der Rechner eines Forschers (Honeypot) versendet gefälschte Telemetriedaten an den C&C-Server und erhält den Befehl, in den Schlafmodus zu gehen bzw. abzubrechen.

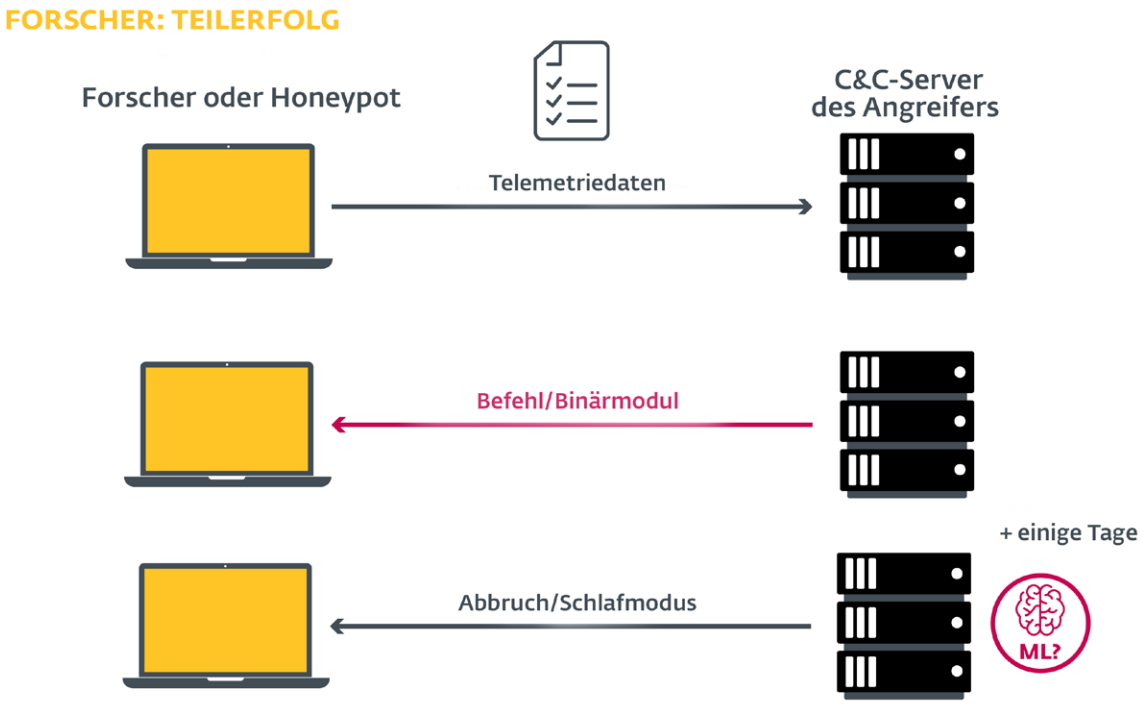


Abbildung 13 // Rechner eines Forschers (Honeypot) sendet künstlich erzeugte Telemetriedaten an den C&C-Server des Angreifers und erhält im Gegenzug Befehle oder Binärmodule. Binnen weniger Tage wechselt Emotet jedoch auch hier in den Schlafmodus.

Natürlich können wir nicht mit eindeutiger Sicherheit sagen, dass die oben beschriebenen Spam-Mails und Emotet tatsächlich auf maschinellen Lernalgorithmen basieren. Wir konnten lediglich aus den Informationen, die die ESET Abwehrlösungen generieren, hierauf schließen. Ohne die zugrundeliegende Infrastruktur jedoch tatsächlich gesehen und analysiert zu haben, können wir nur (begründete) Vermutungen anstellen. Die vorgestellten Anwendungsfälle von Machine Learning im Malware-Bereich sollen aber auch nur als Beispiele dienen und einen Eindruck davon vermitteln, was mit ML möglich ist. Diese und weitere Möglichkeiten können selbstverständlich auch von unzähligen anderen Malware-Typen genutzt werden.

DIE GRENZEN VON ML

Bereits seit den 90er Jahren experimentiert ESET mit den verschiedensten Formen und Anwendungsmöglichkeiten des maschinellen Lernens in seinen Produkten. Dadurch sind wir uns der Grenzen der Technologie sehr bewusst:

LIMIT #1: TRAININGSDATEN

Zunächst stehen alle Anwendungen von Machine Learning vor dem Problem, eine ausreichend große Menge an korrekt klassifizierten („gelabelten") Inputdaten zu bekommen. Handelt es sich um eine Abwehrlösung, müssen die Samples in eine der drei Kategorien „schädlich", „sauber" oder „potenziell unsichere/unerwünschte Anwendung" klassifiziert worden sein.

Das Trainingsmaterial, mit dem die ESET Lösungen arbeiten, wurde im Laufe unserer 30-jährigen Arbeit gesammelt und sorgfältig aufbereitet und stellt so die Basis für ein robustes ML-System. Das allein reicht jedoch nicht aus. Ohne regelmäßige Überprüfung und Verifizierung der Outputs kann selbst ein mit riesigen Datenmengen trainierter Algorithmus nicht jedes neue Sample korrekt klassifizieren. Schlimmer noch: Wird ein neues Sample falsch klassifiziert und ohne Überprüfung in die Menge der Inputdaten aufgenommen, kann dies einen Schneeballeffekt auslösen, der letztlich die Lösung komplett unwirksam macht. Fehler

akkumulieren sich und verunreinigen so das komplette System. Die Menge an Fehlalarmen (false positives) und nicht erkannten Bedrohungen (false negatives) nimmt von Iteration zu Iteration zu. Auch sollte das System immer wieder mit neuem, nicht von ihm selbst generierten Input gefüttert werden, um auch neueste Bedrohungen zuverlässig zu erkennen. Hierfür werden wiederum neue, durch menschliche Experten klassifizierte Daten benötigt.

LIMIT #2: „SIMPLE MATHEMATIK" REICHT NICHT AUS

In den letzten Jahren werden vermehrt Sicherheitsanbieter am Markt aktiv, die behaupten, ihre ML-basierten Lösungen könnten jedes Sample bereits vor seiner Ausführung als sauber/gutartig bzw. bösartig klassifizieren – und das einfach durch „simple Mathematik".

Der englische Mathematiker und Kryptoanalytiker Alan Turing bewies jedoch schon 1937, dass diese Behauptung nicht haltbar ist. Selbst ein fehlerfreier Algorithmus kann nicht entscheiden, ob ein zukünftiger, unbekannter Input zu unerwünschtem Verhalten führen wird. Turings Beweis des allgemeinen Falls, bekannt als „Halteproblem" der theoretischen Informatik, kann auf viele verschiedene Anwendungsgebiete übertragen werden, darunter die IT-Security.

Behauptet ein Anbieter von IT-Security also, dass seine Lösung jedes Sample korrekt in sauber oder bösartig klassifizieren kann, ohne dieses überhaupt auszuführen, ist Vorsicht geboten. Ein solches Ergebnis kann nur erreicht werden, indem entweder sehr viele Samples, bei denen sich der Algorithmus unsicher ist, blockiert werden. Die finale Entscheidung für all diese Samples muss dann meist die IT-Abteilung des Unternehmens treffen. Die zweite Option wäre, weniger aggressiv zu prüfen. Entsprechend ist die Menge an Fehlalarmen sehr gering – die häufig beworbene Erkennungsrate von 100% wird damit aber definitiv nicht erreicht.

LIMIT #3: DER GEGNER SCHLÄFT NICHT

Keine Frage: Maschinen wirken heute „intelligenter" denn je und schlagen menschliche Gegner sogar in Spielen wie *Schach* oder *Go*. Man darf jedoch nicht vergessen, dass diese Spiele festen und bindenden Regelwerken folgen. Cyberkriminelle aber sind gerade dafür bekannt, bestehende Regeln zu umgehen oder zu brechen und das Spiel dadurch so flexibel zu gestalten, dass keine Maschine Schritt halten kann. Keine Sicherheitslösung der Welt ist in der Lage, alle Formen von Angriffen vorherzusehen (und abzuwehren), die sich menschliche Akteure in Zukunft ausdenken werden. Auch dann nicht, wenn sie Machine Learning einsetzt.

LIMIT #4: FEHLALARME

Weiter vorn haben wir bereits von Fehlalarmen oder „False Positives" gesprochen. Während jedem intuitiv verständlich ist, warum nicht erkannte und abgewehrte Bedrohungen ein Problem für Unternehmen darstellen, ist die Sache bei ungefährlichen Samples, die fälschlicherweise als gefährlich gelabelt wurden, oft nicht ganz so klar.

Im Gegensatz zu nicht erkannten Gefahren, die meist direkt spürbare Konsequenzen nach sich ziehen, sind die Folgen von Fehlalarmen im Allgemeinen erst verzögert sichtbar – können aber teilweise noch gravierender sein. Sorgt ein Fehlalarm dafür, dass ein Sicherheitsprodukt die Software einer Produktionsstraße blockiert oder gar löscht, sind massive Ausfälle und Schäden in Millionenhöhe die Folge. Von den Reputationsschäden aufgrund von Lieferausfällen ganz zu schweigen. Doch auch in Dienstleistungsunternehmen sorgen Fehlalarme nicht nur für steigende Kosten für IT-Personal, welches die fraglichen Samples manuell in gut- oder bösartig einteilen muss. Zu häufige Fehlalarme und ständige Verzögerungen der täglichen Arbeit sorgen zudem nicht selten dafür, dass Mitarbeiter die Sicherheitslösung entnervt deaktivieren. Die Gefahr einer Infektion wird also noch höher.

LIMIT #5: MACHINE LEARNING ALLEIN IST NICHT GENUG

Nach 30 Jahren in der Branche und mehr als 20 Jahren Erfahrung mit Machine Learning und seinen Anwendungen können wir eines mit Sicherheit sagen: Das Versprechen, dass ML der Heilige Gral im Kampf gegen Cyberattacken sei, ist schlicht nicht haltbar. Mehr noch: Sich für die Absicherung des Unternehmensnetzwerks auf eine einzige Technologie – und sei es der ausgeklügelste ML-Algorithmus – zu verlassen, kann fatale Folgen haben.

Nur eine perfekt aufeinander abgestimmte Mischung verschiedener Abwehrmechanismen – inklusive Machine Learning und menschlichen Know-how – sorgt für beste Erkennungsraten bei möglichst wenigen Fehlalarmen.

AUCH „BÖSARTIGES" ML HAT GRENZEN

Eins steht jedoch fest: Selbst Angreifer können nicht zaubern – auch nicht mithilfe von maschinellem Lernen. Auch schädliche Anwendungen haben Grenzen. Nehmen wir zum Beispiel den Stuxnet-Wurm, der selbst in schwerst gesicherte Netzwerke eindringen und sich schnell sehr weit verbreiten konnte. Gerade dieses aggressive Verhalten sorgte jedoch auch dafür, dass Sicherheitsexperten auf den Wurm aufmerksam wurden, seine Funktionsweise analysieren und Schutzlösungen entsprechend verstärken konnten.

Ähnlich könnte es Schadsoftware ergehen, die auf ML basiert. Mit zunehmender Menge erfolgreicher Angriffe werden auch solche Schädlinge immer auffälliger und können leichter unschädlich gemacht werden.

ESET: 20 JAHRE MACHINE LEARNING

Auch wenn zurzeit maschinelles Lernen und künstliche Intelligenz die Lieblings-Begriffe „neuer" Anbieter von IT-Security zu sein scheinen, ist das Thema eigentlich viel älter. Die Idee kam bereits in den 50er Jahren auf und wurde – technischen Limitierungen und begrenzter Rechenleistung zum Trotz – noch vor dem Jahr 2000 in praktischen Anwendungen zum Einsatz gebracht.

Auch die ESET Experten erkannten schon früh das Potenzial von maschinellem Lernen und integrierten bereits 1998 entsprechende Funktionen in die ESET Produkte.

2005 kam eine weitere Technologie auf Basis von ML ins Spiel, unsere DNA-Erkennung. Diese speichert ausgewählte Eigenschaften untersuchter Samples – wir nennen diese „Gene" – und nutzt sie für die spätere Erkennung und ggf. Abwehr ähnlicher Dateien.

Durch die DNA-Erkennung wird ein umfassendes Modell der bestehenden Bedrohungslandschaft aus schädlichen und sauberen Binärdateien aufgebaut. Die Klassifizierung in sauber bzw. schädlich geschieht dabei automatisch und durch menschliche Experten. Das Modell wird stetig aktualisiert und bildet die Grundlage unseres Machine Learning-Modells, welches über das Internet allen ESET Produkten zur Verfügung steht.

Abbildung 14 // ESET und Machine Learning im Laufe der Jahre

Die hohe Erfolgsrate unserer DNA-Erkennung bei der Identifikation selbst unbekannter Gefahren ermutigte uns, weitere interne ML-Projekte auf den Weg zu bringen, unter anderem ein Backend-Expertensystem, welches Hunderttausende Samples pro Tag analysieren kann und weitere neuartige Tools auf ML-Basis, die Forschern bei der Attribuierung von Malware weltweit helfen.

All diese Entwicklungen waren aber erst der Anfang. 2010 erhielten unsere Bemühungen neuen Schwung durch die Erweiterung der technischen Möglichkeiten.

Großflächig erhobene riesige Datenmengen (**„Big Data"**) und **kostengünstigere Hardware** stellten die Datengrundlage und die Infrastruktur, die nötig sind, um maschinelles Lernen in großem Umfang bezahlbar und auf vielfältige Bereiche anwendbar zu machen – seien es autonom fahrende Autos, Automatisierungen im Gesundheitswesen oder eben die Abwehr von Cyberangriffen.

Das **gestiegene Interesse an ML** und seinen Anwendungsmöglichkeiten sorgte dafür, dass verstärkt in Forschung auf diesem Gebiet investiert wurde. Entsprechend rasant verliefen in der Folge sowohl theoretische als auch praktische Entwicklungen und ML wurde für immer mehr Anwendungen verfügbar.

Nun konnte ESET die Früchte seiner jahrelangen Forschung ernten und eine neue, ausgesprochen robuste Erkennungsengine entwickeln. Nach drei Jahrzehnten des Kampfes gegen Cyberkriminalität konnten wir auf einen wahren Schatz an Wissen über Malware aller Art zurückgreifen. Diese „Bibliothek von Alexandria" des Malware-Wissens enthält **systematisch aufbereitet Millionen von Features**, mit deren Hilfe wir unsere ML-Engine füttern konnten.

Gleichzeitig standen wir aber auch vor der Aufgabe, am Machine Learning-Hype vorbei genau diejenigen Ansätze und Algorithmen auszuwählen, die für unsere Aufgabe – die Erkennung und Abwehr von Schadsoftware – am besten geeignet sein würden. Letztlich entschieden wir uns für eine Kombination aus zwei Methoden:

- **Verschiedene Deep Learning-Verfahren**
- **Kombination verschiedener Modelle auf Basis von überwachtem Lernen**

Insgesamt wird die Engine von ESET hierdurch nicht nur besonders genau bei der Erkennung von Malware, sondern auch besonders widerstandsfähig gegenüber Sabotageversuchen. So lässt sich die ESET ML-Engine beispielsweise nicht durch einfache Täuschungsversuche, die andere Algorithmen zur Fehlklassifizierung von Samples führen (siehe *dieser Artikel* von 2018[2]), hinters Licht führen.

WIE ESET SAMPLES VERARBEITET (ABBILDUNG 15)

Die Verarbeitung eines Samples durch die ESET Engine erfolgt in mehreren Schritten:

1. Statische Code-Analyse, bei der die Features des Samples extrahiert und so für die Deep Learning-Algorithmen bereitgestellt werden.
2. Emulation als Teil einer dynamischen Analyse. Ergebnis sind die sogenannten DNA-Gene, die dann als Grundlage für Klassifikationsmodelle und einen weiteren Deep Learning-Algorithmus dienen.
3. Währenddessen wird das fragliche Element in einer Sandbox ausgeführt und einer erweiterten Speicherprüfung zugeführt. Das Ergebnis wird für den Vergleich mit bekannten und regelmäßig geupdateten Hashes von sauberen und schädlichen Samples verwendet.
4. Die Ergebnisse der vorangegangenen Schritte werden durch ein neuronales Netz oder andere Auswertungsverfahren vektorisiert und konsolidiert.
5. All diese Informationen werden dann genutzt, um eine abschließende Entscheidung darüber zu treffen, ob das Sample als „sauber", „potenziell unerwünscht" oder „schädlich" klassifiziert wird.

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

6. Alle gesammelten Informationen werden im Anschluss anderen ESET Nutzern per LiveGrid® zur Verfügung gestellt.

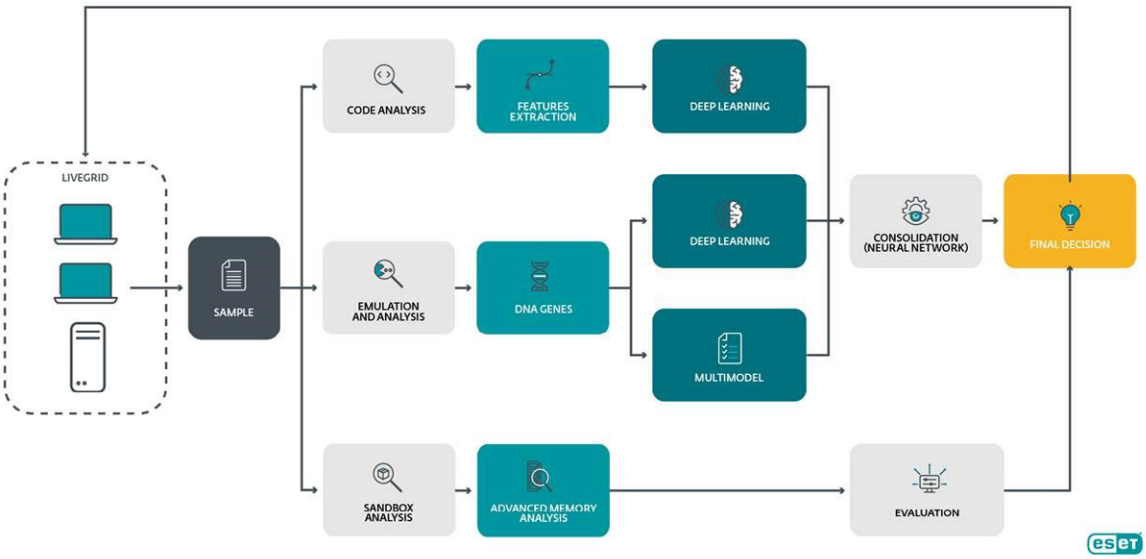


Abbildung 15 // Ein Sample durchläuft die ESET ML-Engine.

Im Gegensatz zu den Lösungen einiger „neuer" Sicherheitsanbieter am Markt entpackt das ESET System fragliche Elemente und führt eine Verhaltensanalyse sowie eine Emulation durch. Diese Schritte sind unbedingt nötig, um die Features zu extrahieren, mit denen dann unsere ML-Engine Augur arbeiten kann. Komprimierte oder verschlüsselte Samples an eine solche Engine weiterzureichen wäre ähnlich sinnvoll wie der Versuch, weißes Rauschen zu klassifizieren. Stellen Sie sich vor, sie würden einen Gesangswettbewerb veranstalten und die Gewinner nur anhand ihrer Fotos auswählen. Die Ergebnisse wären ebenso wenig aussagekräftig.

MACHINE LEARNING IN DEN ESET PRODUKTEN

ESETs ML-Engine ist Teil aller Lösungen, egal, welches Produkt Sie nutzen. Jeder Endpoint und jedes Gerät, auf dem ESET LiveGrid® aktiviert ist, profitiert von der Genauigkeit unserer ML-Engine und der Fähigkeit unserer Lösungen, selbst neueste Gefahren zu erkennen.

Auch Enterprise-Kunden werden in drei Enterprise-Produkten durch maschinelles Lernen geschützt:

1. Der ESET Enterprise Inspector (EEI) ist ESETs Endpoint Detection and Response (EDR)-Tool. Er sammelt Echtzeitdaten über Aktivitäten auf den verbundenen Endpoints und gleicht sie automatisch mit Regeln ab, die auf verdächtige Aktivitäten hindeuten. Die so gesammelten Informationen werden aufbereitet, gesammelt und in durchsuchbarem Format gespeichert. So entsteht eine per Drill-Down zugängliche Sammlung untypischer und verdächtiger Aktivitäten. Der ESET Enterprise Inspector liefert zudem forensische Daten zu vergangenen Vorfällen und gibt Hinweise zu möglichen Gegenmaßnahmen. So lassen sich selbst sogenannte Advanced Persistent Threats (APTs), die sich bereits im Netzwerk befinden, erfolgreich abwehren. Seine Informationen erhält der ESET Enterprise Inspector dabei aus allen ESET Erkennungstechnologien, u.a. Machine Learning.

2. ESET Dynamic Threat Defense (EDTD) arbeitet mit einer cloudbasierten Sandbox, um neue, bisher unbekannte Gefahren zu identifizieren. Damit ergänzt es die ESET Produkte zur Absicherung von Postfächern und Endpoints um eine weitere Schutzschicht. Die Sandbox besteht aus verschiedensten Sensoren, die die statische Codeanalyse um ML, die Prüfung des Arbeitsspeichers und verhaltensbasierte

Analysen erweitern. Im Vergleich zu den einfachen Endpoint-Lösungen nutzt EDTD damit ein weitaus größeres Spektrum an Technologien und ist so wesentlich besser darin, potentiell gefährliche Samples zu erkennen. Sie greift auf die riesige Sammlung aus bereits klassifizierten Elementen zurück, die im Verlauf der letzten 30 Jahre angelegt wurde. Dabei ist es wesentlich effizienter, diese Art Lösung in der Cloud laufen zu lassen, da sie so wesentlich leichter skalierbar ist und kaum Ansprüche an die Infrastruktur des Nutzers stellt.

3. ESET Threat Intelligence liefert Informationen über tatsächlich bestehende oder aufkommende Gefahren. Damit bildet es eine Art Frühwarnsystem, welches über Schadsoftware oder kriminelle Aktivitäten im Internet informiert – und das in genau auf die Bedürfnisse des Kunden zugeschnittener Form. Die Daten werden durch ML und andere ESET Technologien erhoben, im Anschluss analysiert und in für Menschen lesbarer Form dargestellt. Sicherheitsspezialisten und -analysten sind so in der Lage, frühzeitig auf Gefahren zu reagieren.

Zusätzlich können Nutzer ausgewählte Samples zur weitergehenden Analyse an die ESET Erkennungsengine (ebenfalls mit ML) weitergeben und erhalten einen detaillierten Bericht zurück.

Natürlich werden sich das maschinelle Lernen und seine Anwendungsbereiche auch in Zukunft immer weiterentwickeln. Unsere Experten werden deshalb auch zukünftig daran arbeiten, Machine Learning in weitere Produkte des ESET Portfolios zu integrieren bzw. es noch umfassender einzusetzen.

FAZIT

Es ist kaum vorhersagbar, in welche Richtung sich Machine Learning entwickeln wird und ob die Folgen generell eher positiv oder negativ sein werden. Fest steht, dass sowohl Sicherheitsexperten als auch Kriminelle immer mehr Gebrauch von der Technologie machen werden – mit den entsprechenden Folgen für die Sicherheit des Internets und jedes einzelnen Nutzers.

ESET beobachtet die Aktivitäten krimineller Akteure und ist so in der Lage, aufkommende Bedrohungslagen frühzeitig zu erkennen und seine Abwehrstrategien laufend zu verbessern.

Auch wenn die Marketingaussagen „neuer" Sicherheitsanbieter etwas anderes versprechen – mit drei Jahrzehnten Erfahrung auf dem Gebiet der Cybersecurity können wir mit Sicherheit sagen, dass es nicht reicht, sich auf eine Technologie zu verlassen – und sei es Machine Learning oder Deep Learning.

Nur eine Lösung, die viele verschiedene Ansätze in sich vereint, ist gegen Sabotageversuche immun und erzielt hohe Erkennungsraten bei einer gleichzeitig sehr geringen Anzahl an Fehlalarmen. Nur eine solche Lösung kann im Unternehmensalltag praktikabel eingesetzt werden.

ZUSAMMENFASSUNG

Dank Big Data und verbesserter Rechnerleistung ist Machine Learning (ML) in den letzten Jahren zum Mittel der Wahl für unzählige Anwendungsgebiete geworden – darunter IT-Security. Allerdings gilt hier – wie bei jeder anderen Technologie auch –, dass die Einschränkungen und Nachteile ihres Einsatzes nicht außer Acht gelassen werden dürfen.

Dieses Dokument hat gezeigt, welchen Einfluss der Hype um Machine Learning auf die IT-Sicherheitsbranche und auf IT-Entscheider hatte.

Zusätzlich diskutierten wir tatsächlich durch ESET Experten beobachtete Angriffe, die allem Anschein nach ML einsetzten: Verbesserte Spam-Mails in anderen Sprachen als Englisch und Attacken durch Emotet.

Schließlich zeigten wir, dass ESET bereits seit 20 Jahren maschinelle Lernalgorithmen einsetzt und seine Produkte stets mithilfe von ML verbessert.

ÜBER ESET

ESET® ist ein europäisches Unternehmen mit Hauptsitz in Bratislava (Slowakei). Seit 1987 entwickelt *ESET®* preisgekrönte Sicherheits-Software, die bereits über 110 Millionen Benutzern hilft, sichere Technologien zu genießen. Das breite Portfolio an Sicherheitsprodukten deckt alle gängigen Plattformen ab und bietet Unternehmen und Verbrauchern weltweit die perfekte Balance zwischen Leistung und proaktivem Schutz. Das Unternehmen verfügt über ein globales Vertriebsnetz in über 200 Ländern und Niederlassungen u.a. in Jena, San Diego, Singapur und Buenos Aires. Für weitere Informationen besuchen Sie *www.eset.de* oder folgen uns auf *LinkedIn*, *Facebook* und *Twitter*.

ENJOY SAFER TECHNOLOGY™