



# Machine Learning im Kontext von Cyber Security

Masterarbeit  
zur Erlangung des Grades eines Master of Science (M.Sc.) im Studiengang  
Informationssysteme

vorgelegt von  
Kathi Rodi

Matrikelnummer: 3129378

23. Oktober 2019

---

Erstgutachter: Prof. Dr. Reinhold von Schwerin

Zweitgutachter: Prof. Dr. Markus Schäffter

Betreuer: Hans-Martin Münch

---

# **I. Eigenständigkeitserklärung**

Diese Abschlussarbeit wurde von mir selbständig verfasst. Es wurden nur die angegebenen Quellen und Hilfsmittel verwendet. Alle wörtlichen und sinngemäßen Zitate sind in dieser Arbeit als solche kenntlich gemacht.

---

Kathi Rodi, 23. Oktober 2019

---

## Abstract

Machine Learning Ansätze sind im Kontext von Cyber Security essenziell, da es durch immer anspruchsvoller werdende Sicherheitsbedrohungen nicht mehr möglich ist deren Indikatoren manuell zu ermitteln und zu klassifizieren. Diese Aufgabe von Menschen bearbeiten zu lassen wäre deutlich zu kostenintensiv und zu ineffizient.

Anhand einer Literaturrecherche nach Webster&Watson soll herausgefunden werden welchen Mehrwert Machine Learning in Bezug auf Informationssicherheit bieten kann. Dazu werden bestehende Ansätze aus der Industrie sowohl als auch aus der Wissenschaft klassifiziert, wobei die jeweils verwendeten Algorithmen, Features, Evaluationskriterien sowie die durchgeführte Evaluation der jeweiligen Ergebnisse untersucht werden.

Des Weiteren soll eine Klassifikation bestehender, bzw. eine Formulierung neuer Indicator of Compromise (IoC) zur Unterstützung automatisierter Analysen durchgeführt werden.

Zusätzlich wird recherchiert welche Datensätze in Bezug auf Cyber Security bestehen und welche Qualität diese aufweisen, um einer späteren Machine Learning Anwendung zu dienen.

Um einen übergreifenden Datensatz zusammenführen zu können, welcher unter eine Public License gestellt werden könnte, werden öffentlich zur Verfügung stehende Datensätze analysiert.

Ergänzend werden Diskrepanzen zwischen Ansätzen aus der Industrie und der Wissenschaft überprüft, um etwaige Aussagen aus der Industrie herauszufiltern, welche noch nicht wissenschaftlich belegt wurden. Die Ergebnisse aus diesem Teil der Arbeit könnten zu einem späteren Zeitpunkt dazu dienen, einen Prototypen zu implementieren, welcher einer dieser nicht belegten Aussagen aus der Industrie überprüft. Alternativ könnten gefundene und qualifizierte Datensätze, welche noch keiner Untersuchung unterzogen wurden, einem Prototyp als Basis dienen.

Ziel der Arbeit ist es eine umfassende Übersicht über bestehende Machine Learning Ansätze, welche dabei helfen Indikatoren von Cyberangriffen zu klassifizieren, zu gewinnen und einer dieser gefunden Ansätze beziehungsweise einen gefundenen qualifizierten Datensatz wissenschaftlich zu validieren.

# Inhaltsverzeichnis

I.	Eigenständigkeitserklärung . . . . .	ii
	<b>Inhaltsverzeichnis</b>	<b>iv</b>
<b>1.</b>	<b>Einleitung</b>	<b>2</b>
1.	Motivation . . . . .	2
2.	Ziel der Arbeit . . . . .	4
3.	Aufbau der Arbeit . . . . .	4
<b>2.</b>	<b>Forschungsmethoden</b>	<b>6</b>
1.	Literaturreview . . . . .	6
2.	CRISP-DM . . . . .	7
<b>3.</b>	<b>IOC - Bösertiges Verhalten</b>	<b>10</b>
<b>4.</b>	<b>Analyseverfahren</b>	<b>15</b>
1.	Erkennung von Malware - Hybride Analyse (2015) . . . . .	16
2.	Erkennung von Malware - Statische Analyse (2016) . . . . .	16
3.	Erkennung bössartiger XML-basierter Office Dokumente(2016) . . . .	17
4.	Erkennung bössartiger HTTP-Anfragen (2016) . . . . .	18
5.	Klassifizierung von Netzwerkattacken (2017) . . . . .	18
6.	Erkennung von Ransomware - Dynamische Analyse (2017) . . . . .	19
7.	Erkennung von Malware - Imageanalyse (2017) . . . . .	19
8.	Erkennung böswilliger MS Office Macros (2017) . . . . .	19

9.	Klassifizierung von Distributed Denial of Service (DDoS) Attacken (2017) . . . . .	20
10.	Schwachstellen Scanner für Web Applikationen (2017) . . . . .	20
11.	Erkennung von Malware anhand von PE-Header (2017) . . . . .	20
12.	Erkennung von Malware anhand von PE-Header mit erweitertem Feature-Set (2017) . . . . .	20
13.	Erkennung von Exfiltration und Command & Control (C&C)Tunnels (2017) . . . . .	20
14.	Erkennung bössartiger PowerShell-Befehle (2018) . . . . .	20
15.	Klassifizierung von Netzwerkverkehr in 5 Klassen (2018) . . . . .	21
16.	Anomalieerkennung anhand von Systemprotokollen (2018) . . . . .	21
17.	Erkennung von bössartigem Netzwerkverkehr (2018) . . . . .	21
18.	Erkennung von Botnetzen (2018) . . . . .	21
19.	Klassifizierung von Microsoft Malware (2018) . . . . .	21
20.	Klassifizierung von Malware anhand von Datenpaketen (2018) . . .	21
21.	Erkennung von Portscans (2018) . . . . .	22
22.	Klassifizierung von Netzwerkverkehr in 3 Klassen (2018) . . . . .	22
23.	Erkennung bössartiger SQL-Abfragen (2018) . . . . .	22
24.	Erkennung von Low-rate DDoS (LDDoS) Attacken (2018) . . . . .	22
25.	Klassifizierung von Wi-Fi Netzwerkdaten (2018) . . . . .	22
26.	Klassifizierung von Malware anhand einer systematischen Profilerstellung (2019) . . . . .	22
27.	Klassifizierung von Malware - Imageanalyse (2019) . . . . .	23
28.	Erkennung von Fast-Flux (FF) Netzwerken (2019) . . . . .	23
29.	Erkennung von drive-by Download-Attacken bei Twitter (2019) . .	23
30.	Erkennung von Domain Generation Algorithm (DGA) Domains (2019)	23
31.	Erkennung von Phishing Websites (2019) . . . . .	24
32.	Erkennung von Insider Bedrohungen (2019) . . . . .	24
33.	Erkennung von bössartigen PDFs (2019) . . . . .	24

<b>5. Datensätze</b>	<b>25</b>
<b>6. Prototypische Implementierung</b>	<b>26</b>
<b>7. Diskussion</b>	<b>27</b>
<b>8. Fazit</b>	<b>28</b>
<b>9. Zukünftige Forschung</b>	<b>29</b>
<b>A. Literatur</b>	<b>I</b>
<b>Abbildungsverzeichnis</b>	<b>V</b>
<b>Tabellenverzeichnis</b>	<b>VI</b>
<b>B. Abkürzungsverzeichnis</b>	<b>VII</b>
Abkürzungsverzeichnis . . . . .	VII
<b>C. Anlagen</b>	<b>X</b>
Anlagen . . . . .	X

# 1. Einleitung

Bereits im 19. Jahrhundert träumte der Polymath Charles Babbage vom mechanisierten Rechnen. Dieser Wunsch basierte hauptsächlich auf dem Zorn über die Unzulänglichkeit der damaligen analogen mathematischen Anwendungen. Babbage entwickelte ein Konzept für analytische Maschinen, also einen programmierbaren Allzweckrechner. Seine Kollegin, die britische Mathematikerin, Ada Lovelace lieferte die entsprechenden Ideen zur Programmierung seiner Maschine. Allerdings konnte das Konzept der *Analytical Engine* niemals umgesetzt werden und besteht seither, rein als Entwurf. Dennoch macht diese Forschung die beiden bis heute zu Pionieren des modernen Computers und dessen Programmierung.

Babbages Wunschtraum von damals ist nicht nur längst Wirklichkeit geworden, er hat sich in rasendem Tempo weiterentwickelt. Heute können Computer nicht nur fehlerfrei Logarithmen berechnen, sie sind selbst in der Lage einen Großteil unseres Lebens zu digitalisieren. Bankgeschäfte, Einkäufe, die Steuererklärung und bald auch Arztbesuche sind nur ein kleiner Teil dessen, was wir online erledigen. Dabei produzieren wir eine enorme Masse an persönlichen Daten, welche in falschen Händen, zu einem physischen Schaden für uns führen können. Gerade deshalb gilt es diesen Teil unseres Lebens zu schützen. Wie wir unsere physischen Habseligkeiten schützen in dem wir beispielsweise Schlösser verwenden, gilt es ebenso unsere digitalen Artefakte zu schützen, um finanziellen, reputativen sowie physischen Schaden zu verhindern. Studien zeigen allerdings, dass wir der nötigen Sicherheit weit hinterherhinken.

## 1. Motivation

*Cybercrime umfasst die Straftaten, die sich gegen Datennetze, informationstechnische Systeme oder deren Daten richten [...] oder die mittels Informationstechnik begangen werden. (Bundeskriminalamt 2018)*

Das Bundeskriminalamt (BKA) verzeichnete allein im Jahr 2017 knapp 86.000 Fälle von Cybercrime. Davon waren über 1.400 Phishing Angriffe im Onlinebanking bei

denen ein durchschnittlicher Schaden von 4000 € pro Fall entstand (Bundeskriminalamt 2018). Laut einer Studie des Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V. (BITKOM) ist bereits, jeder zweite Deutsche Opfer eines Cyberangriffs geworden, lediglich 18% hätten diesbezüglich angegeben, Anzeige bei der Polizei erstattet zu haben (e.V. 2017). Dies lässt vermuten, dass die Dunkelziffer der tatsächlichen Cybercrime-Straftaten weit über den 86.000 gemeldeten Fällen liegt. AVTest registriert täglich bis zu 350.000 neue Schadhafte Programme (AV-TEST 2019). Das Problem hierbei ist nicht allein die Quantität der Software sondern auch die Qualität. Immer bessere Verschleierungstaktiken sorgen dafür, dass Malware schwerer identifizierbar wird. Sicherheitsüberprüfungen die auf Signaturabgleichen beruhen, funktionieren beispielsweise nur bei bereits bekannten Signaturen, neuartige Malware kann von ihnen nicht erkannt werden. Diese Komplexität und Fülle an Malware überfordert nicht nur Intrusion Detection Systeme, sondern auch Sicherheitsexperten. Wie schon im Jahre 2010 von Evans und Reeder vorhergesagt, fehlt es an Expertise für diese Flut an Angriffen. Da der Mangel an Fachkräften, wenn überhaupt, erst in Jahren ausgeglichen werden kann, bedarf es alternativer Lösungen für die Sicherheit von Heute.

Machine Learning kann der Schlüssel hierfür sein. Diese Technologie kann für die automatische Verarbeitung von Sicherheitsereignissen genutzt werden. Gängige Warnungen können leicht von Machine Learning Verfahren überprüft werden, dadurch haben Sicherheitsexperten mehr Kapazität sich um besondere Warnungen zu kümmern. Des Weiteren ist es schwierig Warnsignale zusehends zu priorisieren und zu kategorisieren. Auch hierbei können Algorithmen helfen. Beispielsweise lässt sich ein System implementieren, welches eine Klassifizierung in gutartig oder bösartig durchführt. Dabei spricht man von einer *binären* Klassifikation. Gleichzeitig ist es möglich, die als bösartig etikettierten Daten in diverse Kategorien einzustufen. Beispielsweise kann Malware, durch *Multi-Klassen Klassifikation*, in Subklassen wie Viren, Würmer, Trojaner und Ransomware aufgeteilt werden, wodurch die spezifische Untersuchung und Bekämpfung effizienter gestaltet werden kann. Eine weitere Fähigkeit von Machine Learning Algorithmen (MLAs) ist das *Clustering*. Diese Technik fasst grundsätzlich ähnliche Inhalte zusammen. Dabei entstehen Gruppen mit Daten die eine hohe interne Homogenität, verglichen mit anderen Gruppen jedoch eine hohe Heterogenität aufweisen. Clustering kann unter anderem dazu genutzt werden, Hypertext Transfer Protocol (HTTP) Verkehr zu analysieren und herauszufinden, um welche Art von Anfragen es sich handelt. Die Requests können beispielsweise zu Botnet-, Mobiltelefon- oder gängige Benutzeranfragen geclustert werden. Dies stellt eine immense Erleichterung für Sicherheitsexperten dar, da sich diese unmittelbar dem potenziell gefährlichen Cluster widmen können. Machine Learning Algorithmen besitzen die Fähigkeit des eigenständigen Lernens. Da sie so zu neuen Erkenntnissen gelangen und nicht auf bereits bekannte Warnun-



gen, wie beispielsweise bösartige Signaturen, angewiesen sind, können mit ihrer Hilfe sowohl Advanced Persistent Threats (APTs) als auch Zero-days erkannt werden. Anhand der dadurch verringerten Antwortzeit auf Attacken, kann nicht nur ein Verlust von Daten, sondern auch ein finanzieller Schaden stark abgemildert werden.

Forschungen belegen die Wirksamkeit von Machine Learning Ansätzen im Bereich Cyber Security und somit die hier aufgeführten Thesen. Um dies zu verdeutlichen, werden adäquate Untersuchungen in Kapitel 4 *Bestehende Analyseverfahren* beschrieben.

## 2. Ziel der Arbeit

Die Ziele dieser Arbeit belaufen sich auf die folgenden vier Punkte:

1. Zunächst soll der Begriff *Indicator of Compromise* geklärt und in Bezug auf Malware untersucht werden.
2. Des Weiteren soll der momentane Stand der Forschung im Bereich Malware Analyse, mit Hilfe von Machine Learning Verfahren, erörtert werden und etwaige Diskrepanzen mit dem Stand der Industrie aufgedeckt werden.
3. Um eine aussagekräftige Malware Analyse zu tätigen, Bedarf es qualitativ hochwertiger Datensätze. Um einen solchen Datensatz zu ermitteln gilt es, die momentan zur Verfügung stehenden Datensätze zu evaluieren.
4. Ferner soll eine prototypische Umsetzung einer Analyse mit einem der evaluierten Datensätze durchgeführt werden.

Die Arbeit soll somit sowohl Sicherheitsexperten als auch Data Scientists dienen, um einen Überblick über den momentanen Stand der Forschung zu liefern. Zu dem soll es diesem Publikum durch die Evaluierung der Datensätze erleichtert werden, eigenständige neue Analysen durchzuführen oder bestehende Analyseverfahren zu optimieren. Die prototypische Implementierung soll diesbezüglich als Beispiel dienen.

## 3. Aufbau der Arbeit

Das erste Kapitel dient der Einführung in das Thema, wobei zusätzlich die Relevanz der Forschung erläutert wird. Ferner werden die Ziele der Arbeit abgesteckt. Das folgende Kapitel *Forschungsmethode* erläutert das fundierte Vorgehen, durch welches

Informationen generiert und Erkenntnisse erlangt wurden. Anschließend werden im Teil *Theoretische Grundlagen* elementare Begriffe, das Thema betreffend, erläutert, um dem Leser eine einheitliche Wissensbasis zu liefern. Der Hauptteil besteht aus drei Teilen: der Untersuchung der *Bestehenden Analyseverfahren*, der Evaluierung der *Datensätze* sowie der *prototypischen Implementierung* eines Analyseverfahrens anhand eines der evaluierten Datensätze. Im Anschluss wird zu den Ergebnissen kritisch Stellung bezogen, sowie ein Fazit gezogen. Abschließend wird ein Ausblick für potenzielle zukünftige Forschungen gegeben.

## 2. Forschungsmethoden

In diesem Kapitel werden die Forschungsmethoden erläutert, auf welcher der Informationsgewinn basiert und wodurch neue Erkenntnisse gewonnen werden konnten. Das Literaturreview wurde zu Beginn der Arbeit durchgeführt, um Informationen bezüglich des Themas zu sammeln und um den momentanen Stand der Forschung zu identifizieren.

Der Cross-industry standard process for data mining (CRISP-DM) wird als Vorgehensmodell ausgewählt, da hierdurch eine strukturierte Vorgehensweise sichergestellt werden kann. Der genaue Aufbau dieses Modells wird in Kapitel 2 beschrieben.

### 1. Literaturreview

Im Rahmen einer Literaturrecherche nach Webster und Watson wurden die wissenschaftlichen Datenbanken ACM Digital Library, ScienceDirect und IEEE sowie die akademische Suchmaschine Google Scholar nach relevanten Inhalten durchsucht. Hierbei wurde darauf geachtet, dass es sich bei den Ergebnissen, um Peer-Reviewed Journals sowie Peer-Reviewed Konferenzen handelt, um eine bestmögliche Qualität der zu verwendenden Quellen zu garantieren. Ferner wurde lediglich nach Publikationen seit 2015 gesucht, um die Aktualität der Ansätze zu gewährleisten. Da sich besonders im Bereich Cyber Security binnen eines Jahres enorme Entwicklungen zeigen, wäre durch das Hinzuziehen älterer Publikationen kein Mehrwert entstanden. Als Suchstring wurde die logische Kombination aus den Begriffen „Machine Learning“OR „Deep Learning“AND „Cyber Security“OR „Information Security“NOT „Android“NOT „IoT“NOT „Mobile“verwendet. Diese Suche ergab insgesamt 308 Treffer. Zusätzlich wurde sowohl eine Vorwärts - als auch eine Rückwärtssuche durchgeführt, welche zu weiteren 24 Treffern führte. Durch die Rückwärtssuche konnten weitere relevante Ansätze von Machine Learning im Bereich Cyber Security, sowie hilfreiche Informationen zu bestehenden Datensets ausfindig gemacht werden. Auch die Vorwärtssuche, welche mit Google Scholar umgesetzt wurde, führte zu hochaktuellen Beiträgen zum Thema. Von den dadurch 332 ausfindig gemachten Quellen wurden 266 anhand Titel, Abstract, Einleitung und Schluss, in Ermangelung von Relevanz oder wegen Überschneidungen bereits

gefundenen Ansätze, aussortiert. Hingegen wurden 66 Quellen für die hier vorliegende Arbeit verwendet. Eine Übersicht über den Prozess der Literaturrecherche kann im Anhang Anlage 1 eingesehen werden. Wie von Webster und Watson empfohlen, wurden die gefundenen Quellen anschließend akribisch in einer Liste nach Inhalt und Relevanz gefiltert. Zunächst wurden die ausgewählten Quellen in vier Themenblöcke aufgeteilt:

- Ansatz inklusive Datenset
- Ansatz ohne Datenset
- Indicators of Compromise (IoCs)
- Datensatz

Die Quellen wurden anschließend den einzelnen Blocks zugewiesen. Zudem wurden weitere Blocks erstellt die jedoch keinen Einfluss auf die Relevanz der Quelle hatten und somit hier nicht gelistet sind. Zu jeder Quelle wurde die Kernaussage notiert, sowie gegebenenfalls bereits inhaltsrelevante Punkte, wie verwendete Machine Learning Verfahren, Namen von Datensets oder interessante Ergebnisse. Anschließend wurde die Relevanz der Quellen untersucht. Um diesbezüglich ein systematisches Vorgehen zu garantieren wurden folgende Relevanzkriterien erstellt:

1. Hoher Themenbezug zu mindestens einem der Themen: IoCs oder Datensets
2. Ausführung eines Ansatzes
3. Ausführung eines Ansatzes inklusive verfügbarem Datenset

Die Quellen wurden anhand dieser Skala bewertet, wobei 1 für eine geringe Relevanz und 3 für eine hohe Relevanz steht. Zusätzlich wurde die Anzahl der Zitationen festgehalten, um die wissenschaftliche Relevanz innerhalb der Forschungsgemeinde zu evaluieren. Einen Ausschnitt der daraus resultierenden Literaturliste kann im Anhang Anlage 2 eingesehen werden.

## 2. CRISP-DM

Bereits im 18. Jahrhundert legte Thomas Bayes mit seinem *Satz von Bayes*, der die Berechnung bedingter Wahrscheinlichkeiten beschreibt, den Grundstein dafür was wir heute *Data Mining*, also den Erkenntnisgewinn aus Daten, nennen. Als in den 1950er Jahren die Produktion kommerzieller Seriencomputer startete konnte die Datenanalyse automatisiert werden. Daraus entwickelten sich die ersten

Neuronalen Netze und Cluster Analysen wie wir sie heute kennen. Einen weiteren Aufschwung erlebte Data Mining in den 1990ern, wo auch der CRISP-DM von DaimlerChrysler, SPSS und NCR entwickelt wurde (SmartVisionEurope 2015). Dieser Prozess beschreibt eine Methodik für Data Scientists, um eine effiziente, robuste und universelle Vorgehensweise zu garantieren (Chapman u. a. 1999). Wie in 2.1 dargestellt, besteht dieses Vorgehensmodell aus sechs Phasen.

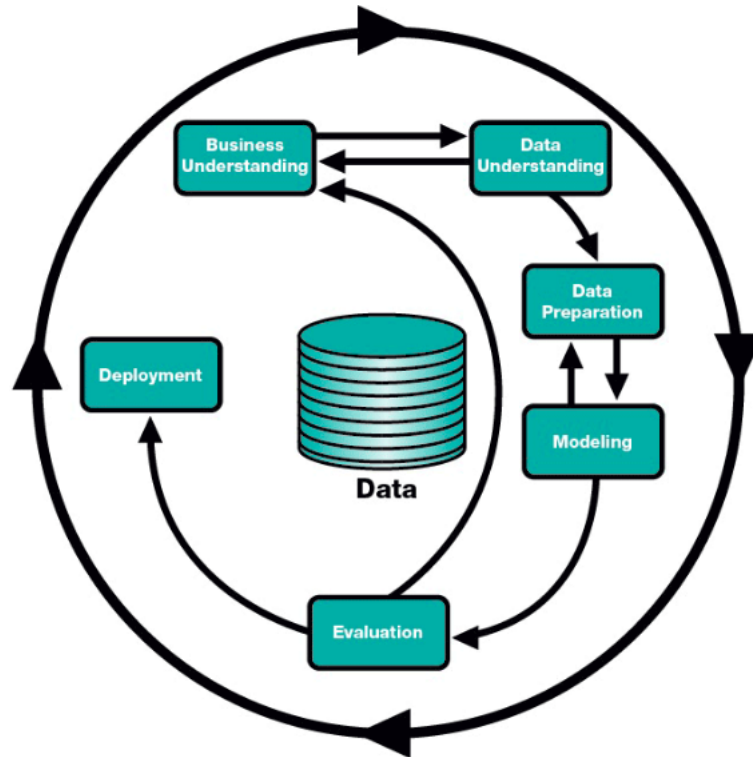


Abbildung 2.1.: CRISP-DM Phasen (SmartVisionEurope 2015)

**Business Understanding:** diese Phase beschäftigt sich mit der Frage nach dem Ziel der Analyse. Dementsprechend, werden die Aufgaben erstellt und ein Plan festgelegt.

**Data Understanding:** die zweite Phase zielt darauf ab Daten zu sammeln und durch ein erstes Screening, deren Qualität festzustellen. Wie die Grafik 2.1 zeigt, kann dies dazu führen, die Ergebnisse aus der ersten Phase noch einmal anzupassen.

**Data Preparation:** nachdem Daten gesammelt wurden, gilt es anschließend diese für Analysen auf zubereiten. Hierbei liegt der Fokus darauf, die bestmögliche Konstruktion des finalen Datensatzes für die anschließende Modellierung zu gewinnen. Dazu ist es nötig relevante Daten auszuwählen und die Daten zu bereinigen. Dazu gehört sowohl das Entfernen und Korrigieren von Datenfehlern, als auch das Schätzen fehlender Daten durch Interpolation beispielsweise.

**Modeling:** diese Phase beschäftigt sich mit der Auswahl einer adäquaten Modellierungstechnik, wie zum Beispiel der Verwendung eines Decision Trees oder eines Neuronalen Netzes. Dazu gehört das kreieren eines Test- und eines Trainingsdatensets, womit verschiedene Modelle getestet werden können. Gegebenenfalls bedarf dies dem Wiederholen der Datenvorbereitung, um ein Datenset nochmals zu justieren.

**Evaluation:** während dieser Phase wird das Modell welches die in Phase eins definierten Ziele am besten erfüllt, ausgewählt.

**Deployment:** in der letzten Phase werden die Ergebnisse aufbereitet und präsentiert und zusätzlich in einem Dokument festgehalten (SmartVisionEurope 2015).

Dieses Vorgehensmodell wurde für diese Arbeit ausgewählt, da es ein strukturiertes Vorgehen ermöglicht und dadurch die Qualität der Ergebnisse gesteigert werden kann. Das *Business Understanding* besteht in dieser Arbeit darin, herauszufinden, was der momentane Stand der Forschung bezüglich Machine Learning im Bereich Cyber Security ist. Anschließend werden bestehende Datensets untersucht, die der späteren Analyse dienen. Um diese Daten anwenden zu können werden diese zunächst in der *Data Preparation* Phase entsprechend aufbereitet. In der nächsten Phase, dem *Modelling* werden diverse Algorithmen auf deren Passgenauigkeit überprüft. Anschließend wird der Algorithmus, welcher die Anforderungen am besten erfüllt, implementiert. Darüber hinaus wird das ganze Vorgehen ausführlich dokumentiert.

F-measure, Precision, Recall

### 3. IOC - Bösartiges Verhalten

Die permanente Steigerung in Größe und Komplexität von Computersystemen, bietet nicht nur einen höheren Nutzen für Kunden, sondern auch mehr Angriffsfläche für Hacker. Dies erschwert die Arbeit von Sicherheitsexperten. Da es darum geht die Kompromittierung eines Systems so früh wie möglich zu erkennen, um potenziellen Schaden zu verhindern, beziehungsweise diesen so gering wie möglich zu halten, arbeiten Experten gegen die Zeit.

Wurde ein System Opfer eines Angriffs, gilt es dieses forensisch zu untersuchen. Normalerweise hinterlässt ein Angreifer Spuren seines Einbruchs. Die Aufgabe der IT-Security ist es, diese zu finden. Diese Hinterlassenschaften werden als *Indicator of Compromise (IoC)* bezeichnet, also Indikatoren, welche darauf hindeuten, dass ein System kompromittiert wurde.

IoCs müssen jedoch differenziert betrachtet werden. Es gibt eindeutige Indikatoren, welche kaum einen Zweifel daran lassen, dass ein System kompromittiert wurde. Angenommen ein Sicherheitsexperte findet Schadsoftware auf einem System und stellt gleichzeitig fest, dass es zu einem Datenupload auf einen nicht identifizierbaren Server kam, welcher von der Malware initiiert wurde. In diesem Fall kann davon ausgegangen werden, dass das System tatsächlich kompromittiert wurde. Der Indikator bildet sich hierbei aus den beiden Indizien: Malware und unautorisierter Upload.

Des Weiteren gibt es Indikatoren welche nicht eindeutig sind. Dieses Phänomen verdeutlicht das folgende Beispiel: angenommen, auf einer Maschine werden Prozesse erkannt, welche nicht von dieser selbst gestartet wurden, sondern durch remote gesendete Befehle. Diese können durch das Windows Tool **PsExec** übermittelt worden sein. Mit Hilfe dessen, lassen sich administrative Tätigkeiten, wie beispielsweise Systemupdates oder Passwort Änderungen, anhand von Remote-Befehlen durchführen. So vorteilhaft dieses Tool in den richtigen Händen erscheint, so gefährlich ist es in den falschen. Angreifer können **PsExec** für bösartige Zwecke missbrauchen. Zwar verlangt der Remote-Zugriff eine IP-Adresse mit korrespondierenden Benutzerinformationen, diese können jedoch durch andere Arten von Angriffen beschaffen werden. Da es sich bei **PsExec** um ein legitimes Tool zur Systemkoordination handelt, wird es von Anti-Viren Programmen nicht erkannt. Dadurch wird die Entdeckung eines Missbrauchs deutlich erschwert. Da die Benutzerinformationen allerdings unverschlüsselt übertragen werden, können immerhin diese über Tools

wie **Wireshark** oder **Tcpdump** abgefangen werden. Der Nachweis über die Nutzung dieses Tools allein reicht also nicht aus, um eine Kompromittierung annehmen zu können.

Bei der Malware spezifischen Analyse gilt es zunächst herauszufinden, was genau passiert ist und welches Schadprogramm für den Angriff verantwortlich ist. Traditionelle Anti-Virus Programme arbeiten basierend auf Datenbanken, in welchen sie bereits bekannte Signaturen und Heuristiken anwenden, um Malware zu identifizieren. Das Problem hierbei ist, dass es für Angreifer ein leichtes ist, ihren Code zu modifizieren, um die Signatur zu verändern, wodurch das Schadprogramm nicht mehr als solches erkannt wird. Verschleierungstaktiken wie diese, lassen sich in drei Gruppen einteilen (P. He u. a. 2017):

- **Packing** Dies Bezeichnet die Technik exekutierbare Dateien zu komprimieren. Um die komprimierte Malware zu erkennen muss diese zunächst entpackt werden. Gleichzeitig ist dies aber auch ein guter IoC, da ausführbare Dateien im Regelfall nicht komprimiert vorliegen.
- **Metamorphismus** Hierbei wird die Erkennung erschwert in dem der Binär-code mutiert wird. Das bedeutet, die Sequenz der Opcodes wird bei jeder Ausführung geändert.
- **Polymorphismus** Eine polymorphe Schadsoftware generiert bei jeder Ausführung eine weitere Version der Malware, sodass eine große Anzahl an divergierender Signaturen für dasselbe Programm entstehen.

Diese Techniken erschweren das Erkennen von Malware anhand gängiger Anti-Virus Programme deutlich. Zukünftig kann Machine Learning hierbei eine große Rolle spielen. Denn wie Han u. a. (2019) bereits erfolgreich untersuchten, ist auf MLAs basierende Erkennungssoftware in der Lage, Malware trotz dieser Verschleierungstechniken zu konstatieren.

Das Erkennen von Malware basiert im Regelfall auf der Untersuchung von Portable Executable Dateien (PE-Dateien). Diese beinhalten ausführbare Daten im Binärformat. Dazu gehören Windows **.exe** Dateien, Objektcode und Dynamic Link Libraries (DLLs). Eine PE-Datei ist folgendermaßen aufgebaut:





Abbildung 0.1.: Aufbau einer PE-Datei (eigene Darstellung)

Die in Abbildung 0.1 grau hinterlegten Bereiche sind für die Analyse von PE-Dateien irrelevant, sie dienen unter anderem lediglich dazu, eine Fehlermeldung auszugeben, falls eine **.exe** Datei in einem Betriebssystem ausgeführt werden soll, mit welchem diese nicht kompatibel ist.

Der Bereich **IMAGE\_NT\_HEADERS** bietet bereits Informationen für eine Analyse. **IMAGE\_FILE\_HEADER** beinhaltet grundlegende Informationen bezüglich der Datei. Beispielsweise wann diese ausgeführt wurde, was einer Analyse sehr nützlich sein kann. Der Sektor **IMAGE\_OPTIONAL\_HEADER** ist entgegen dem was der Name vermuten lässt nicht *optional*. Hier werden wichtige Informationen wie der ProgrammEinstiegspunkt, die Stackgröße zu Beginn sowie die Verwendung eines Graphical User Interface (GUI) (dt. *Grafische Benutzeroberfläche*) oder einer Konsole definiert.

Die grün hinterlegten **IMAGE\_SECTION\_HEADER** bieten die interessantesten Informationen für eine Analyse. Diese *Header* werden vom Compiler generiert und benannt, sodass der Benutzer wenig Kontrolle über die Namen hat. Dementsprechend konsistent ist die Benennung im Regelfall. Im PE-Header finden sich also relevante Informationen wie Imports, Exports, die Namen der verschiedenen Bereiche (blau

hinterlegt), sowie deren Speichergröße auf der Festplatte und im Random-Access Memory (RAM), sowie die Ressourcen welche von einem Programm benötigt werden.

Grundsätzlich gibt es zwei Methoden um eine Malware Analyse durchzuführen: eine statische und eine dynamische.

Die Dynamische Analyse beinhaltet das Ausführen Schadhafter Programme. Dabei wird Malware in einer sicheren Umgebung ausgeführt und so dessen Verhalten analysiert. Dadurch kann im Gegensatz zur statischen Analyse die tatsächliche Verhaltensweise einer Datei untersucht werden, denn nicht jeder String der in einer Binärdatei gefunden wird, muss zwangsläufig ausgeführt werden. Des Weiteren können Logdateien analysiert werden, welche erst durch das Ausführen eines Programms entstehen. Dynamische Analysen werden im Regelfall in einer *Sandbox*, also in einem isolierten Bereich durchgeführt, wodurch kein Schaden an der Umgebung genommen wird. Der Nachteil dieser Analyse besteht darin, dass die Malware die virtuelle Umgebung erkennen kann und sich somit stoppt. Des Weiteren können von der Schadsoftware benötigte Registry Keys oder Dateien in der virtuellen Umgebung fehlen, sodass deren Verhalten nicht korrekt aufgezeichnet werden kann.

Eine weitere Methode zur Malware Untersuchung bietet eine statische Analyse. Hierbei werden PE-Dateien erforscht. Zunächst durchläuft potenzielle Malware diverse Virens Scanner, um die Entdeckung einer bösartigen Signatur zu erhöhen.

Des Weiteren kann *Hashing* zum Einsatz kommen. Dabei wird ein eindeutiger *hash* generiert, welcher verwendet werden kann, um zu recherchieren, ob dieser bereits von anderen AV-Dienstleistern analysiert wurde. Hashing bietet zudem den Vorteil, dass die Datei selbst noch nicht geteilt werden muss. Des Weiteren ist der Austausch eines Hashes auch um einiges schneller als der Upload einer Datei.

Programme verwenden Strings beispielsweise um sich mit einer URL verbinden zu können oder um Textausgaben zu drucken. Diese können einen guten Einblick über das Verhalten einzelner Programme liefern. Beispielsweise können dadurch IP Adressen für Command and Control (C2)-Systeme identifiziert werden, welche von Angreifern zum Verwalten von Remotesitzungen von infizierten Hosts verwendet werden.

Handelt es sich bei Dateien um sogenannte *Packer*, also komprimierte Programme, müssen diese zunächst entpackt werden, um eine erfolgreiche Analyse durchführen zu können. Bei Dateien die relativ wenige Strings enthalten, handelt es sich meist um Packer. Wie die nachfolgende Tabelle zeigt, bietet die statische Analyse eine Vielzahl an Indikatoren, welche darauf hinweisen können, dass es sich bei der untersuchten Datei um Schadsoftware handelt.

Ort	Indikator	Verhalten
System	neue/modifizierte Dateien	Veränderung des Dateisystems durch Malware
System	Registry Einträge	Veränderung/Erstellung von Registry Keys
PE Header	wenige Imports	durch Packer komprimierte Dateien, um die Erkennung und Analyse zu erschweren
PE Header - Imports	<b>SetWindowsHookEx</b>	empfängt Tastatureingaben (Keylogger)
PE Header - Imports	<b>RegisterHotKey</b>	bestimmte Tastenkombination startet Anwendung (Keylogger)
IMAGE_FILE_HEADER	Kompilierungszeit	unsinnige Kompilierungszeit ist verdächtig
Sections	Abweichende Namen	z.B. .srtsa anstatt .data
SECTION .text	divergierende Speichergröße von Virtual Size und Raw Size	Packer extrahiert Code nach .text
SECTION .rsrc	eingebettetes Programm, Treiber	weitere durch Malware gestartete Aktionen

Tabelle 3.1.: Auszug an, durch statische Analysen identifizierte Indikatoren für einen Malware Angriff (In Anlehnung an Sikorski (2012))

Diese Indikatoren aus Tabelle 3.1, sind nur ein kleiner Teil dessen, was bei einer statischen Analyse entdeckt werden kann. Dennoch wird dadurch deutlich, wie hilfreich PE-Dateien im Erkennen von Malware sein können.

Welche Features aus diesen Dateien generiert werden können, um eine aussagekräftige Untersuchung mit Hilfe von MLAs zu generieren, wird in diversen Ansätzen beschrieben, mit welchen sich das folgende Kapitel beschäftigt.

## 4. Analyseverfahren

Das folgende Kapitel beschäftigt sich mit Machine Learning Verfahren, welche für die Erkennung von Cyber Security Angriffen verwendet werden. Diese Verfahren wurden anhand einer umfangreichen Literaturrecherche ermittelt. Jedes Vorgehen wird auf dessen verwendete Algorithmen sowie der ausgewählten Features untersucht. Des Weiteren wird analysiert welche Evaluationskriterien verwendet wurden um die Effektivität zu messen und zu welchem Ergebnis der jeweilige Ansatz führte. Das Vorgehen hierbei, orientiert sich an einem problembezogenen Ansatz. Dazu wird im Folgenden eine Übersicht darüber generiert, welche Probleme aus der Cyber Security bereits von Machine Learning Algorithmen in Angriff genommen wurden. Dies dient einer übersichtlichen Darstellung der Möglichkeiten und zeigt die Diversität auf, in welcher MLAs die Arbeit von Cyber Security Spezialisten unterstützen und verbessern können.

Zunächst werden Verfahren erläutert, welche bereits wissenschaftlich belegt wurden. Wurden, dem Ansatz entsprechende Adaptionen aus der Industrie identifiziert, wurde der jeweilige Ansatz um den industriellen erweitert.

Für die Recherche wurden alle Verfahren in einem Zeitraum von 2015 bis heute berücksichtigt. Diese Periode wurde gewählt, da sich die Zahl der Cyberangriffe, sowie die zur Verfügung stehende Schadsoftware bereits innerhalb weniger Jahre deutlich vermehrt beziehungsweise verändert. Somit soll verhindert werden veraltete Angriffsvektoren zu analysieren, welche bereits von moderneren überholt wurden. Zusätzlich gibt es bereits vergleichbare Arbeiten aus dem Jahr 2016 (s. Buczak und Guven 2016), in welchem Machine Learning Ansätze vor dieser Zeit analysiert werden.

In der folgenden Auflistung wird die Bezeichnung „Erkennung“ für binäre Klassifikation verwendet. Beispielsweise, wenn sich ein Ansatz darauf beschränkt Daten entweder in die Rubrik A „böartig“ oder in die Rubrik B „gutartig“ zu klassifizieren. Erfolgt in einem Analyseverfahren eine Einteilung in mehrere Klassen (mehr als zwei), wird nachfolgend der Begriff „Klassifizierung“ verwendet.

## 1. Erkennung von Malware - Hybride Analyse (2015)

Im Jahr 2015 haben Shijo und Salim einen, auf zwei Analysen basierenden, Ansatz gewählt, um Malware zu erkennen. Zum einen verwendeten sie ein statisches Analyseverfahren bei dem sie *Printable Strings*, also nicht kodierte Zeichenfolgen wie z. B. *FindFirstFile* aus Binärdateien extrahierten. Zum anderen konfigurierten sie eine Cuckoo Sandbox, in der sie Schadsoftware ausführten und deren API Aufrufe in einer Logdatei speicherten.

Sie untersuchten die Ähnlichkeit in API-Aufrufsequenzen anhand von n-Gramm-basierter Ähnlichkeitsmessung. Als Features dienten Tri- und Tetragramme ab einer gewissen Häufigkeit, sowie PrintableStrings ab einer Häufigkeit von zwei. Für die Klassifizierung wurden die Algorithmen Random Forest (RF) und Support Vector Machine (SVM) verwendet. Es wurden jeweils beide Ansätze separat, sowie in Kombination getestet. Analysen mit SVM erzielten eine Genauigkeit von 95.88 % für die statische Analyse und 97.16 % für die dynamische Analyse und waren somit erfolgreicher, als Untersuchungen mit Random Forest. Die besten Ergebnisse erzielte der hybride Ansatz mit SVM mit einer Genauigkeit von 98.71 % und der geringsten False Positive Rate (FPR) von 0.026.

Die Forschung von Shijo und Salim zeigt also, dass mit den von ihnen gewählte Features, mit einem hybriden Ansatz, deutlich genauere Aussagen, als mit rein statischen oder rein dynamischen Analysen, getroffen werden können.

## 2. Erkennung von Malware - Statische Analyse (2016)

More und Gaikwad untersuchten EXE-Dateien auf Schadsoftware. Dazu konvertierten sie die Dateien zunächst in Operation Code (Opcode), also in den Teil der Maschinensprachanweisung der die auszuführenden Operationen angibt, z.B. 55 8B EC 83 EC 5C 83 7D 0C 0F 74 2B 83 7D 0C 46. Das ausgewählte Feature Datenset wurde anschließend nochmals zu einer Attribute-Relation File Format (ARFF) Datei konvertiert, um die Datei nachfolgend mit der Machine Learning Software Weka bearbeiten zu können. In Weka wurden die Algorithmen JRip, C4.5 und IBk verwendet. Wobei es sich bei JRip und C4.5 um Decision Tree (DT) und bei IBk um k-nearest-neighbor (k-NN) Implementierungen handelt. Um die Erkennungsgenauigkeit zu erhöhen, wurden nicht die einzelnen Algorithmen, sondern ein Klassifikatorenensemble angewandt, um Methoden wie Mehrheitsvoting, Veto-Voting und vertrauensbasiertes Veto-Voting verwenden zu können. Ersteres folgt demokratischen Regeln, das heißt, die Klasse mit den meisten Stimmen ist das

Ergebnis. Veto-Voting hingegen basiert auf Annahmen über die Wahl der anderen Algorithmen. Vertrauensbasiertes Veto-Voting ergänzt voriges Voting um eine Vertrauensberechnung, wodurch jedem Algorithmus ein bestimmtes Vertrauensniveau zugeteilt wird.

More und Gaikwad konnten zeigen, dass durch die Verwendung von Veto-Voting eine Genauigkeit von 80.7 % erzielt werden kann. Im Vergleich dazu, lag das beste Ergebnis, welches durch singulären Algorithmeinsatz von IBk erzielt wurde, bei einer Genauigkeit von nur 73.5 %.

Dieses Ergebnis stützt die These von Shijo und Salim aus dem Jahr zuvor, welche zeigten, dass ein nicht-hybrider Ansatz weniger genau ist, als einer, der die statische und die dynamische miteinander verknüpft.

### **3. Erkennung bössartiger XML-basierter Office Dokumente(2016)**

Durch das neue Dateiformat, welches Microsoft 2007 auf den Markt gebracht hat, sollten Sicherheitslücken geschlossen werden. Das Binärformat wurde durch ein XML basiertes Dateiformat ersetzt. Dadurch werden neue digitale Funktionen unterstützt, sowie vertrauensbasierte Bereiche geschaffen, welche das Format weniger riskoreich gestalten sollen. Nichtsdestotrotz können Attacks gegen XML-basierte Office Dokumente gestartet werden. Zu den möglichen Angriffsvektoren zählen beispielsweise macrobasierte Attacks. Durch den Missbrauch von Visual Basic for Applications (VBA) kann die zugehörige shell gestartet werden, um willkürliche Kommandos zu senden. Außerdem können externe Bibliotheken sowie Programme aufgerufen werden, welche Schaden verursachen können. Eine weitere Bedrohung durch den Gebrauch von Macros bildet die Fähigkeit dieser, bössartige Dateien aus dem Internet herunterzuladen.

Cohen u. a. (2016) haben in ihrer Arbeit diese Art von Attackvektoren untersucht, um eine Infizierung durch bössartige Macros frühzeitig zu erkennen. Dazu haben sie zunächst Office Dokumente in eine Liste von Pfaden konvertiert. Diese dienen der Analyse als Features. Dadurch wurde jedoch eine so hohe Anzahl an Features generiert, dass die Untersuchung mit verschiedenen Datensets durchgeführt wurde, welche Top Features von 10 bis 2000 beinhalteten. Um die Feature-Repräsentation, also das Vorhandensein bzw. die Wichtigkeit von Features zu bestimmen, wurden zwei Verfahren angewandt. Zum einen ein binäres Verfahren, welches lediglich die Ab-, respektive Anwesenheit eines Features misst und zum anderen wurde das statistische Verfahren Term Frequency - Inverse Document Frequency (TF-IDF) verwendet,

um die Wichtigkeit eines Terms in Bezug auf ein Dokument zu bestimmen. Anschließend wurden die Daten mit folgenden Algorithmen untersucht: J48, RF, Logistic Regression (LR), Naïve Bayes (NB), Bayesian Network (BN), LogitBoost (LB), Sequential Minimal Optimization (SMO), Bagging und AdaBoost (AB).

Wie die Ergebnisse zeigen, erzielt das Datenset mit den Top 200 Features, welches mit Random Forest analysiert wurde die besten Werte mit einem F-measure von 0.66. Wie sich demonstrieren ließ, ist es möglich bösartige Office Dokumente durch eine Analyse deren Pfade zu erkennen. Die Untersuchung beschränkt sich jedoch auf direkte Gefahren innerhalb von Dokumenten. Indirekte Gefahren, wie etwa die durch weiterführende Links, wurden in dieser Forschungsarbeit nicht berücksichtigt.

## 4. Erkennung bösartiger HTTP-Anfragen (2016)

Pham, Hoang und Vu (2016) haben in ihrer Arbeit ein Intrusion Detection System (IDS) für HTTP - Anfragen entwickelt. Dazu nutzen sie ein von ihnen entwickeltes Modul, um Netzwerkpakete zu erfassen. Dieses Modul basiert auf derselben Erfassungstechnik die **Wireshark** verwendet. Um ein geeignetes Model zu zu evaluieren, welches es ermöglicht die Pakete in Echtzeit zu klassifizieren, wurden diverse MLAs anhand des CSIC 2010 HTTP Datensets von Carmen Torrano Giménez, Alejandro Pérez Villegas (2010) getestet. Dieses besteht aus 223585 Daten die entweder als *normal* oder *anomal* beschriftet sind. Es enthält Attacks wie SQL-Injections, Buffer Overflow und Cross-site scripting (xss). Pham, Hoang und Vu trainierten und testeten die Daten mit den Algorithmen Decision Tree, Random Forest, AdaBoost, Logistic Regression und dem SGDClassifier. Als Evaluationsmethode wurden Precision, Recall und F-measure verwendet. Den höchsten dieser Werte erzielte die Logistic Regression Methode mit einem F1-measure von 96.11 for abnormen und 95.82 für normalen Verkehr. Zukünftig sollen diese Ergebnisse an Paketen, welche durch das eigens entwickelte Modul der Forscher erfasst wurden, evaluiert werden.

## 5. Klassifizierung von Netzwerkattacken (2017)

Yin u. a. (2017) implementierten ein IDS, für welches sie Recurrent neural networks (RNNs) verwendeten. Zusätzlich wurde die Leistung des Models bei binärer als auch bei Multi-Klassen Klassifikation untersucht. Um die Effizienz zu prüfen, wurde ferner ein Vergleich mit diversen MLAs gezogen. Die Analyse basiert auf dem

NSL-KDD Datenset aus dem Jahr 2009. Dieses beinhaltet neben dem normalen Netzwerkverkehr, Daten zu vier verschiedenen Angriffstypen die da wären: Denial of Service (DoS), Remote to Local (R2L), User to Root (U2R) und Probe Attacks. Das Datenset besteht aus 41 Features. Um Vergleiche mit anderen MLAs herzustellen, wurden parallel Experimente mit Artificial Neural Network (ANN), RF, NB, Multi-Layer Perceptron (MLP), SVM, J48, Random Tree und Naïve Bayesian Tree (NBTree) für die binäre Klassifikation durchgeführt. Auf dieselbe Weise wurde die Multi-Klassen Klassifikation überprüft. Als Qualitätsmerkmal der Ergebnisse wurde der Wert *Accuracy* gewählt, welcher die Genauigkeit der Analyse wider spiegelt. Dieser Wert basiert auf dem Prozentsatz der korrekt klassifizierten Daten im Vergleich zur Gesamtheit der Daten. Das Ergebnis zeigt, dass RNNs eine qualitativ hochwertigere Analyse produzieren als die zu verglichenen MLAs. Bei der binären Klassifikation des Testsets erzielte auf RNNs basierende Untersuchungen eine Genauigkeit von 83.28% gefolgt von NBTree mit 82.02%. Bei der Klassifizierung in fünf Klassen erreichte das Recurrent neural network (RNN) mit 81.29% ebenfalls ein besseres Ergebnis als die restlichen Algorithmen. Jedoch fanden Yin u. a. heraus, dass RNNs deutlich mehr Zeit für das Training beanspruchen, dieses Problem soll zukünftig durch Beschleunigung der Graphics processing unit (GPU) behoben werden.

## 6. Erkennung von Ransomware - Dynamische Analyse (2017)

Maniath u. a. (2017)...

## 7. Erkennung von Malware - Imageanalyse (2017)

Roma2017...

## 8. Erkennung böswilliger MS Office Macros (2017)

Bearden und Lo (2017)....



## **9. Klassifizierung von DDoS Attacken (2017)**

Z. He, Zhang und Lee (2017)...

## **10. Schwachstellen Scanner für Web Applikationen (2017)**

Robin Tommy, Gullapudi Sundeeep (2017)...

## **11. Erkennung von Malware anhand von PE-Header (2017)**

Raff, Sylvester und Nicholas (2017)...

## **12. Erkennung von Malware anhand von PE-Header mit erweitertem Feature-Set (2017)**

Kumar, Kuppusamy und Aghila (2017)...

## **13. Erkennung von Exfiltration und C&CTunnels (2017)**

Das u. a. (2017)...

## **14. Erkennung bössartiger PowerShell-Befehle (2018)**

Hendler, Kels und Rubin (2018)...

## **15. Klassifizierung von Netzwerkverkehr in 5 Klassen (2018)**

Ding und Zhai (2018)...

## **16. Anomalieerkennung anhand von Systemprotokollen (2018)**

Brown u. a. (2018)...

## **17. Erkennung von böartigem Netzwerkverkehr (2018)**

Aldwairi, Perera und Novotny (2018)...

## **18. Erkennung von Botnetzen (2018)**

Mathur, Raheja und Ahlawat (2018)....

## **19. Klassifizierung von Microsoft Malware (2018)**

Sabar, Yi und Song (2018)....

## **20. Klassifizierung von Malware anhand von Datenpaketen (2018)**

Yeo u. a. (2018)...

## **21. Erkennung von Portscans (2018)**

Aksu und Ali Aydin (2018)

## **22. Klassifizierung von Netzwerkverkehr in 3 Klassen (2018)**

Teoh u. a. (2018)...

## **23. Erkennung böartiger SQL-Abfragen (2018)**

Jayaprakash und Kandasamy (2018)...

## **24. Erkennung von LDDoS Attacken (2018)**

Siracusano, Shiaeles und Ghita (2018)...

## **25. Klassifizierung von Wi-Fi Netzwerkdaten (2018)**

Qin u. a. (2018)....

## **26. Klassifizierung von Malware anhand einer systematischen Profilerstellung (2019)**

P. He u. a. (2017) Malware Verschleierung durch packing, Metamorphism und Polymorphism.

Han u. a. (2019)...

..... Neueste Forschungen zeigen, dass dynamische und statische Analyseverfahren zu ungenau sind, um neue Schadsoftware in Echtzeit zu erkennen (Vinayakumar u. a. 2019). Dazu bedarf es Deep Learning Verfahren, wie der folgende Ansatz verdeutlicht.

## **27. Klassifizierung von Malware - Imageanalyse (2019)**

Um Polymorphismus, Metamorphismus und Packing zu erkennen ist ein umfangreiches Feature Engineering, sowie beträchtliche Kenntnisse auf Domain-Ebene nötig (Rhode, Burnap und Jones 2018). Zudem können Angreifer der automatischen Malware Erkennung entgehen sobald sie die verwendeten Features kennen (Anderson u. a. 2017). Diesen Problemen wollen Vinayakumar u. a. (2019) mit ihrem Deep Learning Ansatz begegnen.

Dazu verglichen sie klassische Algorithmen für maschinelles Lernen mit Deep Learning Architekturen. Die Vergleiche basieren auf statischen und dynamischen Analysen, sowie auf Bildverarbeitungstechnologien.

...

Wie sich zeigte ist die Imageanalyse schneller als die statische und die dynamische Analyse, da diese auf Rohdaten basiert und komplett auf Zerlegung oder Ausführung von Code verzichtet.

## **28. Erkennung von FF Netzwerken (2019)**

Chen u. a. (2019)...

## **29. Erkennung von drive-by Download-Attacken bei Twitter (2019)**

Javed, Burnap und Rana (2019)... Verschleierte Links können dazuführen dass der Angreifer Fernzugriff zum System des Opfers bekommt von welchem er Daten klauen kann oder dessen Computer in ein Botnetz integrieren kann (Provos u. a. 2007).

## **30. Erkennung von DGA Domains (2019)**

Li u. a. (2019)...

## **31. Erkennung von Phishing Websites (2019)**

Phishing zielt im Vergleich zu anderen Attacken nicht darauf ab Schwachstellen im System auszunutzen, sondern durch Sicherheitslücken beim Menschen an dessen sensitive Informationen wie Benutzernamen und Passwörter zu gelangen. In der Forschung gibt es momentan vier Verfahren, um Phishing Websites zu erkennen: Blacklists, Heuristiken, Inhaltsanalysen und Machinelles Lernen. Blacklists gleichen URLs mit bekannten Phishing Websites ab, Heuristiken verwenden Signaturdatenbanken bekannter Angriffe um sie mit der Signatur eines heuristischen Musters abzugleichen. Inhaltsanalysen versuchen Phishing Websites mit Hilfe bekannter Algorithmen wie TF-IDF zu identifizieren.

Der im Folgende beschriebene Ansatz von Alswailem u. a. (2019) verwendet Machine Learning Verfahren, um Phishing Websites zu erkennen.

## **32. Erkennung von Insider Bedrohungen (2019)**

Le und Nur Zincir-Heywood (2019)...

## **33. Erkennung von böartigen PDFs (2019)**

Jeong, Woo und Kang (2019)...

## 5. Datensätze

## **6. Prototypische Implementierung**

## **7. Diskussion**



## **8. Fazit**

## **9. Zukünftige Forschung**

# A. Literatur

## Bücher und Journals

- Aksu, D. und M. Ali Aydin (2018). “Detecting Port Scan Attempts with Comparative Analysis of Deep Learning and Support Vector Machine Algorithms”. In: *2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT)*. IEEE, S. 77–80 (siehe S. 22).
- Aldwairi, T., D. Perera und M. A. Novotny (2018). “An evaluation of the performance of Restricted Boltzmann Machines as a model for anomaly network intrusion detection”. In: *Computer Networks* 144, S. 111–119 (siehe S. 21).
- Alswailem, A. u. a. (Mai 2019). “Detecting Phishing Websites Using Machine Learning”. In: *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*. IEEE, S. 1–6 (siehe S. 24).
- Anderson, H. S. u. a. (2017). “Evading machine learning malware detection”. In: *Black Hat* (siehe S. 23).
- Bearden, R. und D. C.-T. Lo (2017). “Automated microsoft office macro malware detection using machine learning”. In: *2017 IEEE International Conference on Big Data (Big Data)*. Bd. 2018-Janua. IEEE, S. 4448–4452 (siehe S. 19).
- Brown, A. u. a. (2018). “Recurrent Neural Network Attention Mechanisms for Interpretable System Log Anomaly Detection”. In: *Proceedings of the First Workshop on Machine Learning for Computing Systems*, S. 1 (siehe S. 21).
- Buczak, A. L. und E. Guven (2016). “A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection”. In: *IEEE Communications Surveys & Tutorials* 18.2, S. 1153–1176 (siehe S. 15).
- Bundeskriminalamt (2018). “Cybercrime, Bundeslagebild 2017”. In: (Siehe S. 2, 3).
- Chapman, P. u. a. (1999). “The CRISP-DM user guide”. In: *4th CRISP-DM SIG Workshop in Brussels in March*. Bd. 1999 (siehe S. 8).
- Chen, X. u. a. (2019). “A Deep Learning Based Fast-Flux and CDN Domain Names Recognition Method”. In: *Proceedings of the 2019 2nd International Conference on Information Science and Systems - ICISS 2019*. Bd. Part F1483. New York, New York, USA: ACM Press, S. 54–59 (siehe S. 23).
- Cohen, A. u. a. (2016). “SFEM: Structural feature extraction methodology for the detection of malicious office documents using machine learning methods”. In: *Expert Systems with Applications* 63, S. 324–343 (siehe S. 17).

- Das, A. u. a. (Dez. 2017). “Detection of Exfiltration and Tunneling over DNS”. In: *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. Bd. 2018-Janua. IEEE, S. 737–742 (siehe S. 20).
- Ding, Y. und Y. Zhai (2018). “Intrusion Detection System for NSL-KDD Dataset Using Convolutional Neural Networks”. In: *Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence - CSAI '18*. New York, New York, USA: ACM Press, S. 81–85 (siehe S. 21).
- e.V., B. (2017). *Cybercrime: Jeder zweite Internetnutzer wurde Opfer*. URL: <https://www.bitkom.org/Presse/Presseinformation/Cybercrime-Jeder-zweite-Internetnutzer-wurde-Opfer.html> (besucht am 01.10.2019) (siehe S. 3).
- Evans, K. und F. S. Reeder (2010). *A human capital crisis in cybersecurity : technical proficiency matters : a report of the CSIS Commission on Cybersecurity for the 44th Presidency*. November, S. 35 (siehe S. 3).
- Han, W. u. a. (Jan. 2019). “MalInsight: A systematic profiling based malware detection framework”. In: *Journal of Network and Computer Applications* 125.June 2018, S. 236–250 (siehe S. 11, 22).
- He, P. u. a. (2017). “Model approach to grammatical evolution: deep-structured analyzing of model and representation”. In: *Soft Computing* 21.18, S. 5413–5423 (siehe S. 11, 22).
- He, Z., T. Zhang und R. B. Lee (2017). “Machine Learning Based DDoS Attack Detection from Source Side in Cloud”. In: *2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud)*. IEEE, S. 114–120 (siehe S. 20).
- Hendler, D., S. Kels und A. Rubin (2018). “Detecting Malicious PowerShell Commands using Deep Neural Networks”. In: *Proceedings of the 2018 on Asia Conference on Computer and Communications Security - ASIACCS '18*. New York, New York, USA: ACM Press, S. 187–197 (siehe S. 20).
- Javed, A., P. Burnap und O. Rana (Mai 2019). “Prediction of drive-by download attacks on Twitter”. In: *Information Processing & Management* 56.3, S. 1133–1145 (siehe S. 23).
- Jayaprakash, S. und K. Kandasamy (Apr. 2018). “Database Intrusion Detection System Using Octaplet and Machine Learning”. In: *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*. Iicct. IEEE, S. 1413–1416 (siehe S. 22).
- Jeong, Y.-S., J. Woo und A. R. Kang (Apr. 2019). “Malware Detection on Byte Streams of PDF Files Using Convolutional Neural Networks”. In: *Security and Communication Networks* 2019, S. 1–9 (siehe S. 24).
- Kumar, A., K. Kuppasamy und G. Aghila (2017). “A learning model to detect maliciousness of portable executable using integrated feature set”. In: *Journal*

- of King Saud University - Computer and Information Sciences 31.2, S. 252–265 (siehe S. 20).
- Le, D. C. und A. Nur Zincir-Heywood (2019). “Machine learning based insider threat modelling and detection”. In: *2019 IFIP/IEEE Symposium on Integrated Network and Service Management, IM 2019*, S. 1–6 (siehe S. 24).
- Li, Y. u. a. (2019). “A Machine Learning Framework for Domain Generation Algorithm-Based Malware Detection”. In: *IEEE Access* 7, S. 32765–32782 (siehe S. 23).
- Maniath, S. u. a. (2017). “Deep learning LSTM based ransomware detection”. In: *2017 Recent Developments in Control, Automation & Power Engineering (RDCAPE)*. Bd. 3. IEEE, S. 442–446 (siehe S. 19).
- Mathur, L., M. Raheja und P. Ahlawat (2018). “Botnet Detection via mining of network traffic flow”. In: *Procedia Computer Science* 132, S. 1668–1677 (siehe S. 21).
- More, S. S. und P. P. Gaikwad (2016). “Trust-based Voting Method for Efficient Malware Detection”. In: *Procedia Computer Science* 79, S. 657–667 (siehe S. 16, 17).
- Pham, T. S., T. H. Hoang und V. C. Vu (2016). “Machine learning techniques for web intrusion detection — A comparison”. In: *2016 Eighth International Conference on Knowledge and Systems Engineering (KSE)*. IEEE, S. 291–297 (siehe S. 18).
- Provos, N. u. a. (2007). “The Ghost in the Browser: Analysis of Web-based Malware.” In: *HotBots* 7, S. 4 (siehe S. 23).
- Qin, Y. u. a. (Sep. 2018). “Attack Detection for Wireless Enterprise Network: a Machine Learning Approach”. In: *2018 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. IEEE, S. 1–6 (siehe S. 22).
- Raff, E., J. Sylvester und C. Nicholas (2017). “Learning the PE Header, Malware Detection with Minimal Domain Knowledge”. In: *Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security - AISec '17*. New York, New York, USA: ACM Press, S. 121–132 (siehe S. 20).
- Rhode, M., P. Burnap und K. Jones (2018). “Early-stage malware prediction using recurrent neural networks”. In: *computers & security* 77, S. 578–594 (siehe S. 23).
- Robin Tommy, Gullapudi Sundeep, H. J. (2017). “Automatic Detection and Correction of Vulnerabilities using Machine Learning”. In: *2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)*, S. 1062–1065 (siehe S. 20).
- Sabar, N. R., X. Yi und A. Song (2018). “A Bi-objective Hyper-Heuristic Support Vector Machines for Big Data Cyber-Security”. In: *IEEE Access* 6, S. 10421–10431 (siehe S. 21).

- Shijo, P. und A. Salim (2015). “Integrated Static and Dynamic Analysis for Malware Detection”. In: *Procedia Computer Science* 46.Icict 2014, S. 804–811 (siehe S. 16, 17).
- Sikorski, M. (2012). *Praise for Practical Malware Analysis*, S. 802. URL: <https://lira.epac.to/DOCS-TECH/Hacking/Practical%20Malware%20Analysis.pdf> (siehe S. 14).
- Singh, M., D. Bansal und S. Sofat (2016). “A Novel Technique to Characterize Social Network Users: Comparative Study”. In: *Proceedings of the 6th International Conference on Communication and Network Security*. ACM, S. 75–79.
- Siracusano, M., S. Shiaeles und B. Ghita (2018). “Detection of LDDoS Attacks Based on TCP Connection Parameters”. In: *2018 Global Information Infrastructure and Networking Symposium (GIIS)*. IEEE, S. 1–6 (siehe S. 22).
- Teoh, T. T. u. a. (2018). “Anomaly detection in cyber security attacks on networks using MLP deep learning”. In: *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*. IEEE, S. 1–5 (siehe S. 22).
- Vinayakumar, R. u. a. (2019). “Robust Intelligent Malware Detection Using Deep Learning”. In: *IEEE Access* 7, S. 46717–46738 (siehe S. 22, 23).
- Webster, J. und R. T. Watson (2002). “Analyzing the Past to Prepare for the Future: Writing a Literature Review.” In: *MIS Quarterly* 26.2, S. xiii–xxiii. arXiv: 02767783 (siehe S. 6, 7).
- Yeo, M. u. a. (2018). “Flow-based malware detection using convolutional neural network”. In: *2018 International Conference on Information Networking (ICOIN)*. Bd. 2018-Janua. IEEE, S. 910–913 (siehe S. 21).
- Yin, C. u. a. (2017). “A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks”. In: *IEEE Access* 5, S. 21954–21961 (siehe S. 18, 19).

## Internetquellen

- Carmen Torrano Giménez, Alejandro Pérez Villegas, G. Á. M. (2010). *CSIC 2010 HTTP dataset*. URL: <http://www.isi.csic.es/dataset/> (siehe S. 18).
- SmartVisionEurope (2015). *CRISP-DM Methodology*. URL: <http://crisp-dm.eu/home/crisp-dm-methodology/> (besucht am 18.10.2019) (siehe S. 8, 9).
- AV-TEST (2019). *Malware Statistics & Trends Report*. URL: <https://www.av-test.org/en/statistics/malware/> (besucht am 08.10.2019) (siehe S. 3).

# Abbildungsverzeichnis

2.1. CRISP-DM Phasen (SmartVisionEurope 2015) . . . . .	8
0.1. Aufbau einer PE-Datei (eigene Darstellung) . . . . .	12

# Tabellenverzeichnis

3.1. Auszug an, durch statische Analysen identifizierte Indikatoren für einen Malware Angriff (In Anlehnung an Sikorski (2012)) . . . . .	14
--	----



## B. Abkürzungsverzeichnis

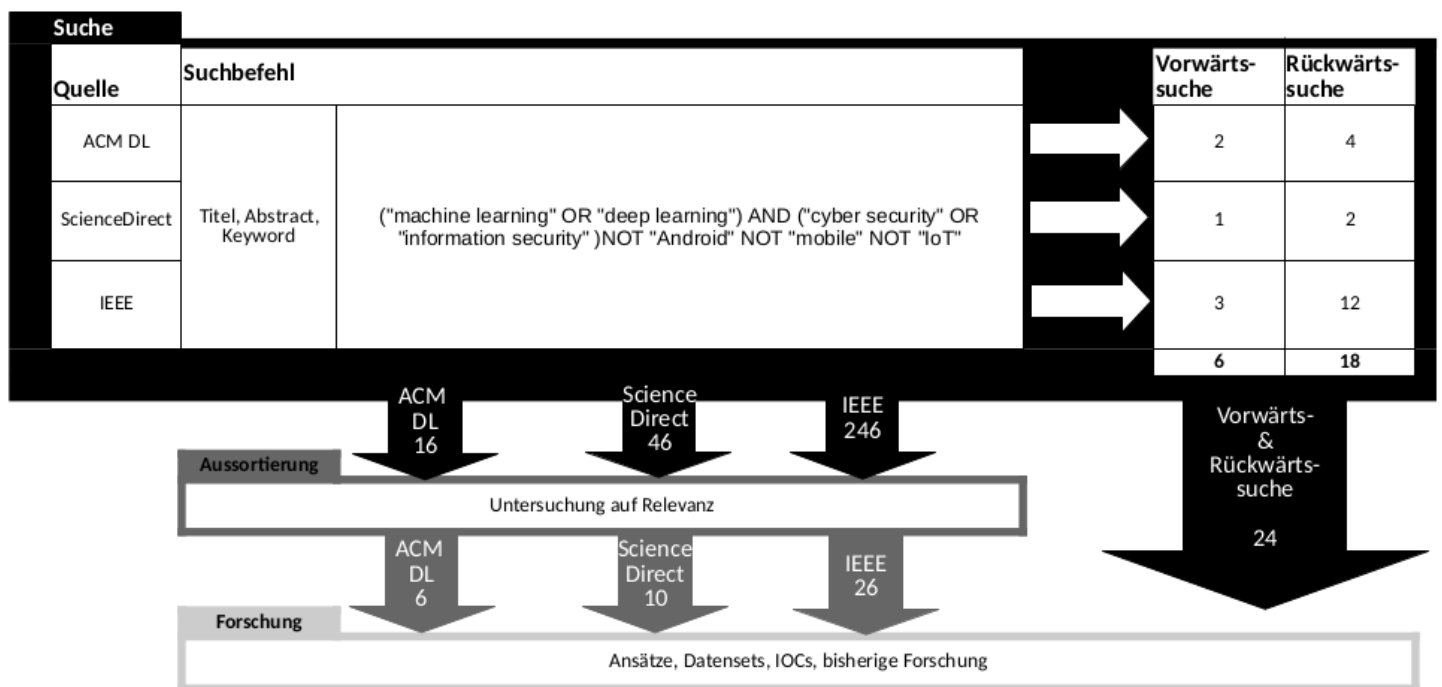
<b>AB</b> AdaBoost .....	18
<b>ANN</b> Artificial Neural Network .....	19
<b>APT<sub>s</sub></b> Advanced Persistent Threats .....	4
<b>ARFF</b> Attribute-Relation File Format .....	16
<b>BITKOM</b> Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V. ....	3
<b>BKA</b> Bundeskriminalamt .....	2
<b>BN</b> Bayesian Network .....	18
<b>C&amp;C</b> Command & Control .....	v
<b>CRISP-DM</b> Cross-industry standard process for data mining .....	6
<b>C2</b> Command and Control .....	13
<b>DoS</b> Denial of Service .....	19
<b>DDoS</b> Distributed Denial of Service .....	v
<b>DGA</b> Domain Generation Algorithm .....	v

<b>DLLs</b>	Dynamic Link Libraries .....	11
<b>DT</b>	Decision Tree .....	16
<b>FF</b>	Fast-Flux .....	v
<b>FPR</b>	False Positive Rate .....	16
<b>GUI</b>	Graphical User Interface .....	12
<b>GPU</b>	Graphics processing unit .....	19
<b>HTTP</b>	Hypertext Transfer Protocol .....	3
<b>IDS</b>	Intrusion Detection System .....	18
<b>IoC</b>	Indicator of Compromise .....	iii
<b>IoCs</b>	Indicators of Compromise .....	7
<b>k-NN</b>	k-nearest-neighbor .....	16
<b>LB</b>	LogitBoost .....	18
<b>LDDoS</b>	Low-rate DDoS .....	v
<b>LR</b>	Logistic Regression .....	18
<b>MLAs</b>	Machine Learning Algorithmen .....	3
<b>MLP</b>	Multi-Layer Perceptron .....	19

<b>NB</b>	Naïve Bayes.....	18
<b>NBTree</b>	Naïve Bayesian Tree.....	19
<b>Opcode</b>	Operation Code.....	16
<b>PE-Datei</b>	Portable Executable Datei	
<b>PE-Dateien</b>	Portable Executable Dateien.....	11
<b>RAM</b>	Random-Access Memory.....	13
<b>RNN</b>	Recurrent neural network.....	19
<b>RNNs</b>	Recurrent neural networks.....	18
<b>RF</b>	Random Forest.....	16
<b>R2L</b>	Remote to Local.....	19
<b>SMO</b>	Sequential Minimal Optimization.....	18
<b>SVM</b>	Support Vector Machine.....	16
<b>TF-IDF</b>	Term Frequency - Inverse Document Frequency.....	17
<b>U2R</b>	User to Root.....	19
<b>VBA</b>	Visual Basic for Applications.....	17
<b>xss</b>	Cross-site scripting.....	18

## C. Anlagen

### Anlage 1: Research Model



Prozess der Literaturrecherche

## Anlage 2: Literaturreview

Legende:						
Related Work	Ansatz+Datensatz	Ansatz ohne Datensatz	Datensatz	IOCs	Motivation	
Autor	Titel	Jahr	Zitate Google Scholar	Publisher	Inhalt	Rating
Raff, Edward and Sylvester, Jared and Nicholas, Charles	Learning the PE Header, Malware Detection with Minimal Domain Knowledge	2017	19	ACM	Malware Detection mit minimalem Domänenwissen wobei ein Teil des PE headers extrahiert wird. Neuronale Netze lernen aus unformatierten Bytes ohne explizite Feature Extrahierung.	3
Sewak, Mohit and Sahay, Sanjay K. and Rathore, Hemant	An investigation of a deep learning based malware detection system	2018	5	ACM	Verbesserungen der Vorhersagen des Malicia Datensets durch Neuronale Netze. Allerdings wurde das Malicia Dataset eingestellt.	2
Hendler, Danny and Kels, Shay and Rubin, Amir	Detecting Malicious PowerShell Commands using Deep Neural Networks	2018	13	ACM	Erkennen bösartiger PowerShell Kommandos mit Hilfe Neuronaler Netze und NLPs	2
Ding, Yalei and Zhai, Yuqing	Intrusion Detection System for NSL-KDD Dataset Using Convolutional Neural Networks	2018	0	ACM	Neuronale Netze wurden auf das NSL-KDD Dataset angewandt welches aus rohen tcpdump Daten besteht	3
Brown, Andy and Tuor, Aaron and Hutchinson, Brian and Nichols, Nicole	Recurrent Neural Network Attention Mechanisms for Interpretable System Log Anomaly Detection	2018	13	ACM	Anomalieerkennung in Systemprotokolle durch RNN (recurrent neural networks)	3
Chen, Xunxun and Li, Gaohao and Zhang, Yongzheng and Wu, Xiao and Tian, Changbo	A Deep Learning Based Fast-Flux and CDN Domain Names Recognition Method	2019	0	ACM	Differenzierung von Fast-Flux domain names und CDN (Content Distribution Network) domain names mit Hilfe von deep Learning	2
Lu Xiaofeng, Zhou Xiao, Jiang Fangshuo, Yi Shengwei, Sha Jing	ASSCA: API based Sequence and Statistics features Combined malware detection Architecture	2018	3	ScienceDirect	Malware Detection von system Files in Windows Systemen durch Machine Learning und Deep Learning. Daten von VirusShare und VirusTotal	3
Quan Le, Oisín Boydell, Brian Mac Namee, Mark Scanlon	Deep learning at the shallow end: Malware classification for non- domain experts	2018	13	ScienceDirect	Daten der Microsoft Malware Classification Challenge von Kaggle werden mit CNN bewertet.	3

Ausschnitt der Literaturreview Liste