NOVA IMS
Information Management School

MGI

Mestrado em Gestão de Informação
Master Program in Information Management

A toolbox for Artificial Intelligence Algorithms in Cyber Attacks Prevention and Detection

A dive into the current developments

Melissa Anita Sandholtet

Dissertation presented as partial requirement for obtaining the master's degree in Information Management, specialization in Information Systems and Technologies Management.

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

A TOOLBOX FOR THE PREVENTION AND DETECTION OF CYBER ATTACKS

A DIVE INTO THE CURRENT DEVELOPMENTS

by

Melissa Anita Sandholtet

Dissertation presented as partial requirement for obtaining the master's degree in Information Management/ Master's degree in Statistics and Information Management, with a specialization in Information Systems and Technologies Management

Advisor: Professor Doutor Vitor Manuel Pereira Duarte Dos Santos

ABSTRACT

This Thesis provides a qualitative view on the usage of AI technology in cybersecurity strategy of businesses. It explores the field of AI technology today, and how it is a good technology to implement into Cyber Security. The Internet and Informational technology have transformed the world of today. There is no doubt that it has created huge opportunities for global economy and humanity. The fact that Businesses of today is thoroughly dependent on the Internet and Information Systems has also exposed new vulnerabilities in terms of cybercrimes performed by a diversity of hackers, criminals, terrorists, the state and the non-state actors. All Public, private companies and government agencies are vulnerable for cybercrimes, none is left fully protected. In the recent years AI and machine learning technology have become essential to information security, since these technologies can analyze swiftly millions of datasets and tracking down a wide range of cyber threats. Alongside With the increasingly growth of automation in businesses, is it realistic that cybersecurity can be removed from human interaction into fully independent AI Applications to cover the businesses Information System Architecture of businesses in the future? This is a very interesting field those resources really need to deep into to be able to fully take advantage of the fully potential of AI technology in the usage in the field of cybersecurity. This thesis will explore the usage of AI algorithms in the prevention and detection of cyberattack in businesses and how to optimize its use. This knowledge will be used to implement a framework and a corresponding hybrid toolbox application that its purpose is be to be useful in every business in terms of strengthening the cybersecurity environment.

KEYWORDS

Cyber Security; Artificial Intelligence; Cyberattacks; Machine Learning

INDEX

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AI	Artificial Intelligence
ALDDoS	Application Layer Distributed Denial of Service
BPNN	Back Propaganda Neural Network
CART	Classification and Regression Tree
CNN	Conventional Neural Network
CFA	Cuttlefish Algorithm
DDoS	Distributed Denial of Service
DoS	Denial of Service
DL	Deep Learning
DNN	Deep Neural Network
DT	Decision Tree
EM-GMM	Expectation Maximization Gaussian Mixture Model
FFSL-EL	Fuzziness-based supervised learning approach through Ensemble Learning
FGLCC	Feature Grouping Linear Correlation-
GBDT	Gradient Boosting Decision Three Algorithm
IDS	Intrusion Detection System
IOT	Internet of Things
ML	Machine Learning
NN	Neural Network
PReLu	A Parametric Rectified Linear Unit
PSO	Transient Search Optimalization
DE	Differential Evolution
ReLU	Rectified Linear Unit
RF	Random Forest
RNN	Recurrent Neural Network

R2L	Remote to Local
KNN	K-Nearest Neighbor
SAE	Stacked Auto-Encoder
SCAE	Stacked Contractive Auto-Encoder
SOFM	Self Organizing Feature Map
SVM	Support Vector Machine
TTP	Multilayer Preprocessing Mechanism
TSO	Transient Search Optimalization
U2R	User to Root

1. INTRODUCTION

1.1. Motivation and Problem Identification

Studies indicate that the cybersecurity in businesses will be replaced by AI technology by 2030 (Devansen, 2020). That made me wonder the state of use of AI in cybersecurity today. How trustworthy the systems are regarding accuracy metrics, and how AI needs to improve to better meet the needs for cybersecurity. The research of AI today is currently researching more advanced and better techniques for AI, since the number of problems in current AI are still significant. The usage of AI in cybersecurity is still a new practice and that makes this study quite interesting in terms of how the future usage of AI applications in cybersecurity will work and how the AI will protect Data Systems and provide businesses from cyberthreats. The truth of today, is that AI can have many problems when trying to protect computers and data. So far, the research in this field focuses on supervised learning of AI how to protect a system against malware and outsider threats (Morgan, 2019).

A recent report new performed by Trend Micro found out that 41% of IT leaders believes that AI will replace their role by 2030 (Devansen, 2020). There is developed Advanced AI and Machine learning applications that Works within the detection and stopping of insider threats. These systems have proven to have the capability of prevention of advanced and sophisticated attacks that moves within a network- potential breaches caused by unknowing access to sensitive information. This can be tackled by intelligent and automated anomaly Detection Systems. AI and Machine Learning can help both analysts and security teams in their daily activities with performing rapid investigations and uncover hidden data patterns through searching on massive amount of logs and event data from applications, endpoints, and network devices. Alongside the societies growing dependents on the Internet and Information Systems, the cyber threats and attacks grows exponentially simultaneously. This leads to a massive cost of cyber security breaches within businesses. The cost of cybersecurity breaches becomes increasingly catastrophic for all businesses (Devanesan, 2020).

This concerning reality made me quite interested about this field, and how we should attack the increasing problem with security branches with the application of AI of today, and in the future. With the increasingly growth of automation in businesses, is it realistic that cybersecurity can be removed from human interaction into fully independent AI applications to fully cover the businesses Information System. Architecture of businesses in the future? This will be investigated and discussed during this thesis.

1.2. OBJECTIVES

The goal of the paper is to design a toolbox application for the use of AI algorithms in cyber-attacks prevention and detection.

To achieve this goal, the following intermediate objectives were defined:

- Investigate the current state of AI algorithms used in in cyber-attacks prevention and detection
- Investigate the main cybersecurity threats
- Propose a toolbox for the use of AI algorithms in cyberattacks prevention and detection
- Validate the toolbox application

This goal achievement will contribute to answer the following research question:

How can AI algorithms help businesses in the detection of cyber-attacks?

This research goal could be unfolded into some other research questions, as follows:

- How can the challenges for businesses regarding cybersecurity be solved by the implementation of artificial intelligence?
- Is there developed trustworthy solutions with AI algorithms in Cybersecurity today?
- What are the major cyber threats for businesses of today and can the implementation of AI in organizations contribute to protect against these threats?

1.3. STUDY RELEVANCE AND IMPORTANCE

Cyberspace provides users with an interactive platform where they can share information and conduct business among other activities. Cybersecurity provides the required preventive methods used to protect data, networks, electronic devices and servers from malicious attacks and unauthorized access. Elements of Cybersecurity involves application security, network security, data security, disaster recovery and business continuity, among others. Common types of cyber threats involve ransomewhere, phishing, malware and social engineering. To combat these different types of threats many tools are available on the market, such as anti-virus/anti malware software, firewalls, encryption methods, two- factor authentication techniques and software updates to improve security. These techniques have in common that they are not satisfactory for tracking and secure cyberspace from various cybercrimes. To be able to identify a wide range of attacks it is necessary to have applications that provide clever- real time decisions. Cyber Defense Systems needs to be adaptable and have a large memory. This is the opportunity enabled using Artificial Intelligence (Naik et al., 2021).

AI and machine learning can assist both analysts and security teams to investigate through a massive amount of log and event data from applications, endpoints, and network devices to conduct rapid investigations that enables them to uncover hidden patterns and find the root causes of the incidents (Stevens, 2020). As the civilizations dependents and reliance to the Internet and the Information Systems grows, the cyberattacks and threats grows exponentially besides it. As the Information systems architectures grow in businesses, as of today there is still doubts on how the AI technology can meet the requirements of the society's cybersecurity needs. The dependency of the internet and Information systems exposes new vulnerabilities in terms of cybercrimes. Therefore, Cybersecurity is now turning into Artificial Intelligence and Machine Learning to mitigate and protect the anomalous behaviors in cyberspace (Stevens, 2020). It seems like our trust and perception of the future for cybersecurity is the use of AI, but as of today there is still a lot of doubts on how the AI technology can meet the requirements of today´s cybersecurity needs. After I finish my resource, this toolbox application could be used by organizations to adapt better security of their systems using Artificial intelligence.

2. METHODOLOGY

The desired output of this study is to design a toolbox application to support AI algorithms in cyberattacks the prevention and detection. The design science methodology was chosen to drive this investigation. The DSR is used as a problem-solving paradigm and the output is to enchase human knowledge and through the creation of innovative artifacts (Brocke et al., 2020) The goal of a DSR project is to extend the capacities of humans and organizations by designing new and innovative artifacts constructs, models, methods and instantiations (Hevner et al. 2013). This is in line with the objective of this paper, which is to of this paper to propose a toolbox application for the use of AI algorithms in cyber-attacks prevention and detection. The design science research methodology can be split into two products- IT Artifacts and processes (Weber et al., 2012). The IT Artifact is the output for the paper.

This dissertation will follow the design science research methodologies presented by Peffers, et al., (2006) and Hevner, et al., (2004). It is important to acknowledge how the proposed methodology will be applied in this research. To do so it is important to explain how this will be applied in the research and what kinds of output that can be received.

2.1. DESIGN SCIENCE RESEARCH (DSR)

There is a variety of steps that is suggested to be used in the literature to solve identified problems. For this paper, the developments made by Hevner et al., (2004) and Peffers et al., (2007) is used. These stages will be conducted to this the investigation and will be explained briefly in the following section.

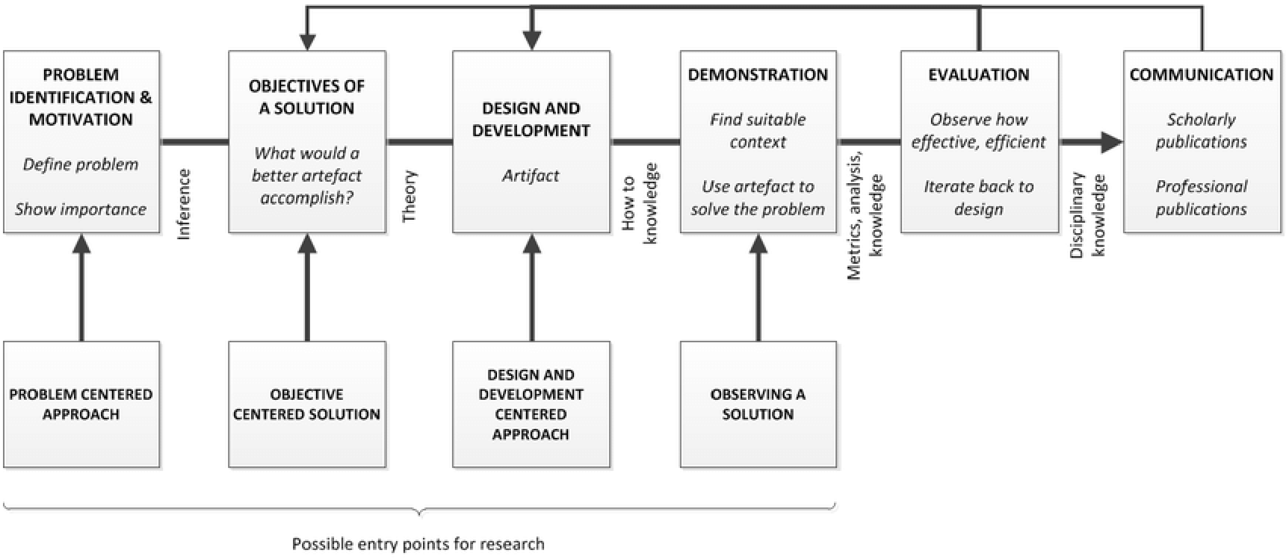

PROBLEM IDENTIFICATION & MOTIVATION
Define problem
Show importance
Inference
OBJECTIVES OF A SOLUTION
What would a better artefact accomplish?
Theory
DESIGN AND DEVELOPMENT
Artifact
How to knowledge
DEMONSTRATION
Find suitable context
Use artefact to solve the problem
Metrics, analysis, knowledge
EVALUATION
Observe how effective, efficient
Iterate back to design
Disciplinary knowledge
COMMUNICATION
Scholarly publications
Professional publications
PROBLEM CENTERED APPROACH
OBJECTIVE CENTERED SOLUTION
DESIGN AND DEVELOPMENT CENTERED APPROACH
OBSERVING A SOLUTION
Possible entry points for research

Figure 1- DSRM Process Model Adaption (Peffers et al., 2007)

2.1.1. Problem Identification and Motivation

This step is defining the specific research problem and justify the value of a solution. The problem definition will be used to create an artifact. The value of a solution is aiming to motivate the researcher and the audience to understand the problem and accept the outcome of the solution (Peffers et al., 2007).

2.1.2. Define The Objectives

When the problem is identified, the objectives of a solution (qualitative or quantitative) need to be defined of what is feasible and possible (Peffers et al., 2007).

2.1.3. Design and Development

The artifact is created in this step. This step determines the artifact´s desired functionality and its architecture. Decisions regarding Function and form is made so that requirements are set to the artifact. Resources that need to be done before moving from objectives to design and development include knowledge of theory that can be brought to bear in a solution. (Peffers et al., 2007).

2.1.4. Evaluation

This step reflects the degree of which the outcome of a solution matches the expected value. The utility, quality and efficiency must be demonstrated through well- executed evaluation methods. The essential feedback is given to the specific artifact, but also to the construction process itself. There is a variety of evaluation approaches available in the literature (Cleven et al., 2009). Attributes such as functionality, performance, reliability and accuracy can be used as comparisons on what was initially proposed and what was made. The chosen evaluation approach must be coherent with the artifact and its corresponding evaluation metrics (Hevner et al., 2004). One possible outcome of the evaluation process is that the artifact needs to be improved. Researchers may decide to look over the process again to improve the process. Researchers may also decide to leave the improvements to other projects and move on to the final step (Peffers et al., 2007).

2.1.5. Communication

In the last step, the communication is used to diffuse the resulting knowledge. The problem and its relevance are communicated to the respective audience, together with the artifact and attributes. (Peffers et al., 2007). The communication must emphasis the importance of the problem and the effectiveness of the solution realized in the artifact (Hevner et al., 2004).

2.1.6. Environment

The environment is defining the problem space which resides the area of interest. For IS research it is composed of people, organizations, and their existing or planned technologies (Hevner et al., 2004). It includes the goals, tasks, problems and opportunities that define business needs as there are perceived by people within the organization. Business needs are assessed and evaluated by inside the context of organizational strategies, structure, culture, and existing business processes. Business needs are positioned to existing technology infrastructure, applications, communication architectures and development capabilities. All of this will together define the business needs, or the "problem" as perceived by the researcher (Hevner et al., 2004).

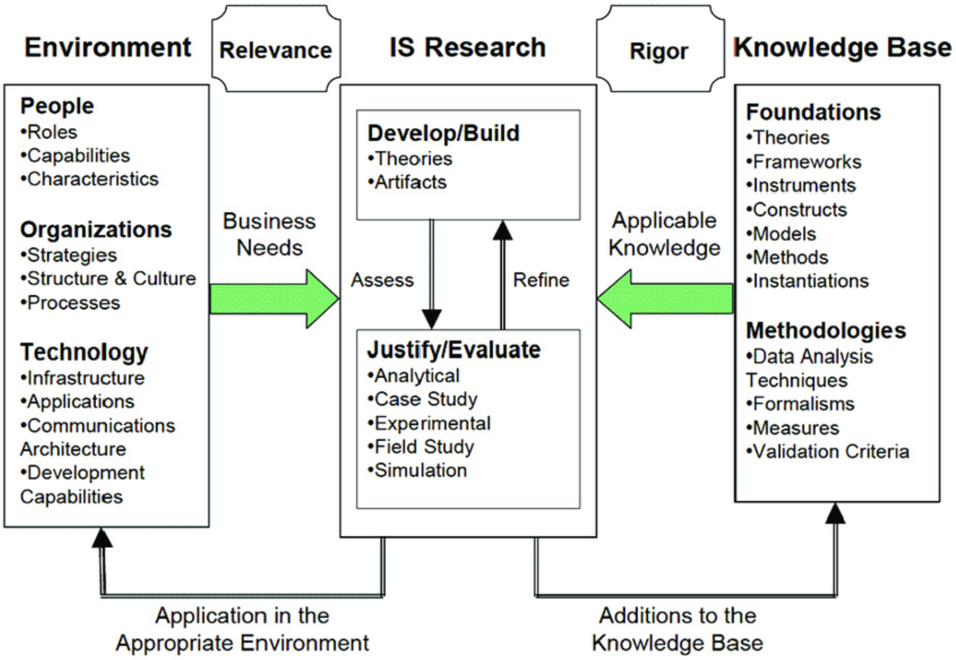


Figure 2- Information Systems Research Framework (Hevner et al., 2004)

2.2. RESEARCH INTERVIEWS

A research interview can be defined as a qualitative method to provide a deep understanding of participants experiences, opinions and point of view for a specific subject (Gill et al., 2008). Research interviews is one of the most used methods for data collection (Rowley, 2012) Interviews are discussions that is usually conducted one- to- one between an interviewer and an individual. The purpose of the interviews is to gather on the specific sets of defined topics. Interviews can be used to gather background information and/or to gather export knowledge of individuals (Harrell & Brandley, 2009).

2.2.1. Interviews Preparation

Accurate preparation by the interviewer’s side is a critical preparation and should not be underestimated. Successful interviews start with careful planning that include the area and specific scope of the questions. Conducting background reading and of literature concerning the specific research area and how to conduct research interviews is necessary before the development of the research questions. It will also facilitate the research plan. Testing devices should be tested to avoid conflicts of the interview (McGrath et al., 2018).

2.2.2. Participants Identification

In the interview process, suitable participants are important to enhance the quality of the insights. Diverse experiences, backgrounds and roles may provide more insights and interesting findings to the research (Rowley, 2012). The authors recommend to revealing the participants specific background, such as academic experience and professional role to provide more authority to the feedback. It is important to inform the participants about the scope of the research. They should also give their opinions and acceptance to be recorded (McGrath et al., 2018).

2.2.3. Interview Guide

An interview guide should be prepared in advance of the interview. When the designing an interview guide it is important to ask questions that are likely to peruse as much information as yield as much of the information about the study phenomenon as possible, and to be able to address the aims and objectives of the study. In qualitative research interviews, good questions are defined as open-ended, neutral and understandable. The length of an interview varies depending on the topic, researcher and participants (McGrath et al., 2018). When conducting the actual interview, it is important for the interviewer to be well known with the interview schedule. By that, the process will appear more natural and less rehearsed. The researcher should also possess the right repertoire of skills and techniques to comprehensive data are collected during the interview (McGrath et al., 2018).

2.2.4. Interview Conduction

Listening and questioning is two of the most fundamental skills when conducting interviews, according, alongside with a natural attitude and use of silence (McGrath et al., 2018). The interviewer should not play a passive role, but rather contribute with his/her knowledge to provide insights. There should be paid careful attention to make a positive environment to obtain a good relationship between the actors involved. That will help to avoid possible bias in the collected data (McGrath et al., 2018).

2.2.5. Interview Transcript

Writing down recorded data may be time consuming, but if it is done properly if will give great benefits to the researcher. It allows the researcher to dive into the data and validate it properly (McGrath et al., 2018). The author also recommend that the transcription occurs as soon as possible after completion of the interviews.

2.3. RESEARCH STRATEGY

- **Activity 1: Problem Justification and motivation**

Enchasing the knowledge and understanding of the capabilities regarding AI and machine learning are becoming essential to information security, since these technologies can analyze swiftly millions of datasets and tracking down a wide range of cyber threats (Morgan, 2019) Understanding the current technologies and how to optimize to enable AI algorithms to be used to prevent and detect cyberattacks is possible through the construction of a toolbox that emphasizes the current technologies and its further application on technologies deployment.

- **Activity 2: Define the objectives for a solution**

When the problem justification and motivation is defined, the objectives are prepared. This step focuses on inferring the objectives of a solution from the problem definition to what is possible and feasible AI algorithms for the implementation of the toolbox application. The objectives will support the investigation on how the toolbox application is expected to support solutions to problems not previously supported or addressed. The objectives support the defined problem justification which results from the research question: "How AI algorithms can help businesses in the prevention and detection of cyberattacks?".

- **Activity 3: Environment**

This activity focus on the business need of a toolbox that is composed with AI algorithms that can be used in the prevention and detection of cyberattacks. The business, economic and technical aspects of AI in cybersecurity is investigated and discussed in this section.

Activity 4: Design and Development

This is the activity where the artifact is applied. The toolbox application to support AI algorithms in the prevention and detection of cyberattacks, is developed through the investigation of current technologies, applied knowledge of tested datasets with algorithms proven to work in actions that is tested in open-source datasets. To be able to build the toolbox application the respective literature is build according to the defined objectives.

- **Activity 6: Evaluation**

The observation on an artifact can be evaluated in many different forms, depending on the nature of the problem venue and the artifact. The activity involves comparing the objective of a solution to the actual results from the use of artifact in an observation. The artifact, which is the toolbox application will be evaluated in deep interview with relevant people in the field.

- **Activity 7: Communication**

A scientific publication is a way to communicate the proposed artifact to the public. In that way, the details of the development can be shared, and it will also open for other improvements suggestions.

3. ENVIRONMENT

3.1. CYBERSECURITY- IMPACT ON ENTERPRISE

Before investigating the appropriate scientifical papers to build my toolbox, it is important to study the environment in which the toolbox could be effective. The environment to be studied in this paper is the businesses that is dealing with information technology in their daily bases and is having their core operations digitalized. There is a demanding need for intuitive and automated systems- level approach control the overall security characteristics of Information Systems (Iguer et al., 2014). Every organization is at risk of data breaches, system hacks, malware or ransomewhere attacks. A survey from Statista that collected number of US data breaches between 2013-2019 showed that in the half of 2018, 308 out of 688 reported data breaches impacted business. (Statista, 2020). Banking, credit and financial institutions was at top three with 84 reported breaches. Personal information and credit are the prime targets for cyberattacks. A survey made by Accenture found in 2016 revealed that 78% of financial institutions were confident in their cybersecurity strategy, yet 1 of every 3 institutions is successfully attacked- with an average of 85 breaches attacks per year. The E-commerce and retail industry is also considered at high risk. These companies are threatened through omnichannel access and supply chain networking, that is holding a large collection of personal and financial data (Manship, N.A). Cyberattacks have increased a lot in frequency the past year. It has been increased in the sophistication of assaults, and in the intensity of the damage that occurs to the organization’s data and systems. A new report issued by Trend Micro found that 41% of IT leaders believe AI will replace their role by 2030. Until the recent years, automation has not been considered as a “fix-all” solution in cybersecurity, the human interaction with advances knowledge and expertise have been considered the most important aspect of Information security (Davanesan, 2021). The threat landscape of cyberattacks is rapidly changing alongside with the digitalization processes of businesses. The increasing digitalization and particularly the use of internet has exposed organizations to a variety of new risks resulting from attacks though digital interfaces. These include denial- of- services (DoS) attacks on networks, data breaches on corporate devices, and data viruses that can sabotage computer infrastructures. The sabotage of systems in order to compromise services and system integrity and theft of customer data is all kids of acts that result in harm to an enterprise which is highly dependent on the digital technologies to conduct their business (Agrafiotis, 2018). As companies increase their reliability on Information Technology and Information Systems to collect, process, store and transmit data, the knowledge and expertise of cybersecurity becomes more important than ever for individuals and companies. If companies don´t comply with the security needs, companies will experience more data branches that will be ranging from minor disruptions to devastating consequences for individuals and organizations (Callen-Naviglia & James, 2018).

3.2. CYBERSECURITY- ECONOMIC IMPACT

According to a report by Deloitte from 2020, the average company will spend between 6% and 4% of their annual IT budget on cybersecurity each year. Most companies spent approximately 10% of their IT Budget on cybersecurity. Summing up, the companies should take their IT budget and multiply it by 0,10 to get a view of what the cost of cybersecurity will be. The report reported that the average company spends 3,2% of its total revenue on IT costs. The findings hold true across industries and companies of all sizes. The report shows that a small company will generally have a budget less than $5 million. For midsize organizations the spendings is between $5-20 million. While larger organizations will normally spend $20-50 million per year. The report revealed that large businesses spent between $2 million and $5 million on cybersecurity per year. Midsize businesses spent between $500,000 and $2 million on cybersecurity per year, while small businesses spent $500,000 or less on cybersecurity per year. It has already been shown that cyberattacks tends to be much more expensive. Besides of the economic loss, the damage to a company´s reputation is uncalculated (Lewis & Smith, 2020). A report from MacAfee estimated since 2018 that the cost of global cybercrime reached over $1 trillion. McAfee also estimated that the monetary loss from cybercrime at approximately $945 billion. The global spending was added to this, which was expected to exceed $145 billion in 2020. As of 2020, that was $1 trillion dollar drag over the economy. Only small proportions of organizations have a solid plan to prevent and respond to IT Security incidents (Lewis & Smith, 2020). The demanding needs of cybersecurity has led many organizations to pay greater attention to cybersecurity investment decisions, especially to procure the appropriate level of these investments (Chronopoulos et al., 2018).

3.3. CYBERSECURITY- THREATS

The common defense mechanisms used by organizations today are not sufficient enough to protect their cyber environment from the evolving security vulnerabilities. In fact, cybersecurity has become a global interest and importance for all organizations. It spans securing information by detecting, preventing and responding to cyberattacks (Humayun, 2020). A recent report from Embroker team from 2021 rates cyberattacks as the fifth top rated risk. The risk continues to grow. The risk will not decrease for the following years. Cybercrime is up 600% as a result of the covid-19 pandemic and almost every industry must embrace new solutions in their cybersecurity strategy. (Embroker Team, 2021). The report performed Embroker in 2020 estimated that cybercrime will cost companies worldwide $10,5 trillion annually by 2025. That is an increase of cost from $3 trillion in 2015.

(Embroker, 2020). Cyberattacks is targeting every business, and especially small to medium size businesses are becoming more frequently targeted. A research report on cost of cybercrime performed by Accenture found that 43% of cyberattacks are aimed at small businesses, while only 14% are prepared to defend themself (Accenture, 2019). The cost of cyberattacks can extend for months to years, including significant expenses that companies are not aware of and have not included in their budget or cybersecurity strategy. include in their strategies. The short and long-time cost of cyberattacks include loss of data, revenue losses from system downtime and business disruption notification cost. Cyberattacks may also damage a brand´s reputation (Embroker Team, 2021). The ever increasing cyberattacks requires the cyber security and forensics specialists to make greater attention regarding the analysis, detection and defense against the cyber threats in real-time. In practice, it is not possible to deal with the large numbers of attacks without deeply dive into the attacks features and taking the necessary actions. This is where the essence of cyber intelligence come to place. This kind of intelligence will not be possible without the use of AI, machine learning and advance data mining techniques to collect and analyze, to be able to truly understand the cyberattack evidence (Conti et al., 2018).

3.4. CYBERATTACKS- IMPACT ON ENTERPRISE

There is no doubt that there are issues related to digital security today. Cybersecurity is a concern of ever business. As of today, AI has been embedded in Information Systems across all industries (Laato et al., 2020). At the forefront of AI today are Systems that learn how to perform a specific task without being specifically programmed for that task. AI systems have the capability to learn the mappings from their inputs to their output, can learn more abstract representations of their inputs to their outputs and perform this process through multiple layers of input- to- output mapping. Each layer with increasingly abstract representations. Alongside with the availability of large datasets and the ever-increasing computing power, there has been a growing use of data-driven artificial intelligence systems, which has shown their successful application in diverse areas. The advances and performance improvements in computing power have given the capability to retrieve an enormous amount of data and information available in a wide range of platforms has become an important component in decision making processes. The continues increase in security problems has given rise to a growing demand for Artificial Intelligence (AI) approaches that have a significant impact on detecting simple security risks, as well as sophisticated cyber-attacks (Parades et al., 2021).

4. LITTERATURE REVIEW

To be able to get a broader understanding of the subject of this study, a theoretical overview is necessary. Therefore, a review of important concepts of AI and Cybersecurity- and why how AI can be used in the context of cybersecurity will be further consolidated in this section. The following section will take a closer look at: Concepts of AI with its technical aspects and capabilities in cybersecurity.

4.1. CYBERSECURITY CONTEXT

As to understand the topic that will be addressed in this study, it is important to acknowledge the concepts of the different use and potential of the use of AI in the detection and prevention process of cyberattacks. The investigation will also address the evolution of this use of AI to detect and prevent cyberattacks in the cybersecurity field. How AI can be used in the detection and prevention of cyberattacks is a discipline that is a hot topic for researchers.

The purpose of Cybersecurity is to provide protection to computer systems, networks and electronic data from information disclosure and illegitimate access (Schatz et al., 2017). According to Sarker et al., (2021) cybersecurity is dealing with security of anything around the cyberworld. That includes Information security, network security, IOT Security, cloud security, infrastructure security and operational security. Cybersecurity is characterized by its collection of methods that involved everything regarding protection of personal, governmental, industrial and other information cyber criminals. (Srivastava et al., 2021). From 2016 to 2021, The investment on worldwide cybersecurity is expected to increase more than $1 trillion (Morgan, 2019). In today's society, Internet usage is an essential daily routine. The amount of data the society exchange daily is huge and enormous. On the other hand, the number of cyberattacks is continuing to increase at a considerable rate. Every few months, cybercriminals double the potency of their customized attacks for half of the price (Akhata et., 2021). As cyberattacks evolve and become more sophisticated and automatic, traditional cybersecurity becomes inefficient (Truong et al., 2020). Traditional approaches of cybersecurity that includes network protection systems and computer security systems provides inefficient security against the continuously evolving and creative attempts of today's cyberattacks (Kabbas et al., 2020). Therefore, there is a high demand to find a way to combat with the developing cyber threats. This is where Artificial Intelligence (AI) is one of the approaches that can be used efficiently in Cybersecurity.

4.1.1. Artificial Intelligence in Cybersecurity

Over time, the conventional cybersecurity architectures, like firewalls, encryption and virus software have decreased their protection and failed as malicious attackers and/or hackers have found a way around them. The traditional technologies used in cybersecurity, focuses mainly on the past and is totally dependent on known cyber-attacks. When new attacks appear, the traditional systems are unable to detect them, because they are not able to spot changes, and detect them. This is leaving a blind spot during the appearance of unusual or "new" attack the system is not trained to spot and detect. Complemented by its learning capabilities, AI have the capability to help the certain the maintenance of cybercrime (Tyugu, 2011). AI have the capability to detect new and complex variations in attack flexibility (Tuang et al., 2020). AI technologies have been utilized in gathering data that are valuable in the prevention of cyber-attacks. (Poonia, et al., 2011). AI the ability to deal with a large amount of security data (Truong et al., 2020). AI will include self-contained security systems that can detect and response to attacks. The amount of data breaches received daily is already hard to handle for security personnel, however automatically detecting and prevention by responding to threats has helped to reduce expert's workload. AI is found to be a great method to manage these attacks (Akhtar & Fent, 2021). When a significant amount of security data generated and transferred through the network daily, the network security analyst will find it difficult to keep up the analysis with the large volume and phase of the data that is circulating through the network. The network analyst will find it difficult to monitor and consider attack elements quickly and corrects. This is where AI applications will assist by increasing the frequency of which suspicious type of behavior is mentioned and detected (Akhtar & Fent, 2021). This can assist network security officers by reacting to situations they have not seen or handled before faster, reducing the need for manual time-consuming analysis performed by security analysts (Akhtar & Fent, 2021). Regular network traffic is studied by AI security systems over time which will lead to detection of the threats over time. This will lead AI to be able to make a baseline of what are the normal patterns. If any change of inconsistency is found in the normal patterns, the AI security systems will detect the attacks (Akhtar & Fent, 2021).

4.1.2. Machine Learning Technologies

Machine Learning serve as a branch of AI and is closely related to computational statistics, which also focus on prediction-making by computers. Machine Learning is defined by Arthur Samuel as a "Field of study that gives computers the ability to learn without being explicitly programmed". Machine Learning`s primary focus is classification and regression based on features that is previously learned from training data. This is a part of supervised learning. ML can also be unsupervised and have the capability to be trained to learn baseline behavioral profiles entities and thereby used to find anomalies (Xien et al., 2018).

Machine Learning (ML) has gained a large interest in many applications and fields of study, particularly in cybersecurity. ML can be used to analyze and classify bad actors in a huge set of available data. Machine Learning can help project traffic trends and spot anomalies in network behaviors. Machine Learning approaches are becoming enablers for security solutions in computer networks due to its capabilities to traffic information in order to detect abnormal patterns. Detecting zero-day intrusion have been a goal for cybersecurity for a long time, especially in intrusion detection (Delease et al., 2019).

Delease et al., 2019 defined as “a special sort of technique in machine learning for the extraction of features, better learning, and perception of machines”. The technology behind Deep learning uses multiple consecutive layers for their algorithmic operations. These layers are connected in an interlinked fashion where each layer receives input as the output of the previous layer. This is an efficient algorithmic usage with excellent advantage for the hierarchical extraction of features that are best in representing data rather than features that are manual in deep learning aspects. It uses specified architectures of Artificial Neural Network Namely Multilayer perception, Convolutional Neural Networks and Recurrent Neural Networks (Delease et al., 2019).

In the recent years, the developments of AI-based cybersecurity architectures has increased significantly, the developments efforts toward the implementation of models that would guarantee high degree of scalability. One of such architectures is the Artificial Deep Neural Network (ANN). The utilization of ANN in cybersecurity is reinforced by their ability to create smarter learning patterns that can reflect acceptable/normal network behavior (Abdiyeva-Aliyeva et al., 2021).

Lazeer et al., 2021 defines Artificial neural networks as “Imitations of a neuron in the human brain.”. These are dedicated to solving machine learning problems, and a neuron is considered as a unit which is expressed by an activation function (Lazaar, 2021). Neural Network Can be divided into four main categories: (1) Feed Forwarded Neural Networks, (2) Recurrent Neural Networks, (3) Resonance Neural networks and (4) Self organized neural networks. The has been an increasingly focus and interest dedicated to neural networks during the last decade, and it have been deployed heavily to classify objects or to make predictions of data. Various applications of neutral networks are found in image recognition, classifications of text or images, identification of objects, data prediction and filtering datasets. (Lazaar, 2021). The field of cybersecurity has been given a great boost with the field of deep neural networks, since it has proven to serve as a good reinforcement against cyberattacks. One of the reasons for that is due to the scalability of such framework, that can helps human cyber security experts to cover a wide range of perspectives as such systems swing into action to independently detect and prevent cyberattacks. (Abdiyeva-Aliyeva et al., 2021).

Deep Neural networks, which is a form of Artificial Neural Networks (ANN) have in the recent years been widely employed in the area of cybersecurity. The utilization of ANNs in cybersecurity is supported by their ability to create learning patterns can have the capability to reflect acceptable/normal network behavior. Besides Its abilities It is critical to train the DNN model with the cybersecurity data so that the model is examining in cyberattack patterns- that will determine the efficiency of the model. When the DNN has a comprehensive architecture, the architecture is developed for the detection and prevention of Malware and DDoS attacks (Buczak, et al., 2015).

4.1.3. Intrusion Detection Systems

Intrusion detection is the process of monitoring systems and networks for abnormal activities or intrusion attempts. The aim of the system is to collect events and report them to the administrator. Intrusion detection systems are mechanisms that enchase the security when implemented in parallel with anti-malware, firewalls and access controls (Verwoerd & Hunt, 2002). Intrusion Detection systems can either be software or hardware. The purpose of an IDS it to automate the process of monitoring events to reveal intrusions. Intrusion Detection Systems is categorized into different groups depending on the type of event it monitors, or how it is deployed. The different types is: Network-based IDS, Host-based IDS, Hybrid-based IDS, Signature-based IDS, Anomaly-based IDS, Passive-based IDS, and Active-based IDS. (Lazaar, 2021). The Challenges that are brought by security breaches for businesses can be devastating and even destroy a business`s reputation.

Therefore, there is a great importance to have a good Intrusion Detection System in place. The Intrusion Detection should be able to detect and protect from such security breaches. Intrusion Detection Systems are feasible solutions for cybersecurity problems, but they can face challenges without the right implementation. Unfortunately, Anomaly based IDS can have a high rate of false positives (FP) and require considerable computational requirements (Kaja et al., 2019).

4.1.4. Network Intrusion Systems

In this paper, I study is concentrated around network traffic data and therefore the recent developments of Network Intrusion Detection Systems are studied further. Niyaz et al., (2015) defines Network Intrusion Detection Systems (NIDs) as "*The essential tools for the network system administrators to detect various security breaches inside an organization ".* NIDS monitors and analyzed network traffic that enters or exists from network devices of an organization. The NIDS raised alarms if an intrusion is observed. Intrusion Detection can have various methods and NIDSs are categorized into two main classes: Signature Based NIDS (SNIDS) and Anomaly detection-based NIDS (ANIDS). In the SNIDS, Snort attack signatures are pre-installed to preform pattern matching for the traffic against installed signatures to detect an intrusion in the network. On the other hand, the ADNIDS observes a deviation from normal traffic patterns by classifying network traffic as an intrusion when it observes a deviation (Niyaz et al., (2015). When SNIDS is known for good results and efficiency in the detection of known attacks and shows high detection accuracy with less false-alarm rates. On the other hand, ADNIDS, is known to be suitable for the detection of unknown and new attacks. ANIDS is known to produce high false-positive rates, but it has wide acceptance among the research community for its theoretical potential in the identification of novel attacks (Niyaz et al., 2015).

4.1.5. Cyberattacks

In order to understand the different developments and technologies associated with AI in the prevention and detection process of cybersecurity it is important to understand the variety of different cyberattacks and cybercrimes that currently exists and is a threat to businesses of today. There exists a variety of cyberattacks that currently exists. For the purpose of this paper, -- main categories that is found to be the most prominent on businesses will be described.

Denial of Services (DoS)

For more than one decade, Denial of Service attacks has caused severe property loss and private leakage in banking, transportation and other crucial services. A Denial of Services (DoS) Attacks is an attack which the perpetrator deliberately overloads an individual's or organization´s web servers by repeatedly accessing them. The perpetrators typically use an automated program (Joy, 2020). It includes spamming, a computer resource, such as web server or computer memory) until the computer overloads and crashes. This is causing it to slow down so much that the user cannot use the machine. (Chan et al., 2019). Such attacks typically aim to render a website or computer network inaccessible for a period. This type of attack is often used in the contacts of corporate sabotage, computer-based activism and or hacktivism and cyberwarfare. In many cases, perpetrators of such attacks use a computer program that infect random unsecured computers and run undetected in the background. (Joy, 2020).

Remote to Local (R2L)

This attack type is where the perpetrator sends specific packets to a machine over a network to gain the local access of the network. In these cases, the attacker typically does not have an account on the physical machine but is still able to send the packets to the system over a network to gain the local access of the network (Chan et al., 2019).

User to Root (U2R) attack

In this attack, the perpetrator attempts to get access right from the host and can gain root access to the machine. To be able to do this the attacker must exploit the system through sniffing passwords or do stereotypical hacking (Chan et al., 2019).

Malware

Malware is defined by combining the two words malicious and software. The term Malware is used to indicate any unwanted software. Idika et al., (2007) defined malware as "any code added, changed or removed from a software system in order to intentionally cause harm or subvert the intended function of the system". Malware is categorized by its ability of self- execution, corruption, propagation and replication of a computer system. The Corruption of a computer System can have big consequences of its information confidentiality, integrity and denial of services (Namanya et al, 2018).

Botnets

Bots is a part of malware, and are programs designed to performs specific operations. Bots can be designed for legitimate purposes. A bot is malicious when it's designed to form Botnets. Botnets is defined as "*A network of host computers (zombies, bot) that is controlled by an attacker or botmaster*" (Namanya et al., 2018). The purpose of designing botnets is to run malicious software under the command of a botmaster (Bertino et al., 2017).

Ransomewhere

In the recent year, the amount of ransomewhere attacks has increased explosively and is used by criminals to generate revenue through extortion. Ransomewhere is a part of social engineering- that serves as key part of criminal´s attacks strategy. Kane et al., 2018 defines ransomewhere as a software that renders a victim´s computer or data unusability. Many internet users seem to be unaware of ransomewhere attacks, and therefore do little to protect themselves. Ransomewhere is expected to continue to evolve beyond the capability of the defense solutions of present day (Kane et al., 2018).

Unknown and Zero Day attacks

Zero-Day attacks is a cyberattack that exploits a vulnerability that has not yet been disclosed publicly. A zero-day attacks serve as a serious threat to the Internet security since it enable zero-day vulnerabilities in computer systems. In Zero-Day Attacks, Attackers take advantage of the unknown domain of zero-day exploits and thereby use them in conjunction with sophisticated and targeted attacks. The goal is to achieve stealthiness with respect to the standard techniques of Intrusion Detection. (Bilje et al., 2012). Present research illustrate a variety of issues since it is not able to provide a complete solution for the detection and analysis of zero-day attacks (Kaur & Singh,2015). The increase in the number of such attacks and its new varieties of poses a tremendous challenge for IDS solutions that rely on a database of historical attack signatures. Therefore, the industrial demand for robust and trustworthy IDSs that are capable of flagging zero-day attacks is growing. The Current research on Zero-day detection available, that is often outlier-based suffers from high false-negative rates. This reduces their practical use and performance (Hindy et al., 2020).

4.2. Prisma Methodology

The previous section of concepts defined the scope of this investigation and specified the specific keywords for the use of this scientific review. A systematic Literature review was performed to understand the most recent development in the use of AI in cybersecurity´s literature, regarding technology and successful use. The following research questions were developed to conduct this process:

RQ1	“What are developed AI systems and algorithms developed in the prevention and detection of Cyberattacks”?
RQ2	What are the algorithms accuracy and detection rate and what type of attacks can they detect?
RQ3	Which AI methods and algorithms could be recommended to be used in cyber-attacks prevention and detection?

Table 1- Systematic reviews research questions.

Answering these questions required a selection of the most relevant studies in this area.

Therefore, keyword was carefully selected (table 2). The aim of the keywords was to consult the available literature in this area and proceed to a responsible selection of articles and scientific studies available. The terms “AI”,” cyberattacks”, “Intelligent detection system”, “Algorithm’s “Detection”, “prevention” was treated as keywords.

Keyword	Title
AI	Algorithm
AI	Machine Learning
Intrusion Detection System	Prevention, Detection

Table 2- Systematic Review´s Keywords and Synonyms.

Boolean Queries were designed to include at least one of the above expressions in the abstracts, titles or keywords of the search articles. These queries guaranteed the collection of relevant data related to the domain of this study. Only scientific documents were considered before 01.01.2017 and between 2017 and 2022, aiming to cover the latest AI developments in the cybersecurity field.

The query string was prepared as follows:

(”AI” or Algorithm”) AND (” Detection or Prevention”) AND (” Cyberattacks”) NOT (Review or overview”) AND (” Machine learning” or neural network”) NOT (” Review OR Overview”)

The databases included in the studies are reflected in table 3.

Databases	Domain
Nova Discovery	https://www.novadiscovery.com/
Web of Sciences	https://webofscience.com/
Scorpus	https://scorpus.com/

Table 3- Systematic Reviews- databases and domains

For the aim of selecting the most relevant articles for this study, inclusion and exclusion criteria were defined as mentioned in Table 4:

Inclusion Criteria	Exclusion Criteria
Algorithms for Cyberattacks prevention and detection	Publications before 2017
Only finished Articles.	Language different than English
Deep Learning, Neural Network and Machine Learning technologies for Cybersecurity	Articles that review/ Overview
Hardware security, Network security, Malware Security, database security and cloud security	Non-Academic papers (e.g., Magazines reports, newspapers).

Table 4- Systematic review- Inclusion and Exclusion Criteria.

After inserting the search strings in the chosen databases, one got all the identified articles through the database search which resulted in a total of (n=1084). This is the identification phase of the Prisma Flowchart. Then, moving to the screening phase, the first step is to remove duplicates. One removed a record of (n=330), the remained articles (n-884) were analyzed. In step 2, a total of 208 articles were removed due to their topic within the AI technologies and algorithms out of scope of this study. After this removal (n=676) were analyzed. In step 3: AI developments within very specific domains were removed (e.g., Vehicle's systems, water distribution systems, healthcare systems) and (n=376) articles remained. The last abstract screening was based on articles that did not include developments and/or were based on other authors developments (n-200). After the screening process, a total of (n-23) articles remained and carefully analyzed. The method retrieved 23 articles to be included in this study. The studies are referred to as D1- D23 in following sections of the paper.

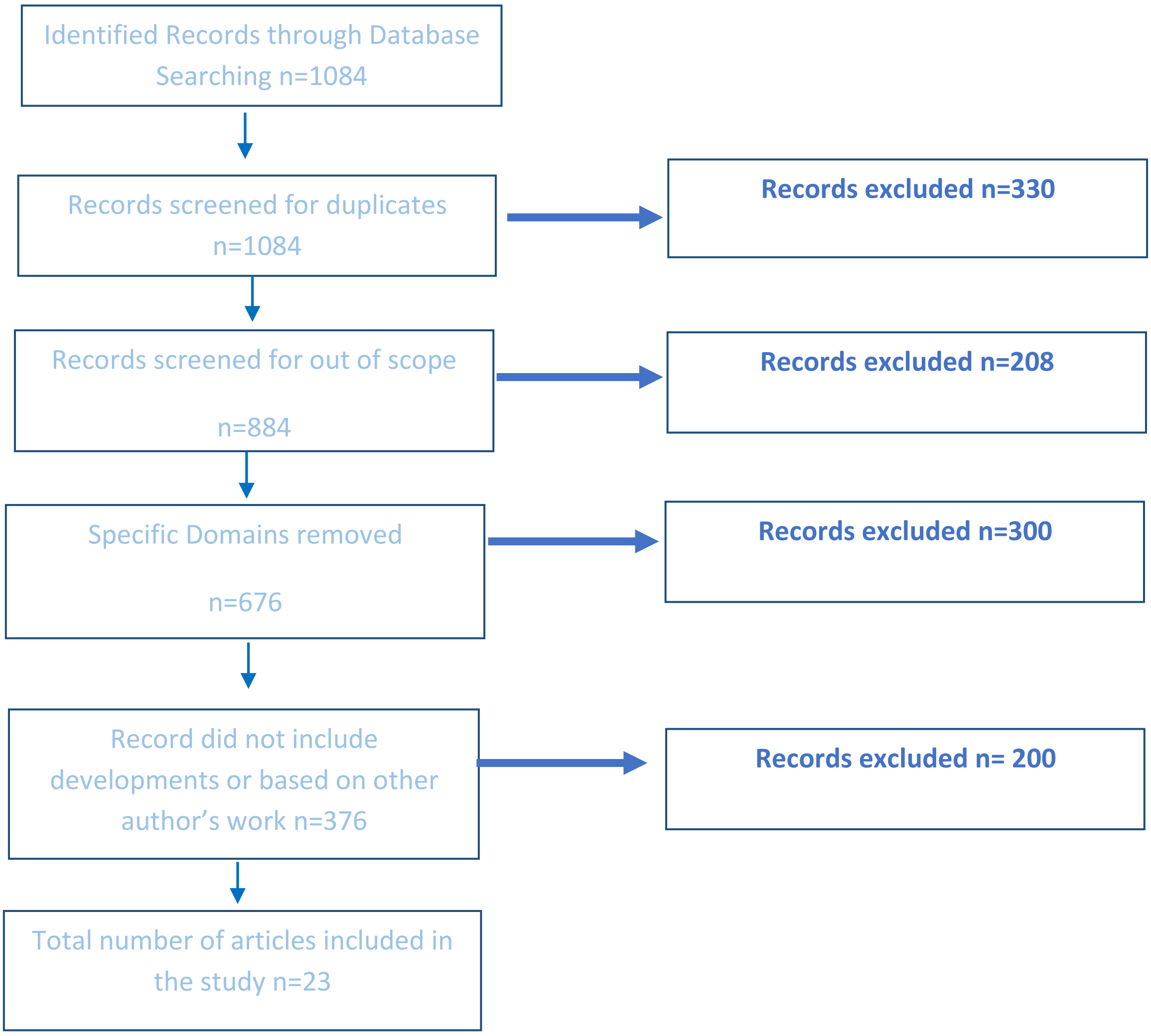

Identified Records through Database Searching n=1084
Records screened for duplicates n=1084
Records excluded n=330
Records screened for out of scope n=884
Records excluded n=208
Specific Domains removed n=676
Records excluded n=300
Record did not include developments or based on other author's work n=376
Records excluded n= 200
Total number of articles included in the study n=23

Figure 3- Prisma Flowchart

4.2.1. Analysis per Topic

There is a variety of cybersecurity models with the incorporation of AI technologies that have been proposed by different researchers. For the scope of my investigation, I focus on the developments within Neural Networks, Machine Learning and deep learning techniques as this is the most used techniques in the field. In this theoretical section I will explain the different concepts and explain my research findings regarding the techniques, algorithms and models developed for the aim of preventing and detection cyber-attacks according with developed research questions. This Information will further be used to develop a toolbox that contain algorithms and frameworks that can have the capabilities to be a help with showing previous developments and their corresponding results for intrusion detection systems combined with AI algorithms in prevention and detection of cyberattacks for businesses.

The studies of D1-D14 focused on the detection of attacks using Machine Learning, Deep Learning and Neural networks techniques and respective algorithms to detect a variety of different cyber-attacks. The result of the detection rate varies. The lowest detection rate is 84% with S10, compared with the highest achieved accuracy rate with the use of Random Forest algorithm and Support Vector Machine both for the distinguishing of malicious or normal network traffic.

D1 and D3 uses deep learning algorithms to classify Malware and DDos attacks with an accuracy rate within 97, 7 to 99,79% within their selected database. D2 and D4 uses neural network and artificial neural network to classify the distinguishing between malicious and normal network traffic. While D2 uses three-layer-propagation Neutral network (BPNN) for the classification of abnormal activity, D4 uses Deep Neural Networks (DNN) with multiple fully connected layers (Multi-Layer Perception). D2 achieved an accuracy rate of 99, 23% for the detection of application layer DDos attack. With the usage of the more complex DNN, S4 were able to achieve a higher accuracy rate of 99,97% for DDos attacks, and 99,44% for Malware attacks.

D5 and D6 put their focus around the feature selection process for their proposed IDS. Both studies use feature selection algorithms in the purpose of higher the accuracy rate of their proposed IDS. D5 uses the Fisher Score Feature Selection Algorithm with the combination of supervised learning algorithms, whiles D6 uses a new filter-based feature grouping method based on linear correlation coefficient (FGLCC). The best accuracy is in D5 where the feature selection with the combination with decision three algorithms gains the highest accuracy. The IDS in D13 do also highlight the feature selection process. In this IDS the authors propose a feature optimalization technique with the use of Machine Learning and Deep learning for malicious PowerShell script detection and gains an accuracy of 98%. While most of the studies put their efforts against anomaly network detection, S11 developed an IDS named “PhishHaven” with the focus on Efficient Real-Time AI phishing URLs Detection. The system was able to gain an accuracy of 98% with the use of a variety of machine learning techniques.

D15-19 combines the detection and prevention of cyber-attacks by adapting hybrid models for both purposes. The variety of techniques used is the Hybrid Machine Learning, Conventional machine Learning, Multi-layer perception and recurrent neural network. The accuracy rate of the five studies varies from 82% to the best result -96%. Whereas the best accuracy rate is found in the proposed IDPS that were built with the use of the machine learning algorithm- Random Forest in S17, where the authors proposed the system “JARVIS”- an Intrusion Detection and prevention System for DDoS attacks. All the systems are proposed to target DDoS attacks, except S17 who proposed Recurrent

Neural Network (RNN) to a behavioral- based model for early-stage network prediction with the accuracy rate of 96%.

D20-24 is developments of IDS for cloud and database management systems. Their approaches have an accuracy rate that vary from 64,96 to the best result – 99,98%. In this section the accuracy rate varied a lot. The reason will also lay within the complexity within cloud and DBMS. S20, 22, 23,24 is all proposed IDS in the Cloud Environment. The highest accuracy for the IDS is found in S22, whereas the authors used a feature selection method and intrusion detection using a combination of machine learning techniques such as neural network, fuzzy logic, Support Vector Machine and accomplice the highest accuracy rate with the neural network algorithm with 99,98% accuracy.

5. PROPOSAL

This field is constantly evolving, because of the risk every business is facing. The risk of Cyberattack is with businesses all the time and the damage can be critical. Therefore, the research area is big, and there are many available developments out there. Since this is cyber security is an area in high demand, organizations would be to evolve their competences regarding this topic. During the study I found that the biggest challenge regarding this topic is the competences out there. Defining Cyber-attacks, developing suitable algorithms, define the optimal number of features for the model and have the right data to test it against is a challenge and a process that demands high competences, time and cost. There is also a great idea to invest in in house or hire competence in the development of algorithms and tools for prevention and detection of cyber-attacks. Today's cyberenvironment is evolving fast, and new cyber-attacks develops every day, therefore the use of AI and Machine Learning is a great tool to use to protect organizations cyber environment. To propose an algorithm to use in prevention and detection in case of cyberattack, the technical expertise inside the organization is crucial. No algorithm or tool is perfect, and it always must be monitored and continuously developed. Although during my thesis investigation there is developed a variety of techniques and algorithms with a good accuracy rate for the prevention or protection of different type of attacks, and to be used in different kinds of systems.

5.1. ASSUMPTIONS

After my investigation of the main areas and concepts of the terms of AI techniques in Cybersecurity prevention and detection of cyber-attacks, as well as specific developments in this field, and how many different techniques, methods and results that is tested and published in this field. It was possible to have a broader understand of what needs to be included in the definition of the main concept behind this master´s thesis. There is a lot of different developments that used different algorithms in the use of AI in both prevention and detection of a cyber-attacks, along with a variety of cyber-attacks that these algorithms were tested against. The studies in the literature review covered developments of AI algorithms, feature selection and optimalization methods and testing to develop intelligent IDS systems for the prevention and detection of cyber-attacks. The Developments available gave me a clear view to define a framework to promote and support businesses in the implementation of AI in Intrusion Detection Systems to support businesses cybersecurity environment. The Proposal includes a flowchart where the user is guided through a specific sequence where the final output would be an architecture for the prevention and detection of cyberattacks. The user will get a recommendation of algorithms to use in the prevention and detection on different attack classes. This toolbox is developed as a support tool for security analyst or managers, with a certain expertise in AI and Machine Learning. Although not all businesses have that kind of information in house- it can also be an inspiration for organizations that wants to develop their skills and introduce more intelligent techniques for the protection of their cyber environment in the long term.

5.2. DEVELOPMENTS SOLUTIONS

D1: Covensky et al., (2018) proposed a tool for the differentiation of known and previously unseen malware families by using Deep Learning. Their method presents an approach for file-signature differentiation between malware of known and unknown families. A deep Neural Network were trained to classify known malware families. A second classifier were trained by using the output of the previous stage, this classifier learns to differentiate between known and novel malware families. The authors used a rich database that consists of a thousand`s of different malicious files. The model was able to achieve an accuracy of 97,7% when classifying between seen and unseen malware families (Covensky et al., 2018)

D2: Jiang et al., (2018) proposed a hybrid detection model detect four variations of ALDDoS attacks. The system aimed to improve the accuracy of detection and reduce the time and complexity training user behavior model by using time windows. The experiment conducted a comprehensive dataset for DDoS detection attacks, by applying three layers of back propagation Neural Network (BPNN) for the classification. When comparing their approach with traffic based and hybrid KNN model, their model achieved to detect DDoS attacks with accuracy and precision (Jiang et al., 2018)

D3: Fatani et al., (2021) proposed an IDS for IOT Systems using deep learning and metaheuristic (MH) optimalization algorithms. The developed system uses a Convolutional neural network (CNN) as a feature extractor technique to obtain relevant features from the input data. The authors developed a new feature selection method using a new variant of the transient search optimization (TSO) algorithm using the deferential evolution (DE) algorithm. Their results showed that the proposed TSODE significantly outperformed the traditional TSO since the application of the DE operators has improved the exploitation and exploration phases of the traditional TSO. The developed IDS scheme based on CNN and TSODE is significantly enhanced classification accuracy (Fantani et al., 2021).

D4: Chamou et al., (2019) developed an anomaly detection system combined with deep learning methods. Their developed model used network flows to classify cyber threats into multiple fully connected layers (Multi-Layer Perceptron). Though classify attacks, the authors used two deep neural networks (DNN) with multiple fully connected layers (Multi-Layer Perceptron). The result of the proposed anomaly detection implementation system was achieved with an accuracy rate of 99,97% for DDoS detection and 99,44% for malware detection. (Chamou et al., 2019).

D5: Rhode et al., 2018 developed a novel malware prediction model based on Recurrent Neural Network (RNNs) that significantly reduces dynamic detection time, to less than 5 s per file, whilst retaining the advantages of a dynamic model. This offers the new ability to develop methods that can predict and block malicious files before they execute their payload completely, preventing attacks rather than having to remedy them. The authors were able to achieve a detection accuracy of 94% with just 5 s of dynamic data using an ensemble of RNNs and an accuracy of 96% in less than 10 s, whilst typical file execution time for dynamic analysis is around 5 min (Rhode et al., 2018).

D6: Aksu et al., (2018) conducted a comparative study using the 3 machine learning techniques. They used the CICID2017 Dataset and the supervised machine learning algorithms Support Vector machine (SVM), k nearest Neighbor (KNN) and decision Three for binary classification between DDoS attacks. In order to reduce the dataset dimension to gain higher accuracy for the given techniques they used Fisher Score Algorithm in order to reduce the dataset Dimension and select the most appropriate features. They reduced the features from 80 to 30, and the non-related features were eliminated. Results showed that although the dataset was reduced by 60% selecting the best features, the success of KNN increased. (Aksu et al., 2018).

D7: Mohammadi et al., (2019) proposed a unique filter-based selection algorithm using the KK Cup 99 dataset. The algorithm was proposed with the goal of increasing the detection rate in feature selection while reducing the false positive rate. The authors also proposed an effective- feature grouping based on Linear correlation coefficient (FGLCC). This feature grouping method is well known for its simplicity and low computational cost. The proposed algorithm- FGLCC led to better result in detection, accuracy and false positive rate compared to similar algorithms. FGLCC detect attacks with an average accuracy of 91%. The authors also managed to improve the performance of the algorithm by combining it with the CFA Algorithm. With reduction of feature to only 10, that accuracy increased, and authors were able to achieve an accuracy of 95,23%. The proposed algorithm significantly reduced the false positive rate, attack detection rate and accuracy are increased for different types of attacks (Mohammadi et al., 2019).

D8: Nevrus et al., 2019 proposed an Intelligent intrusion detection System that consists of a two-stage architecture based on machine learning algorithms. The algorithms evaluated was the Random Forest, J48, Adaptive Booster and Naïve Bayes. All the algorithms had an accuracy rate of over 90%. In the first stage the IDS uses K-means to detect attacks and the second stage uses supervised learning to attacks and eliminate the number of false positives. The proposed IDS with the respective techniques resulted in a computationally efficient IDS that was able to detect and classify attacks at a 99,97% accuracy with the use of the Random Forest algorithm, while also lowering the number of false positives to 0. (Nevrus et al., 2019).

D9: Delplace et al., (2019) compared different machine learning algorithms for the aim of cyber-attack detection. Their project aimed to build and compare models that can detect botnets in a real network traffic represented by NetFlow datasets. The different algorithms were tested was the logistic regression, support vector machine, Random Forest, Gradient Boosting and a Dense Neural Network. The three algorithms; Random Forest, Gradient boosting and dense Neural Network outperformed the other algorithms in detecting botnets. (Delplace et al.,2019).

D10: Mazini et al., (2019) highlight the concerning problem with the current Intrusion Detection Systems (IDS). Although the Systems have been considered as a main component of a safe network, one of the concerns is the false alarm report of intrusion to the network and intrusion detection accuracy that happens when it is exposed to a high volume of network data. With this problem in mind, the authors proposed a new reliable method for an anomaly network-based IDS. They used Artificial bee colony (ABC) and AdaBoost algorithms in order to achieve a high detection rate with low false positive rate. The authors used the ABC Algorithm for feature selection and the AdaBoost algorithm to evaluate and classify features. Their proposed system was tested on two well-known cybersecurity datasets- NSL-KDD and ISCXIDS2012. Their simulations confirmed that the model has a significance difference from other Intrusion detection Systems. The model accomplished great performance in different attack-based scenarios and achieved a good accuracy and detection rate. (Mazini et al., 2019).

D11: Another development using Machine learning algorithm is proposed by Park et al., (2018). The authors proposed an IDS for the classification of malware, unknown attacks and Shellcode. The authors analyzed the recognition performance by applying the Random Forest Algorithm to various datasets constructed from Kyoto 2006+ dataset. This development showed a great performance rate in the detection of unknown attack class with an accuracy of 90%. This suggest that there is a distinct pattern for the unknown attacks (Park et al., 2018).

Throughout the years different machine learning deep learning-based approaches have been proposed for designing defensing mechanisms against various phishing attacks. Recently, researchers showed that phishing attacks can be performed by employing a deep neural network-based phishing URL generating system called Deep Phish. To prevent this kind of attacks (Sameen et al., 2021) proposed an ensemble machine learning-based phishing detection system called Phish haven to identify AI- generated and human-crafted phishing URLS. Through experiments they analyzed their solution with a dataset if 100% phishing and normal URL´s. The results showed that their solution can achieve 98% accuracy, outperforming existing lexical human-crafted phishing URL´s detecting systems. Theoretical analyses also showed that their solution could detect future AI- generated phishing URL´S based on their selected lexical features with 100% accuracy.

D12: Seo et al., (2021) proposed a Real- Time Network intrusion prevention system based on Hybrid Machine learning. Their system is built upon a two-level detection approach supporting real-time processing with a high detection accuracy. The authors proposed a two-level intrusion detection approach supporting real-time processing with a high detection accuracy. It exploits packet- and flow-based classifiers to compensate for the performance and accuracy. The level 1 classifier extracts some selected features from the packet first to promote the fast classification, achieving real-time attack detection. The level 2 classifier only handles flows that were not classified by the level 1 classifier; therefore, the traffic is small enough to be processed by a time-intensive machine-learning-based classifier. Such a unique structure of the two-level classifier can provide classification speed and accuracy simultaneously.

D13: Amarasinghe et al., 2019 proposed a vulnerability detection, prevention and prediction system based on the development of a rich database that was developed with a variety of attacks using data mining techniques. The proposed approach is an automated system that consists of a mechanism to deploy vulnerabilities and a rich database with known vulnerabilities. The Convolutional Neural Networks detects the vulnerabilities, and the artificial intelligence-based generative models do the prevention process and improves reliability.

D14: Patil et al., 2022 proposed an Intrusion Prevention System named" JARVIS" to prevent and detect Dos attacks from traffic data. The developed model- JARVIS is a machine learning model deployed by using Random Forest. The model detected any attacks and suggested rules that can be deployed on snort (an open-source real time intrusion prevention and detection system). The model has can detect any attacks and suggest rules that can be deployed on Snort to prevent the attacks. Their developed machine learning model successfully detected incoming attacks with good accuracy and suggested rules in the interface which allowed the user to deploy them and prevent the attack from causing further damage (Patil et al., 2022).

D15: Krishna et al., 2020 proposed an Intrusion detection and prevention system that resulted in efficient and rapid detection of DOS, Probe, R2L and U2R attacks. The system was also proven to prevent these attacks. The intrusion is detected using the Deep Learning model- Multi-Layer Perceptron. Appropriate data from network is captured and thereby stored as a csv file and fed to the implemented Deep learning model to predict attacks in a real time manner. Secondarily, intrusion is prevented using a script that runs in the background. A script is developed to perform the prevention phase by taking appropriate decisions on the prevention function to be performed for DDoS, U2R, R2L and Probe attacks. The decision of the prevention is made by using data from the classification part. This is achieved through the Multi-Layer Perceptron model. The proposed Intrusion Detection System and the Intrusion Prevention system are combined as a single system to achieve the aim of intrusion detection and prevention tasks in a faster and efficient manner (Krishna et al., 2020).

D16: Igbe et al., (2020) Proposed an IDS a system for detecting denial of service (DoS) attacks in a network using a combination of the dendritic cell algorithm (DCA) and the negative selection algorithm (NSA). The proposed system classifies incoming network traffic into either of two classes: "normal" or "DoS attack." The authors approach follows a majority voting technique by creating multiple instances of the DCA and the NSA algorithm and assigning weights to their respective output. The results showed that the system is very effective in detecting DoS/DDoS attacks with very high accuracy (Igbe et al., 2020).

D19: Song et al., (2021) proposed a feature optimization technique for AI-based approaches to increase the accuracy of malicious PowerShell script detection. The authors statically analyze the PowerShell script and preprocess it with a method based on the tokens and abstract syntax tree (AST) for feature selection. The tokens and AST represent the structure of the PowerShell script. Performance evaluations with optimized features achieved a detection rate of 98% in both machine learning (ML) and deep learning (DL) experiments. Among the different ML and DL algorithms tested, the ML model with the 3-gram of selected five tokens and the DL model based on AST 3-gram deliver the best performance (Song et al., 2021).

D18: Aung, et al., (2018) proposed a Hybrid Intrusion Detection System using two data mining algorithms (K-Means and K-nearest Neighbors) to classify malicious and normal activities. The algorithms were able to reduce time complexity of the system. The system detects DoS, Probe, R2L, U2R with great accuracy (Aung et al., 2018).

Specific developments- Database and Cloud Security developments

There is also a big concern that can lead to several losses in businesses' cloud computing environment. Intrusion detection is one of the technologies to protect a cloud computing environment for malicious attacks.

D19: Wang et al., (2020) proposed a Hybrid system that uses a stacked constructive auto-encoder (SCAE) for feature reduction and the SVM classification algorithm for the detection of malicious attack in the cloud. They used the NSL-KDD and KDD Cup 00 Intrusion detection dataset with great performance results. (Wang et al. 2020).

D20: (Said et al.,2020). The authors proposed a Danger theory (DT) based database intrusion System to identify abnormal insider user behavior to prevent and mitigate data breach. The authors proposed a hybrid immune algorithm based on DT-ID Algorithm to enhance the performance of DT-DIDS with good results. (Said et al., 2020).

D21: Javadpour et al., (2017) proposed a new method for IDS to increase the accuracy of Intrusion Detection in cloud computing. The authors combined the methods of feature selection of linear correlation and mutual information with great results. In their experiment, the KD99 Database was used. The authors tested different classification algorithms including decision tree, random forest, CART algorithm and Neural Network. Their methods obtained good results, with the best accuracy of 99,8 % with the neural network method (Javadpour et al., 2017).

D22: Staphanakis et al., (2020) proposed a novel hybrid system approach for detecting anomalies during typical cloud operation. Their proposed approach is using the Self-Organizer Feature Map algorithm with multiple inputs. Their proposed method is based on several monitoring sites of the network and using them to train a SOFM. The results indicate that the proposed method performs well in comparison with conventional techniques (Staphanakis et al., 2020).

D23: Gao et al., (2018) proposed the FSSL- EL algorithm for network intrusion detection on Cloud Based Robotic System tested on the NSL-KDD dataset. Their approach is based on a novel fuzziness-based semi-supervised learning approach via ensemble learning. Their results deliver a promising result compared with other state-of-the-art methods (Gao et al., 2018).

5.3. FRAMEWORK FOR DEVELOPMENT OF AI IN CYBERSECURITY

Finding a way to implement AI to prevent and detect different attack classes in an IDS is presented as a framework that businesses can use a guide when introducing or advancing its usage of AI techniques in their systems. The Framework is proposed as a dynamic framework. Since the purpose of AI is to always evolve. The steps and descriptions will follow below. This research will not enter in the details of the implementing the process. This decision is made since that requires an investigation of each company´s requirements and would be out of the scope of this research.

The framework highlights the developments steps that is important when building Intelligent models in the prevention and detection of Cyberattacks. The insights and inspiration are based on the highlighted steps provided in the literature review.

Step 1: Retrieve Network Traffic Data

Step 2: Monitoring

Step 3: Identify Cyberattack(s)

Step 4: Data preprocessing- Choose feature extraction and selection method

Step 5: Training/Learning

Step 6: Requirement analysis- Functional and Nonfunctional Requirements

Step 7- Choose classification algorithms for detection/prevention

Step 8- Define Architecture

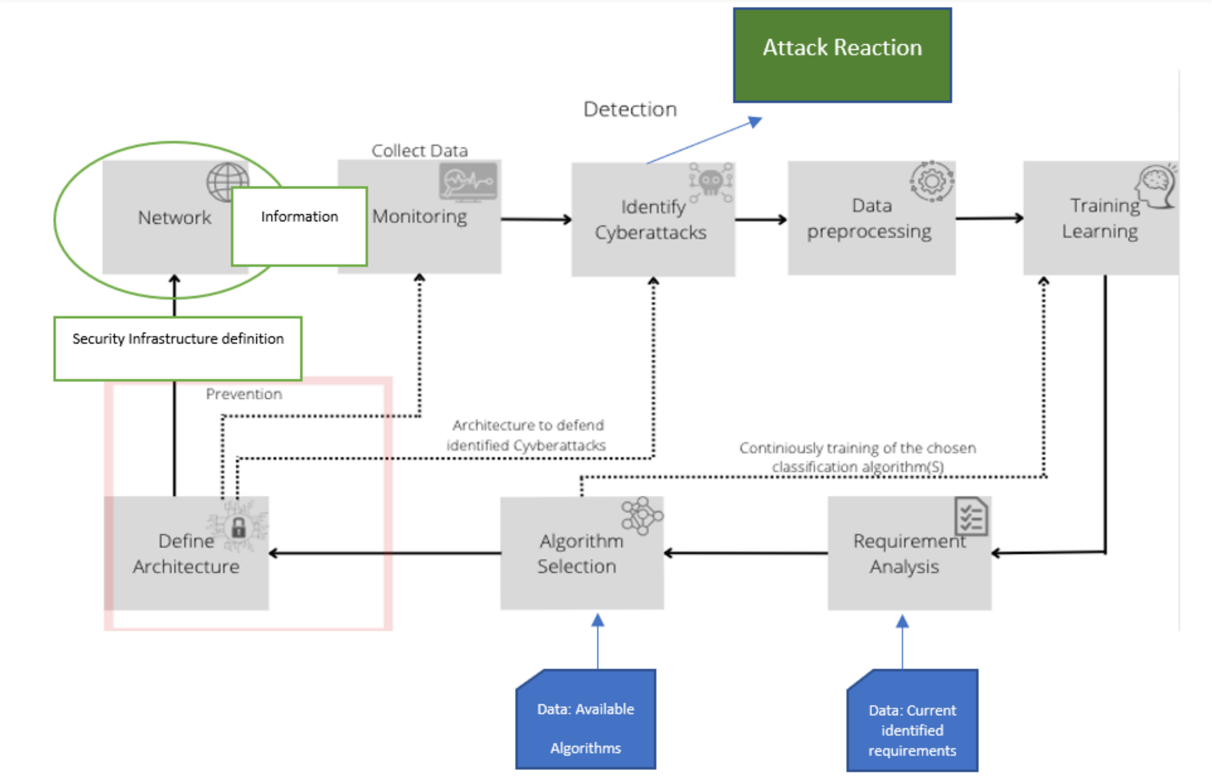

Attack Reaction
Detection
Collect Data
Network
Information
Monitoring
Identify Cyberattacks
Data preprocessing
Training Learning
Security Infrastructure definition
Prevention
Architecture to defend identified Cyvberattacks
Continiously training of the chosen classification algorithm(S)
Define Architecture
Algorithm Selection
Requirement Analysis
Data: Available Algorithms
Data: Current identified requirements

Figure 4- Framework for development of AI in Cybersecurity

Step 1: Retrieve Network Traffic Data

Network Traffic Data is the amount of data which moves across an organizations network. It is necessary to retrieve, process and have a clear method of monitoring network availability and activity to identify anomalies for analysis purposes.

Step 2: Monitoring- Collect data from the network.

According to the literature review, Collection of data and monitoring is an important step to be able to do the best and accurate training for your algorithms. it's important to collect the network traffic data that is moving across the organizations network. For testing purposes, it is important to extract real intrusion data, to be able to test and train your algorithms on the correct data for it to be efficient. There is a variety of large datasets available for the purpose of testing algorithms for detection of a various of different attacks. The monitoring of data is the purpose of an Intrusion Detection system. IDS monitors analyze network traffic that enters or exists from network devices of an organization. The Intrusion Detection Systems raise alarms if an intrusion is observed.

To be able to train and learn the developed algorithms for prevention/ detection it is important to pick a suitable dataset. A big risk when implementing data for the to the development of cyberattacks is the available data to test your algorithms. The dataset needs to contain cybersecurity data and contain the attack types that the organization wants to target and be big enough so that the algorithms can have enough testing so that your results can be accurate and be implemented to the real world of cyberattack, ones the organizations decide to run the development and implement the defense algorithms into its own systems and data. Do be able to do the best and accurate training for your algorithms it's important to extract the right and most relevant attributes from the dataset that is explained further in the Data PreProcessing step.

Step 3: Identify Cyberattack(s)

The organization needs to identify the cyberattack it is facing or want to find controls to for defense or prevention, that is important because different types of cyber-attacks need different types of technologies and methods for defense. According to the LR there is a lot of differences regarding accuracy and success detection according to which type of attack, needs different types of steps and processes.

Step 4: Data Preprocessing

The preprocessing step is an essential step for all implementations of algorithms. The preprocessing is the step of preparing the data. The data preprocessing step consists of different methods such as consists of different techniques such as data cleaning, instance selection, normalization, one hot encoding, transformation and feature extraction and selection techniques. The product of data preprocessing is the final training set. Data preprocessing techniques enables to machine learning algorithms process the data for the creation of a predictive model. Moreover, these techniques usually increase the accuracy of the model. Therefore, the data must be cleansed from errors and transformed the raw data into a predefined format. (Patil et al., 2020)

Step 5: Training- Learning

The training stage is where the algorithms is learned on a training dataset. The set of data used for training should consist of a variety of real intrusions. This is important for the model to be able to train to distinguish between intrusions and anomalies from normal network traffic. The Learning phase should result in a model that is able to determine whether the network traffic data is normal or an intrusion. (Aung et al., 2018). This step is important to build an accurate model for the purpose of the detection and prevention of different attack classes.

Step 6: Requirement Analysis

As of all software developments projects, it is important to access a requirement analysis, that is common practice in all software development projects. When the organizations decide to implement AI techniques in its cyber environment it is Define a clear path for the development process regarding functional and non-functional requirements on how the solution should work, and non-functional requirements for the solution.

Step 7: Algorithm(s) Selection

Pick the right classification algorithm for the detection of the cyberattack(s) according to the knowledge from the learning and testing stage. In that previous step data was tested and trained with different algorithms and the accuracy rates is extracted.

Step 8: Define Architecture

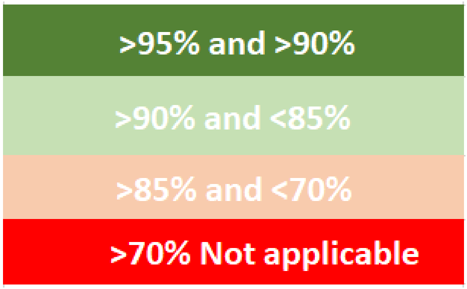
After all those previous steps, the architecture of the defense and prevention mechanisms is ready. This procedure is a continuously step since AI algorithms is continuously learning. The architecture is also active in the monitoring of data that enters the system, and will identify cyberattacks, this stage is also part of the prevention of cyberattacks.

5.4. Toolbox Application

Conditions are met to propose a toolbox for AI implementation guide to help organizations transform their traditional cybersecurity environment into an intelligent one using Intelligent (AI) algorithms. This framework will help to select the most adequate technologies to be recommended according to which cyberattacks the organizations if facing and what is their targeted system.

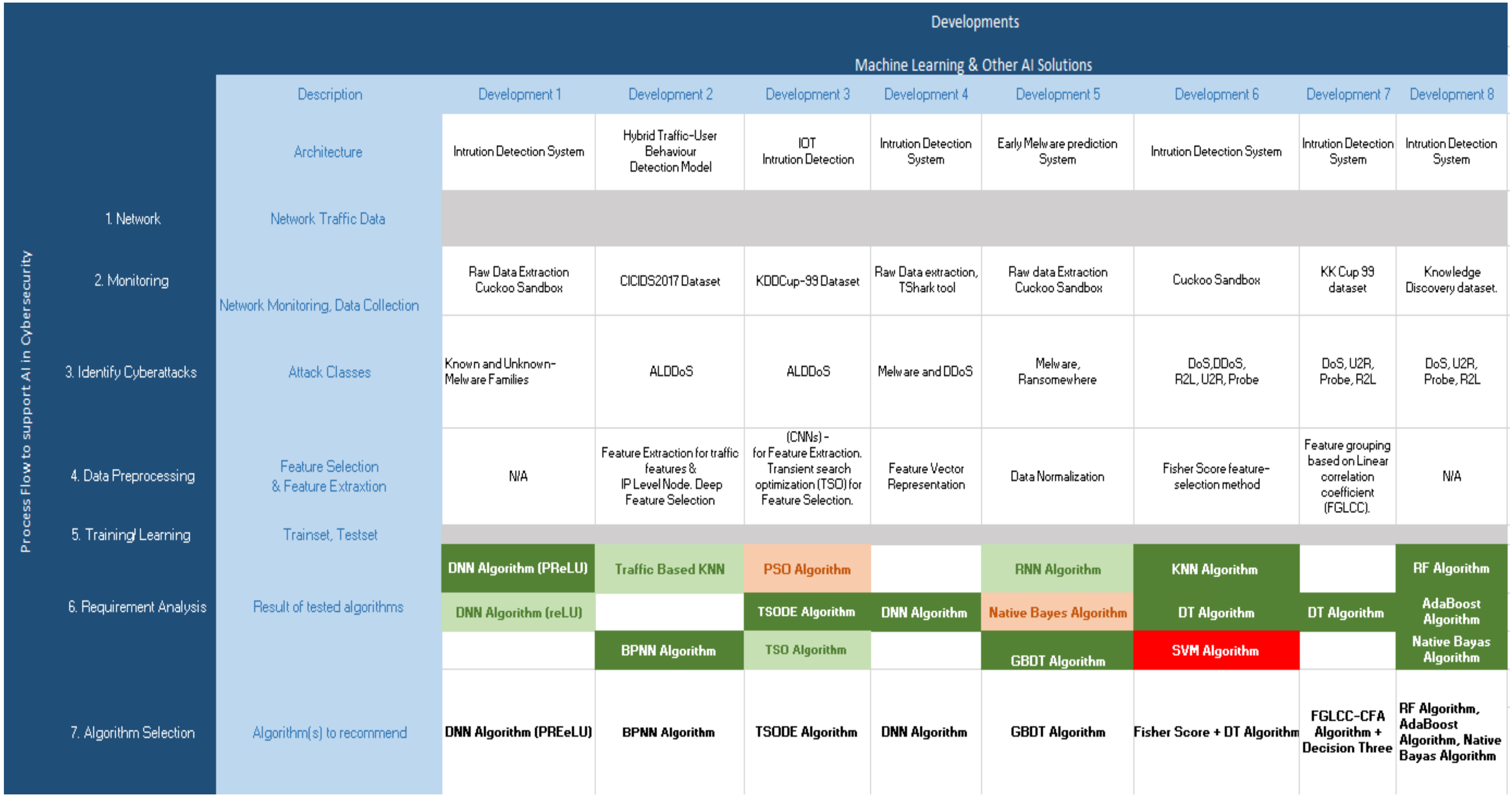
This process requires some steps that is reflected in a figure 4: The below conceptual toolbox presents the process of developing AI algorithms to use for the prevention and detection of cyberattacks are paired with relevant Algorithms and their area of usage and architecture. It shows the typology between the area of use and the algorithms capacity found and corresponding the cybersecurity infrastructure solutions. The toolbox works as follows: the user is required to follow the structure as in the proposed framework. (1) The user will see an architecture presented in the first row and in the second step (2) the toolbox will presents datasets and data collection methods that has been used for training and testing according to which (3) cyber-attack that final algorithm is proposed to detect. (4) Alongside the sequence the user will be recommended different types of data preprocessing techniques and different algorithms that has been used I the testing of detection and or/prevention of different types of attacks, alongside with its accuracy. The algorithms used in the prevention and or detection of cyber-attacks is classified through a color scheme with the green (marks the best accuracy) and red (marks the weakest accuracy). The accuracy marked in red is classified in this toolbox as not applicable to use. In the last step of the toolbox, an recommendation is made for the type of algorithm(s) to use for specific architectures and cyberattacks. The last line of algorithms presents the best algorithm for its presented and is therefore the one to recommend.

The Step (1) Network Traffic Data and (5) Training/Learning is marked in grey, as there is no methods or techniques included in these steps. The decision to exclude these from the study is the fact that step (1) Network Traffic Data apply to all the developments presented. For Step (4) the reason is that this step is very individual for each development phase and therefore it is not suitable to include for the purpose of this Toolbox. The purpose of the toolbox is to be transparent and show available solutions with great results for various type of attack and technologies.



🟩 >95% and >90%
🟩 >90% and <85%
🟧 >85% and <70%
🟥 >70% Not applicable

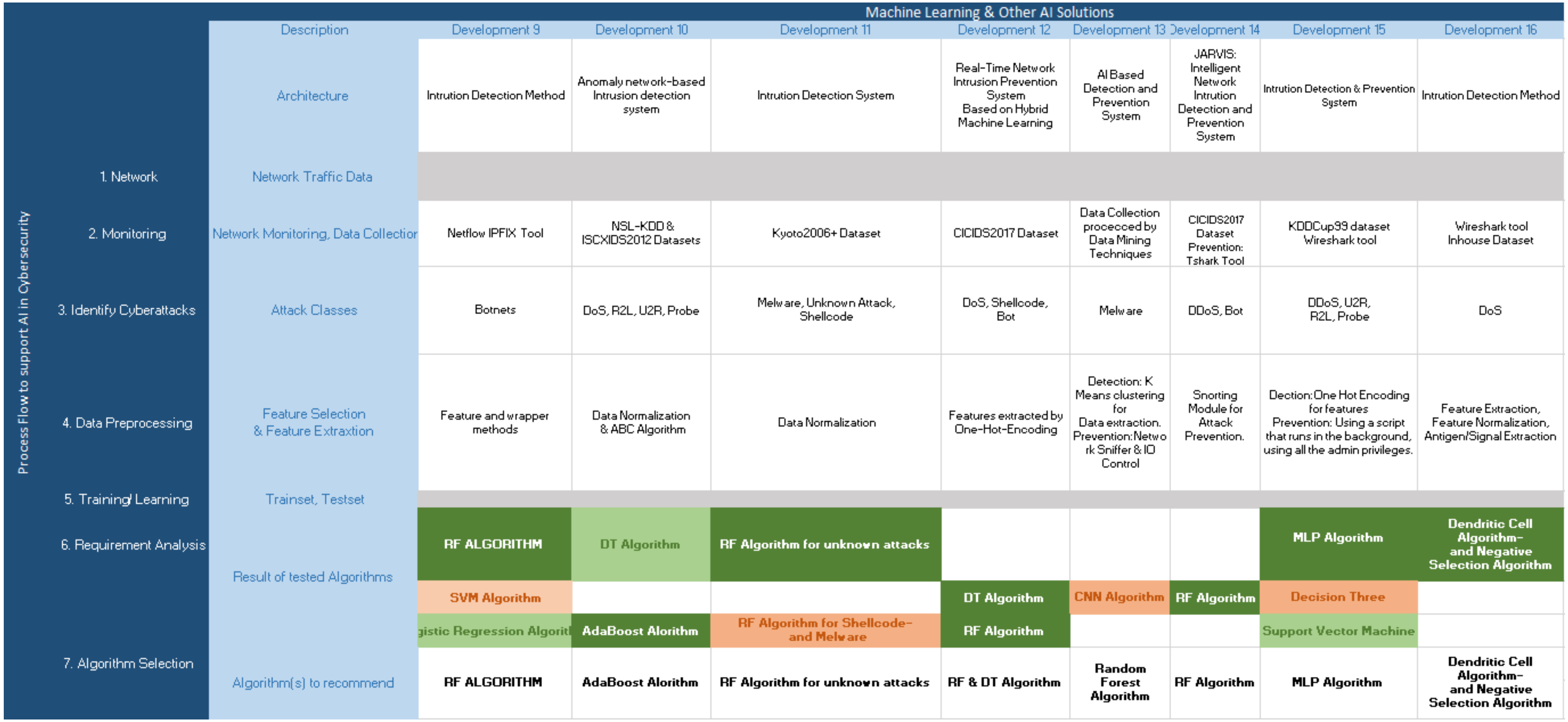
Figure 5- Toolbox- Color Scheme



Process Flow to support AI in Cybersecurity

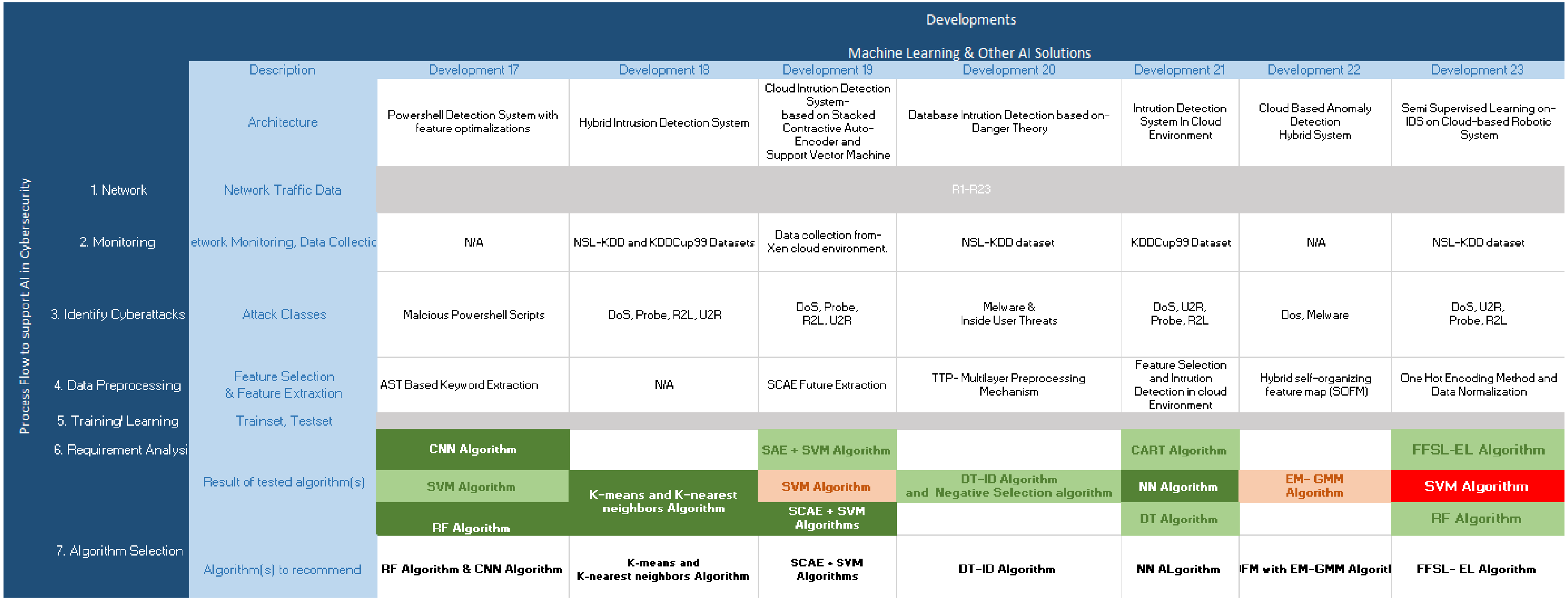
		Developments Machine Learning & Other AI Solutions							
	Description	Development 1	Development 2	Development 3	Development 4	Development 5	Development 6	Development 7	Development 8
	Architecture	Intrution Detection System	Hybrid Traffic-User Behaviour Detection Model	IOT Intrution Detection	Intrution Detection System	Early Melware prediction System	Intrution Detection System	Intrution Detection System	Intrution Detection System
1. Network	Network Traffic Data								
2. Monitoring	Network Monitoring, Data Collection	Raw Data Extraction Cuckoo Sandbox	CICIDS2017 Dataset	KDDCup-99 Dataset	Raw Data extraction, TShark tool	Raw data Extraction Cuckoo Sandbox	Cuckoo Sandbox	KK Cup 99 dataset	Knowledge Discovery dataset.
3. Identify Cyberattacks	Attack Classes	Known and Unknown-Melware Families	ALDDoS	ALDDoS	Melware and DDoS	Melware, Ransomewhere	DoS,DDoS, R2L, U2R, Probe	DoS, U2R, Probe, R2L	DoS, U2R, Probe, R2L
4. Data Preprocessing	Feature Selection & Feature Extraxtion	N/A	Feature Extraction for traffic features & IP Level Node. Deep Feature Selection	(CNNs) - for Feature Extraction. Transient search optimization (TSO) for Feature Selection.	Feature Vector Representation	Data Normalization	Fisher Score feature-selection method	Feature grouping based on Linear correlation coefficient (FGLCC).	N/A
5. Training/ Learning	Trainset, Testset								
6. Requirement Analysis	Result of tested algorithms	DNN Algorithm (PReLU)	Traffic Based KNN	PSO Algorithm		RNN Algorithm	KNN Algorithm		RF Algorithm
		DNN Algorithm (reLU)		TSODE Algorithm	DNN Algorithm	Native Bayes Algorithm	DT Algorithm	DT Algorithm	AdaBoost Algorithm
			BPNN Algorithm	TSO Algorithm		GBDT Algorithm	SVM Algorithm		Native Bayas Algorithm
7. Algorithm Selection	Algorithm(s) to recommend	DNN Algorithm (PREeLU)	BPNN Algorithm	TSODE Algorithm	DNN Algorithm	GBDT Algorithm	Fisher Score + DT Algorithm	FGLCC-CFA Algorithm + Decision Three	RF Algorithm, AdaBoost Algorithm, Native Bayas Algorithm

Figure 6- Toolbox Application for AI Algorithms in Cyber attacks Prevention and Detection (1)



Process Flow to support AI in Cybersecurity	Machine Learning & Other AI Solutions								
	Description	Development 9	Development 10	Development 11	Development 12	Development 13	Development 14	Development 15	Development 16
	Architecture	Intrution Detection Method	Anomaly network-based Intrusion detection system	Intrution Detection System	Real-Time Network Intrusion Prevention System Based on Hybrid Machine Learning	AI Based Detection and Prevention System	JARVIS: Intelligent Network Intrution Detection and Prevention System	Intrution Detection & Prevention System	Intrution Detection Method
1. Network	Network Traffic Data								
2. Monitoring	Network Monitoring, Data Collectior	Netflow IPFIX Tool	NSL-KDD & ISCXIDS2012 Datasets	Kyoto2006+ Dataset	CICIDS2017 Dataset	Data Collection proceeced by Data Mining Techniques	CICIDS2017 Dataset Prevention: Tshark Tool	KDDCup99 dataset Wireshark tool	Wireshark tool Inhouse Dataset
3. Identify Cyberattacks	Attack Classes	Botnets	DoS, R2L, U2R, Probe	Melware, Unknown Attack, Shellcode	DoS, Shellcode, Bot	Melware	DDoS, Bot	DDoS, U2R, R2L, Probe	DoS
4. Data Preprocessing	Feature Selection & Feature Extraxtion	Feature and wrapper methods	Data Normalization & ABC Algorithm	Data Normalization	Features extracted by One-Hot-Encoding	Detection: K Means clustering for Data extraction. Prevention: Network Sniffer & IO Control	Snorting Module for Attack Prevention.	Dection: One Hot Encoding for features Prevention: Using a script that runs in the background, using all the admin privileges.	Feature Extraction, Feature Normalization, Antigen/Signal Extraction
5. Training/ Learning	Trainset, Testset								
6. Requirement Analysis	Result of tested Algorithms	RF ALGORITHM	DT Algorithm	RF Algorithm for unknown attacks				MLP Algorithm	Dendritic Cell Algorithm- and Negative Selection Algorithm
		SVM Algorithm			DT Algorithm	CNN Algorithm	RF Algorithm	Decision Three	
		gistic Regression Algorit	AdaBoost Alorithm	RF Algorithm for Shellcode- and Melware	RF Algorithm			Support Vector Machine	
7. Algorithm Selection	Algorithm(s) to recommend	RF ALGORITHM	AdaBoost Alorithm	RF Algorithm for unknown attacks	RF & DT Algorithm	Random Forest Algorithm	RF Algorithm	MLP Algorithm	Dendritic Cell Algorithm- and Negative Selection Algorithm

Figure 7- Toolbox Application for AI Algorithms in Cyber attacks Prevention and Detection (2)



			Developments						
			Machine Learning & Other AI Solutions						
		Description	Development 17	Development 18	Development 19	Development 20	Development 21	Development 22	Development 23
		Architecture	Powershell Detection System with feature optimalizations	Hybrid Intrusion Detection System	Cloud Intrution Detection System- based on Stacked Contractive Auto-Encoder and Support Vector Machine	Database Intrution Detection based on-Danger Theory	Intrution Detection System In Cloud Environment	Cloud Based Anomaly Detection Hybrid System	Semi Supervised Learning on-IDS on Cloud-based Robotic System
Process Flow to support AI in Cybersecurity	1. Network	Network Traffic Data	R1-R23						
	2. Monitoring	etwork Monitoring, Data Collectic	N/A	NSL-KDD and KDDCup99 Datasets	Data collection from-Xen cloud environment.	NSL-KDD dataset	KDDCup99 Dataset	N/A	NSL-KDD dataset
	3. Identify Cyberattacks	Attack Classes	Malcious Powershell Scripts	DoS, Probe, R2L, U2R	DoS, Probe, R2L, U2R	Melware & Inside User Threats	DoS, U2R, Probe, R2L	Dos, Melware	DoS, U2R, Probe, R2L
	4. Data Preprocessing	Feature Selection & Feature Extraxtion	AST Based Keyword Extraction	N/A	SCAE Future Extraction	TTP- Multilayer Preprocessing Mechanism	Feature Selection and Intrution Detection in cloud Environment	Hybrid self-organizing feature map (SOFM)	One Hot Encoding Method and Data Normalization
	5. Training/ Learning	Trainset, Testset							
	6. Requirement Analysi	Result of tested algorithm(s)	CNN Algorithm		SAE + SVM Algorithm		CART Algorithm		FFSL-EL Algorithm
			SVM Algorithm	K-means and K-nearest neighbors Algorithm	SVM Algorithm	DT-ID Algorithm and Negative Selection algorithm	NN Algorithm	EM- GMM Algorithm	SVM Algorithm
			RF Algorithm		SCAE + SVM Algorithms		DT Algorithm		RF Algorithm
	7. Algorithm Selection	Algorithm(s) to recommend	RF Algorithm & CNN Algorithm	K-means and K-nearest neighbors Algorithm	SCAE + SVM Algorithms	DT-ID Algorithm	NN ALgorithm	FM with EM-GMM Algorit	FFSL- EL Algorithm

Figure 8- Toolbox Application for AI Algorithms in Cyber attacks Prevention and Detection (3)

5.5. EVALUATION

In order to this step of the design science Research (DSR) technology, the validation phase is perused. The main objective of the validation step was to retrieve thoughts, views and validation of the proposed flowchart and toolbox, as well as insights from expert working in the field of cybersecurity and discuss the important of the prevention and detection of cyber-attacks in a business, and the current use of AI and Machine Learning Models in a business's Environment. As a result, it was arranged three interviews with field experts to discuss and validate the proposed flowchart and toolbox. The interview participants were a Vulnerability Management Team Leader with 5 years of experience within Cyber Security, an operational engineer with 7 years' experience of maintain cybersecurity systems, the Third person were a person graduated from a master's degree of information security with 2-year experience.

	Individual Interviews	
Participants	Job Role	Years of Experience
1	Vulnerability Management Team Leader	5 years
2	Operational Engineer	7 years
3	Information Security Graduate	2 years

Table 6- Individual Interview`s Participants

Interview Questions	
1.	Do you consider the proposed framework as useful and why? If not, why do you believe it is not?
2.	Would you consider implementing the proposed framework?
3.	Do you consider the proposed toolbox as useful and why?
4.	Would you consider implementing the proposed toolbox?
5.	Do you have any recommendation or suggestions for? further improvements of the proposed flowchart and toolbox?
6.	Would you like to share your professional experience in this area?

Table 8- Interview Questions

1) **Do you consider the proposed flowchart as useful and why? If not, why do you believe it is not?**

- Participant 1: “From an operational point of view I think it could be useful to implement to present a standardized process for implementing AI Algorithms to detect and prevent attacks in an organization. It was simple, yet understandable and effective way to visualize a workflow for the implementation of AI to prevent and detect attacks. “
- Participant 2: “I agree with framework, and is a good way to present a process flow for the implementation of AI algorithms and additional techniques to engineers when implementing AI models in detection of cyber-attacks, although in a real environment it is also important to consider start the process by doing a risk analysis, that is an important first step when building a Cybersecurity architecture”

- Participant 3: “From a theoretical point of view, I believe it is a good and efficient way to visualize and highlight the steps that is important to consider when implementing AI in Cybersecurity.

2) Would you consider implementing the proposed framework?

- Participant 1: “From an operational point of view, as an operational engineer- I think the flowchart could be useful and I would consider implementing it”.
- Participant 2: “The framework reflects the step of implementing a standardized way to testing and learning AI algorithms to use in an IDS System in a correct way. I agree with the process flow and the fact that it is a dynamic process. I believe the proposed flowchart could be very useful for researchers in this field”. For developers that wants to implement a Machine Learning model for the detection and prevention of attacks in a testing environment, it is a good framework. Although, from my perspective as a manager, there is more steps to consider in a framework for this purpose in the real world. The most important steps to consider is the risk analysis at the beginning of any cyber security development projects, and an architecture is very expensive to implement”.
- Participant 3: Yes, I agree with the processes in the framework and would consider to implement it.

3) Do you consider the proposed toolbox as useful and why? If not, why do you believe it is not?

- Participant 1: “Yes, I believe it is a good fit for its purpose”. I think it is the simplest way to present the process”. To develop a toolbox is an agile process and it will develop over time, as of now I agree with the representation of the toolbox”. As the toolbox develops it could include more stages. I like the way it represents the different attacks and the techniques and algorithms that is most effective to detect it”. The toolbox presents an effective way to lookup the most effective algorithms available”. In the future I would consider adding a new step within the toolbox and that is to categorize the cyber-attacks inside the OSI module. That would make it clearer to managers in which layer that is attacked.
- Participant 2: “Yes, the toolbox can be useful for developers and researchers within the field”. I like the approach and it is an efficient way to look up a problem and find a solution when a development is considered in an organization”. Apart from that, I would prefer the first step to be the objective of the implementation and the risk assignment, and how this is measured”. The topic of implementing AI algorithms in cybersecurity is very interesting and demanding, and it involves a lot of risk and additional cost. It is also very hard to implement

new techniques and algorithms in an existing system, for this purpose the business would need to implement a whole new system. “

- Participant 3: “I believe the toolbox could work well for its purpose. I think it is a very transparent way to show different techniques that has proven to deliver good results. This can be a great starting point for researchers, but also organizations that is considering implementing intelligent techniques in their cyber security strategy. The development in the toolbox is also very interesting and I am surprised by the algorithm`s good results.”

4) Would you consider implementing the proposed toolbox?

- Participant 1: “Yes, for the purpose of showing the available algorithms to be used for classification of different attacks and categorize the most efficient ones. This can be helpful in the research stage in an organization, and for Developers. You proposed a simple and accurate view of the development steps for the implementation phase.”
- Participant 2: “Yes, it is simple to understand and look up algorithms and methods to use to attack and prevent an attack. This can save developers for time since it reduces the time of research.”
- Participant 3: “I believe the toolbox is a good support for security analysts when considering implementing new intelligent techniques for attacks detection. The toolbox is easy to understand, and it represents the available methods out there in an efficient way. I especially like those recommendations of algorithms at the end, so that the viewers can see the most effective algorithms for different purposes.

5) Do you have any recommendation or suggestions for further improvements of the proposed flowchart and toolbox?

- Participant 1: For the framework, I would like to risk assignment as a first step.
  Regarding the toolbox, it could be nice to see a more detailed view for the attack classes- example malware attacks can be very different and divided into many different attacks with different signatures. It could also be very interesting to go deeper into the preprocessing steps, as you have included the techniques, I would also like to see the features selected to have a view of which features that is effective when classifying different attacks.
- Participant 2: In the toolbox, it could be an idea to add the ISO- OSI Model. That can help managers to see in which type of layers the attacks appear in and help them to identify a decision.

- Participant 3: Naturally, as the technologies evolves- the toolbox should be developed with more techniques and algorithms. The field develops very quickly and therefore it would be naturally that this toolbox will need include more attacks Solutions' and additional of algorithms in the future.

5.6. Discussion

In this section, one will perform an analysis of the inputs and feedbacks obtain during the previous section will incorporate to reach a conclusive discussion on the evaluation of the proposed framework and toolbox that is proposed to contribute within the study of AI algorithms used in cyberattacks prevention and detection.

The analysis will be performed based on value, feasibility of implementation and improvements based on feedback collected from the validation phase. Then an extensive evaluation of the proposed framework will be done on the collective knowledge based on the previous analysis.

Regarding the value of the proposed framework and the additional toolbox, all the interview participants agreed that the proposals is very valuable for its purpose of contribute to extensive knowledge in the field and can be a useful source for developers and researchers in the field.

All the participant´s with their different expertise and role in the field of cybersecurity agreed on the lack of resources to look up useful solutions to their problems and the lack of applications to use as a reference tool when starting on an AI journey within the protection of sufficient attacks. While companies have few resources dedicated to this type of work, when they decided to move their traditional cyber security solutions into more AI centered solutions, the existence of the proposals made in this study, can reduce a lot of resource cost for organizations.

Summarizing, the proposals was considered valuable to promote automatic processes and implementation of intelligence algorithms in the field of cybersecurity.

Regarding the suggestions for improvements the interview participants agreed on that the proposals could include more details in the futures, especially the toolbox where more attacks classes could be considered, and there would also be natural to develop the toolbox with more intelligent algorithms once they are tested and accepted in the research field. On the management side, the cost and risk assessment should also be in consideration to make the decision process of implementing this framework and toolbox into an organization.

6. CONCLUSION

The objective of this study was to propose a framework and an additional toolbox to contribute to foster the process of implementing AI algorithms in the prevention and detection of cyber-attacks was achieved.

After a careful review of a systematic literature review, a framework to build an infrastructure for AI algorithms and methods in the prevention and detection of attacks. A toolbox that contained methods and algorithms to use for the prevention and detection of different cyber-attacks were also proposed. The toolbox was built to reduce developers with the research cost of finding the right intelligent methods and algorithms to best fit their problem.

To evaluate the proposals, interview participants with different expertise in the field of cyber security provided feedback and valuable insights.

To conclude this study, the process of implementing AI in Cybersecurity is a very in-demand topic for organizations. It is identified with the problem of the little existence of fully trustworthy methods and practices for the implementation. Therefore, when interviewing specialists in the area, a great enthusiasm and adherence in responding, participating, and helping to improve the method proposed in this work was received.

6.1. LIMITATIONS

Regarding the developed framework and toolbox, it is important to mention that in addition to the attributes outlined, there are also limitations inherent such as:

- As the literature review is based on a collection of experimental developments within the field and not tested in organizations- the processes in the framework may be a little laboratory, in organizations cyber security architecture there will be more processes to consider before considering adapting specific intelligent methods or systems, such as resources available, cost of the implementation and risk.
- The developments included in the toolbox is experimental developments and is not tested live as of my knowledge from the literature review. Therefore, there will be limitations on its feasibility in real system environments. With that mentioned, the results are tested in realistic environment with large datasets of real cybersecurity data to mirror the reality in cyber security environments as close as possible.
- Cost of the implementation of artificial intelligence. The implementation of a security architecture such as an Intrusion Detection Systems is very expensive, and there is a lot of considerations to make that is out of the scope of my thesis.

- In order to be able to make a detailed analysis of the effectiveness of the proposed methods, it should have been applied in multiple contexts where later the results should be evaluated. However, in the context of this work, it was not possible to reach due to the necessary time window.

- To obtain more reliable feedback and closer to reality, the interviews could include more specialists within the field of cyber security, however, this work represents only a starting point and can be extended to a deeper study.

6.2. Recommendations for Future Developments

For the future work, an obvious important pinpoint is that the framework and toolbox is not tested in an real environment- inside an organization. Regarding other recommendations for future work, it would be interesting to add the more information regarding the attack types and its influence on the organizations- Such as the ISO-OSI Model and its layer so that the possible users of the toolbox could have a clearer view of where in the information layers the attack is affecting- this could be especially useful for managers to have a clearer view of the cyber-attacks affection on their systems and how to prevent and detect them. There could also be useful to add risk assignments on each attack types and on individual organizations level. That can be an additional help for organization to choose the additional controls regarding the intelligent methods adoptions. There could also be interesting to preform analysis of the cost of the specific methods included in the toolbox.

BIBLIOGRAPHY

Abdiyeva-Aliyeva, G., Hematyar, M., & Bakan, S. (2021c). Development of System for Detection and Prevention of Cyber Attacks Using Artifıcial Intelligence Methods. *2021 2nd Global Conference for Advancement in Technology (GCAT)*. https://doi.org/10.1109/gcat52182.2021.9587584

Accenture. (2020b, August 14). *Eighth Annual Cost of Cybercrime Study*. Accenture. https://www.accenture.com/us/en/insights/security/eight-annual-cost-cybercrime-study

Agrafiotis, I & Nurse, J & Goldsmith, M & Creese, S & Upton, D. (2018) A taxonomy of cyber-harms: Defining the impacts of cyber-attacks and understanding how they propagate*, Journal of Cybersecurity, Volume 4, Issue 1, 2018, tyy006.* https://doi.org/10.1093/cybsec/tyy006

Ahmad, A. B. Abdullah and A. S. Alghamdi, Remote to Local attack detection using supervised neural network, *2010 International Conference for Internet Technology and Secured Transactions, 2010,* pp. 1-6.

Krishna, A. Lal M.A., A. J. Mathewkutty, D. S. Jacob and M. Hari*,* Intrusion Detection and Prevention System Using Deep Learning, *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020*, pp. 273-278. doi: 10.1109/ICESC48915.2020.9155711.

Aksu, D., Ustebay, S., & Aydin, M.A & Atmaca, T. (2018). Intrusion Detection with Comparative Analysis of Supervised Learning Techniques and Fisher Score Feature Selection Algorithm. In *International symposium on computer and information sciences* (pp. 141-149). 10.1007/978-3-030-00840-6_16.

Amarasinghe, A. M. S. N., Wijesinghe, W. A. C. H., Nirmana, D. L. A., Jayakody, A., & Priyankara, A. M. S. (2019). AI Based Cyber Threats and Vulnerability Detection, Prevention and Prediction System*.* In *2019 International Conference on Advancements in Computing (ICAC)* (pp. 363-368). IEEE.

A. M. S. N. Amarasinghe, W. A. C. H. Wijesinghe, D. L. A. Nirmana, A. Jayakody and A. M. S. Priyankara, "AI Based Cyber Threats and Vulnerability Detection, Prevention and Prediction System. In 2019 *International Conference on Advancements in Computing (ICAC)* (pp. 363-368). 10.1109/ICAC49085.2019.9103372.

Aung, Y. Y., & Min, M. M. (2018). Hybrid intrusion detection system using K-means and K-nearest neighbors' algorithms. In *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)* (pp. 34-38). IEEE.

Belani, M. (2021). The Use of Artificial Intelligence in Cybersecurity: A Review. Computer.Org.

Brendel, A.B., Lembcke, T. B & Kolbe, L. M. (2022). Towards an Integrative View on Design Science Research Genres, Strategies, and Pivotal Concepts in Information Systems Research*. ACM SIGMIS Database*: *the DATABASE for Advances in Information Systems*, 53(4), 9-23.

Bertino, E., & Islam, N. (2017). Botnets and internet of things security*. Computer, 50(2)*, 76-79.

Chan.L., Morgan.L., Simen, H., Alshabanah., F, Ober.,D, Genry., D, Min, D., & Cao, R (2019). Survey of AI in cybersecurity for Information Technology Management. In *2019 IEEE Technology & engineering management conference (TEMSCON)*, (pp.1-8). IEE

Chronopoulos, M., Panaousis, E., & Grossklags, J. (2017). An Options Approach to Cybersecurity Investment*. IEEE Access, 6*, 12175–12186. https://doi.org/10.1109/access.2017.2773366

Conti, M & Dargahi, T & Dehghantanha, A. (2018). Cyber Threat Intelligence: Challenges and Opportunities. *Cyber Threat and Intelligence*, 1-6. https://doi.org/10.1007/978-3-319-73951-9

Cordonsky, I., Rosenberg, I., Sicard, G., & David, E. O. (2018). DeepOrigin: End-To-End Deep Learning For Detection Of New Malware Families. In *2018 International Joint Conference on Neural Networks (IJCNN)*. https://doi.org/10.1109/ijcnn.2018.8489667

Chamou, D., Toupas, P., Ketzaki, E., Papadopoulos, S., Giannoutakis, K. M., Drosou, A., & Tzovaras, D. (2019). Intrusion Detection System Based on Network Traffic Using Deep Neural Networks. In *2019 IEEE 24th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*. https://doi.org/10.1109/camad.2019.8858475

Delplace, A., Hermoso, S., & Anandita, K. (2020). Cyber Attack Detection thanks to Machine Learning Algorithms. *arXiv preprint arXiv:2001.06309.*

Devanesan, J. (2021). Will AI replace cybersecurity teams completely? Tech Wire Asia. https://techwireasia.com/2021/03/will-ai-replace-cybersecurity-teams-completely/

Embroker Team. (2021). 2021 Must-Know Cyber Attack Statistics and Trends. https://www.embroker.com/blog/cyber-attack-statistics/

Tyugu, E. (2011). Artificial intelligence in cyber defense*, In 3rd International Conference on Cyber Conflict (ICCC 2011),* pp. 1-11, 2011. IEE

Fatani, A., Abd Elaziz, M., Dahou, A., Al-Qaness, M. A. A., & Lu, S. (2021). IoT Intrusion Detection System Using Deep Learning and Enhanced Transient Search Optimization*. IEEE Access, 9, 123448–123464.* https://doi.org/10.1109/access.2021.3109081

Gao, Y., Liu, Y., Jin, Y., Chen, J., & Wu, H. (2018). A Novel Semi-Supervised Learning Approach for Network Intrusion Detection on Cloud-Based Robotic System*. IEEE Access, 6, 50927–50938.* https://doi.org/10.1109/access.2018.2868171

Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). Methods of data collection in qualitative research: interviews and focus groups*. British Dental Journal, 204(6)*, 291–295. https://doi.org/10.1038/bdj.2008.192

Harrell, M. C., & Brandley, M.A. (2009). Data Collection Methods. Semi-Structured Interviews and Focus Groups. *Rand National Defense Research Inst Santa Monica ca.*

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 75-105.

Hou, Y., Zhuge, J. W., Xin, D., & Feng, W. (2014, May). SBE− A Precise Shellcode Detection Engine Based on Emulation and Support Vector Machine. In *International Conference on Information Security Practice and Experience* (pp. 159-171). Springer, Cham.

Igbe, O., Ajayi, O., & Saadawi, T. (2017). Detecting Denial of Service Attacks Using a Combination of Dendritic Cell Algorithm and the Negative Selection Algorithm. In *2017 IEEE International Conference on Smart Cloud (SmartCloud).* https://doi.org/10.1109/smartcloud.2017.18

Jiang, J., Yu, Q., Yu, M., Li, G., Chen, J., Liu, K., Liu, C., & Huang, W. (2018). ALDD: A Hybrid Traffic-User Behavior Detection Method for Application Layer DDoS*. In 2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/ 12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE).* https://doi.org/10.1109/trustcom/bigdatase.2018.00225

Javaid, A., Niyaz, Q., Sun, W., & Alam, M. (2016, May). A deep learning approach for network intrusion detection system. *In Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS)* (pp. 21-26).

Javadpour, A., Kazemi Abharian, S., & Wang, G. (2017). Feature Selection and Intrusion Detection in Cloud Environment Based on Machine Learning Algorithms. In *2017 IEEE International Symposium on Parallel and Distributed Processing With Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC).* https://doi.org/10.1109/ispa/iucc.2017.00215

Kaja, N., Shaout, A., & Ma, D. (2019). An intelligent intrusion detection system. *Applied Intelligence, 49(9),* 3235–3247. https://doi.org/10.1007/s10489-019-01436-1

Krishna, A., Lal M.A., A., Mathewkutty, A. J., Jacob, D. S., & Hari, M. (2020). Intrusion Detection and Prevention System Using Deep Learning. *In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*. https://doi.org/10.1109/icesc48915.2020.9155711

Laato, S., Farooq, A., Tenhunen, H., Pitkamaki, T., Hakkala, A., & Airola, A. (2020). AI in Cybersecurity Education- A Systematic Literature Review of Studies on Cybersecurity MOOCs. In *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT).* https://doi.org/10.1109/icalt49669.2020.00009

Marques, P., Rhode, M., & Gashi, I. (2021). Waste not: Using diverse neural networks from hyperparameter search for improved malware detection. Computers &Amp; Security, 108, 102339. https://doi.org/10.1016/j.cose.2021.102339

Mazini, M., Shirazi, B., & Mahdavi, I. (2019). Anomaly network-based intrusion detection system using a reliable hybrid artificial bee colony and AdaBoost algorithms*. Journal of King Saud University - Computer and Information Sciences, 31(4)*, 541–553. https://doi.org/10.1016/j.jksuci.2018.03.011

McGrath, C., Palmgren, P. J., & Liljedahl, M. (2019). Twelve tips for conducting qualitative research interviews. *Medical Teacher, 41 (9)*, 1002-1006. https://doi.org/10.1080/0142159X.2018.1497149

McAfee (2020) The Hidden Costs of Cybercrime. Smith, Z & Lewis, A. https://www.mcafee.com/enterprise/en-us/assets/reports/rp-hidden-costs-of-cybercrime.pdf

Mohammadi, S., Mirvaziri, H., Ghazizadeh-Ahsaee, M., & Karimipour, H. (2019). Cyber intrusion detection by combined feature selection algorithm. *Journal of Information Security and Applications, 44,* 80–88. https://doi.org/10.1016/j.jisa.2018.11.007

Naik, B., Mehta, A., Yagnik, H., & Shah, M. (2021). The impacts of artificial intelligence techniques in augmentation of cybersecurity: a comprehensive review. *Complex & Intelligent Systems, 1-18.* https://doi.org/10.1007/s40747-021-00494-8

Namanya, A. P., Cullen, A., Awan, I. U., & Disso, J. P. (2018, August). The world of Malware: An overview. *In 2018 IEEE 6th International Conference on Future Internet of Things and Cloud (FiCloud)* (pp. 420-427). IEEE.

Igbe,O., Ajayi,O., & Saadawi, T. (2017). "Detecting Denial of Service Attacks Using a Combination of Dendritic Cell Algorithm and the Negative Selection Algorithm. *In 2017 IEEE International Conference on Smart Cloud (SmartCloud), 2017,* (pp. 72-77). doi: 10.1109/SmartCloud.2017.18.

Park, K., Song, Y., & Cheong, Y. G. (2018b). Classification of Attack Types for Intrusion Detection Systems Using a Machine Learning Algorithm. In *2018 IEEE Fourth International Conference on Big Data Computing Service and Applications (BigDataService).* (pp. 282-286). https://doi.org/10.1109/bigdataservice.2018.00050

Patil, A. P., Premkumar, H., M H M, K., & Hegde, P. (2022b). JARVIS: An Intelligent Network Intrusion Detection and Prevention System. In *2022 IEEE Fourth International Conference on Advances in Electronics, Computers and Communications (ICAECC).* (pp. 1-6). https://doi.org/10.1109/icaecc54045.2022.9716622

Paredes, J. N., Teze, J. C. L., Simari, G. I., & Martinez, M. V. (2021). On the Importance of Domain-specific Explanations in AI-based Cybersecurity Systems (Technical Report). *arXiv preprint arXiv:2108.02006.*

Rhode, M., Burnap, P., & Jones, K. (2018). Early-stage malware prediction using recurrent neural networks. *Computers &Amp; Security*, 77, 578–594. https://doi.org/10.1016/j.cose.2018.05.010

Rowley, J. (2012). Conducting research interviews. *Management Research Review*. https://doi.org/10.1108/01409171211210154

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research*. Journal of Management Information Systems, 24* (3), 45-77. https://doi.org/10.2753/MIS0742-1222240302

Said, W., & Mostafa, A. M. (2020). Towards a Hybrid Immune Algorithm Based on Danger Theory for Database Security. *IEEE Access, 8,* 145332–145362. https://doi.org/10.1109/access.2020.3015399

Sameen, M., Han, K., & Hwang, S. O. (2020). PhishHaven—An Efficient Real-Time AI Phishing URLs Detection System. *IEEE Access, 8,* 83425–83443. https://doi.org/10.1109/access.2020.2991403

Seo, W., & Pak, W. (2021). Real-Time Network Intrusion Prevention System Based on Hybrid Machine Learning. *IEEE Access, 9,* 46386–46397. https://doi.org/10.1109/access.2021.3066620

Song, J., Kim, J., Choi, S., Kim, J., & Kim, I. (2021). Evaluations of AI-based malicious PowerShell detection with feature optimizations. *ETRI Journal, 43*(3), 549–560. https://doi.org/10.4218/etrij.2020-0215

Statista (2020). *Number of Data Breaches in the United States from 2013 to 2019, by industry.* https://www.statista.com/statistics/273572/number-of-data-breaches-in-the-united-states-by-business/

Stephanakis, I. M., Chochliouros, I. P., Sfakianakis, E., Shirazi, S. N., & Hutchison, D. (2019). Hybrid self-organizing feature map (SOM) for anomaly detection in cloud infrastructures using granular clustering based upon value-difference metrics. *Information Sciences, 494,* 247–277. https://doi.org/10.1016/j.ins.2019.03.069

Stevens, T. (2020). Knowledge in the grey zone: AI and cybersecurity. *Digital War, 1(1–3)*, 164–170. https://doi.org/10.1057/s42984-020-00007-w

Shandilya, S. K., Upadhyay, S., Kumar, A., & Nagar, A. K. (2022). AI-assisted Computer Network Operations testbed for Nature-Inspired Cyber Security based adaptive defense simulation and analysis. *Future Generation Computer Systems, 127*, 297–308. https://doi.org/10.1016/j.future.2021.09.018

Vavra, J., & Hromada, M. (2019). Evaluation of Data Preprocessing Techniques for Anomaly Detection Systems in Industrial Control System. *DAAAM Proceedings,* 0738–0745. https://doi.org/10.2507/30th.daaam.proceedings.101

Verwoerd, T., & Hunt, R. (2002). Intrution Detection techniques and approaches*. Computer communications, 25*(15), 1356-1365.

Vom Brocke, J., Hevner, A., & Maedche, A. (2020). Introduction to design science research. In *Design science research. Cases* (pp. 1-13). Springer, Cham.

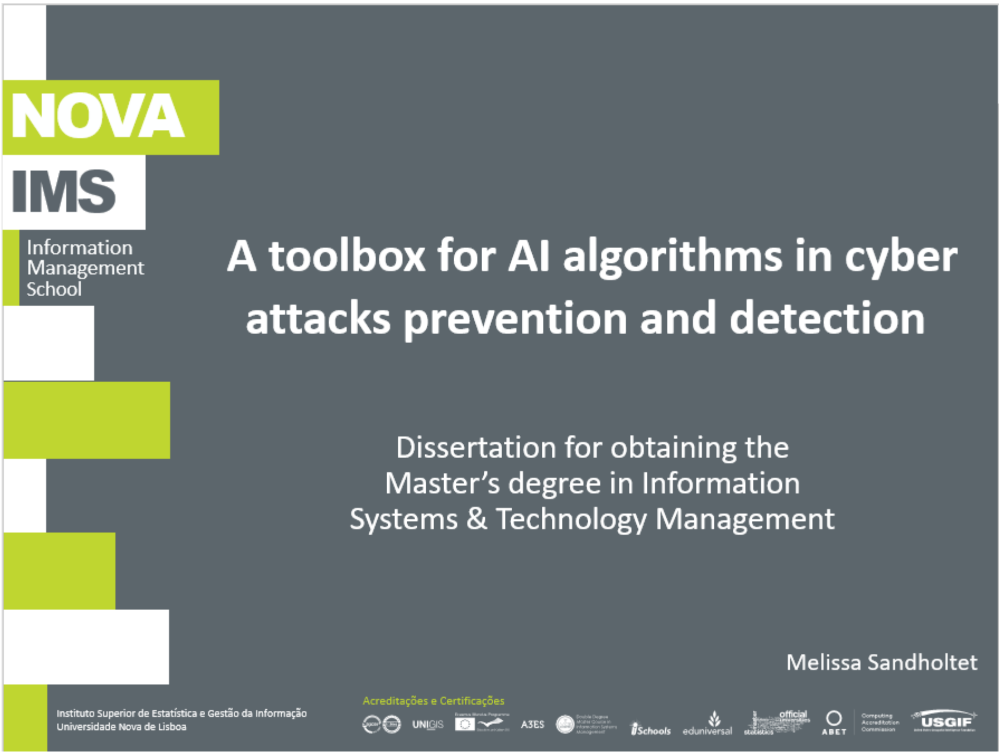
Wang, W., Du, X., Shan, D., Qin, R., & Wang, N. (2022). Cloud Intrusion Detection Method Based on Stacked Contractive Auto-Encoder and Support Vector Machine*. IEEE Transactions on Cloud Computing*. https://doi.org/10.1109/tcc.2020.3001017

What is a cyberattack? | IBM. (n.d.). https://www.ibm.com/topics/cyber-attack

Yar, M., & Steinmetz, K. (2019). Cybercrime and Society (3rd ed.). *SAGE Publications*.

APPENDIX

Experts Interview Guide



NOVA IMS Information Management School

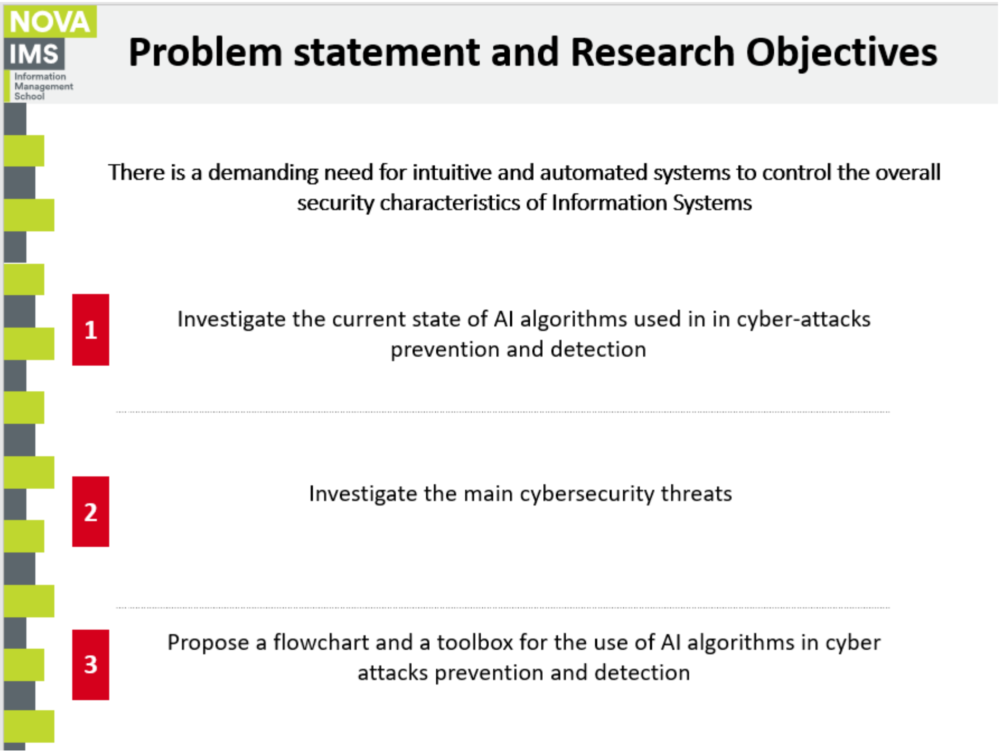
A toolbox for AI algorithms in cyber attacks prevention and detection

Dissertation for obtaining the
Master's degree in Information
Systems & Technology Management

Melissa Sandholtet

Instituto Superior de Estatística e Gestão da Informação
Universidade Nova de Lisboa

Acreditações e Certificações



NOVA IMS Information Management School

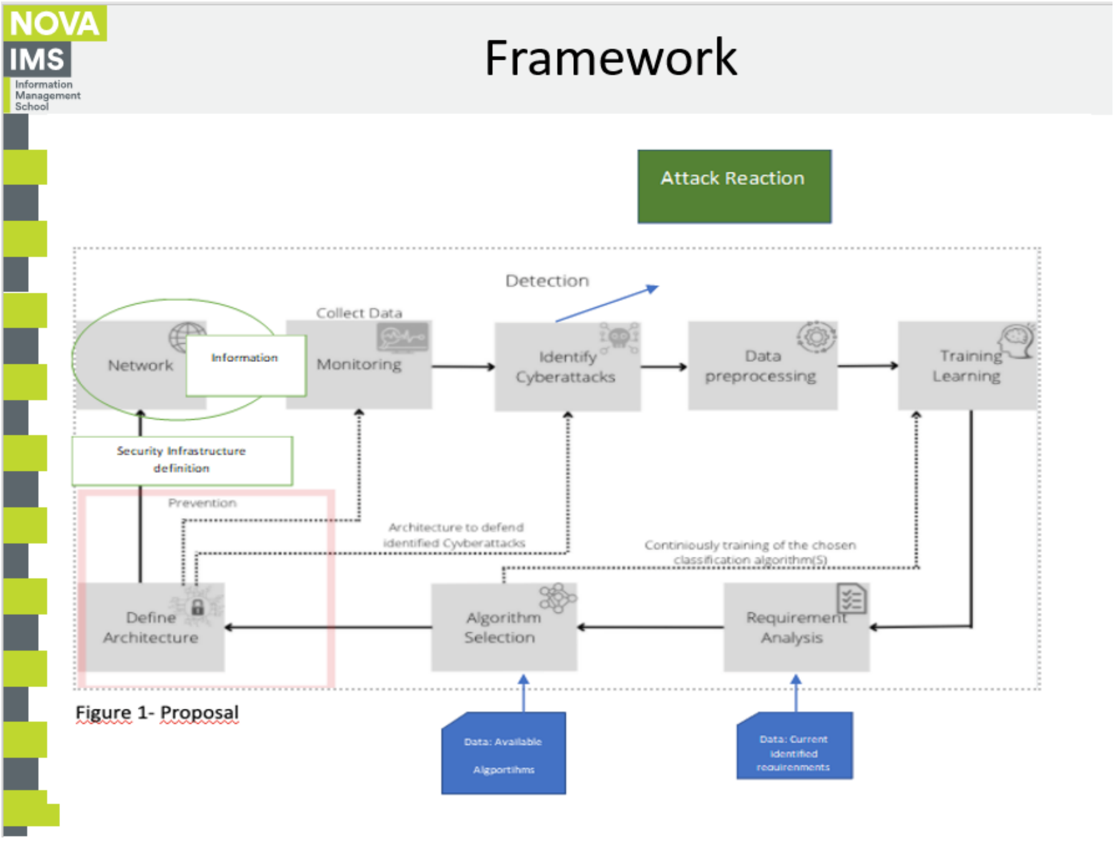
Problem statement and Research Objectives

There is a demanding need for intuitive and automated systems to control the overall security characteristics of Information Systems

1 Investigate the current state of AI algorithms used in in cyber-attacks prevention and detection

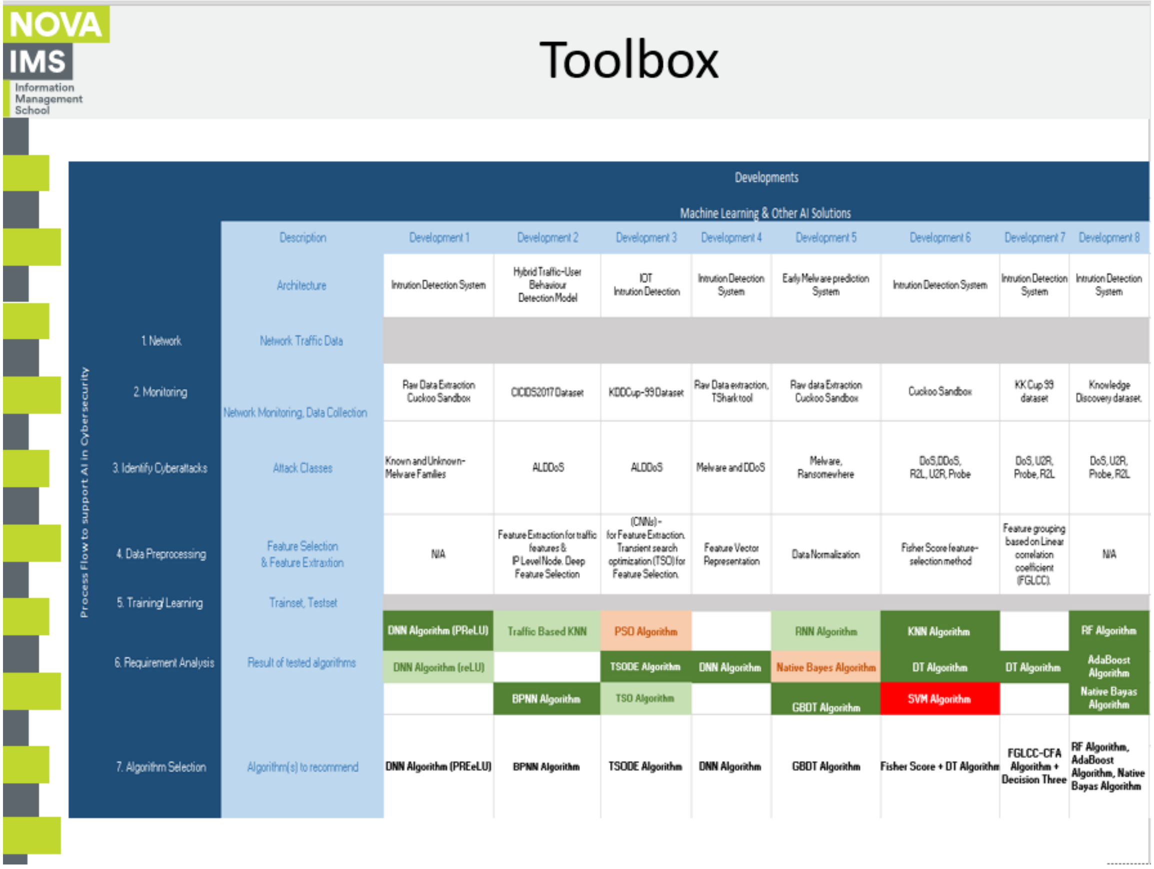
2 Investigate the main cybersecurity threats

3 Propose a flowchart and a toolbox for the use of AI algorithms in cyber attacks prevention and detection



NOVA
IMS
Information Management School
Framework
Attack Reaction
Detection
Collect Data
Network
Information
Monitoring
Identify Cyberattacks
Data preprocessing
Training Learning
Security Infrastructure definition
Prevention
Architecture to defend identified Cyberattacks
Continiously training of the chosen classification algorithm(S)
Define Architecture
Algorithm Selection
Requirement Analysis
Data: Available Algporithms
Data: Current identified requirements

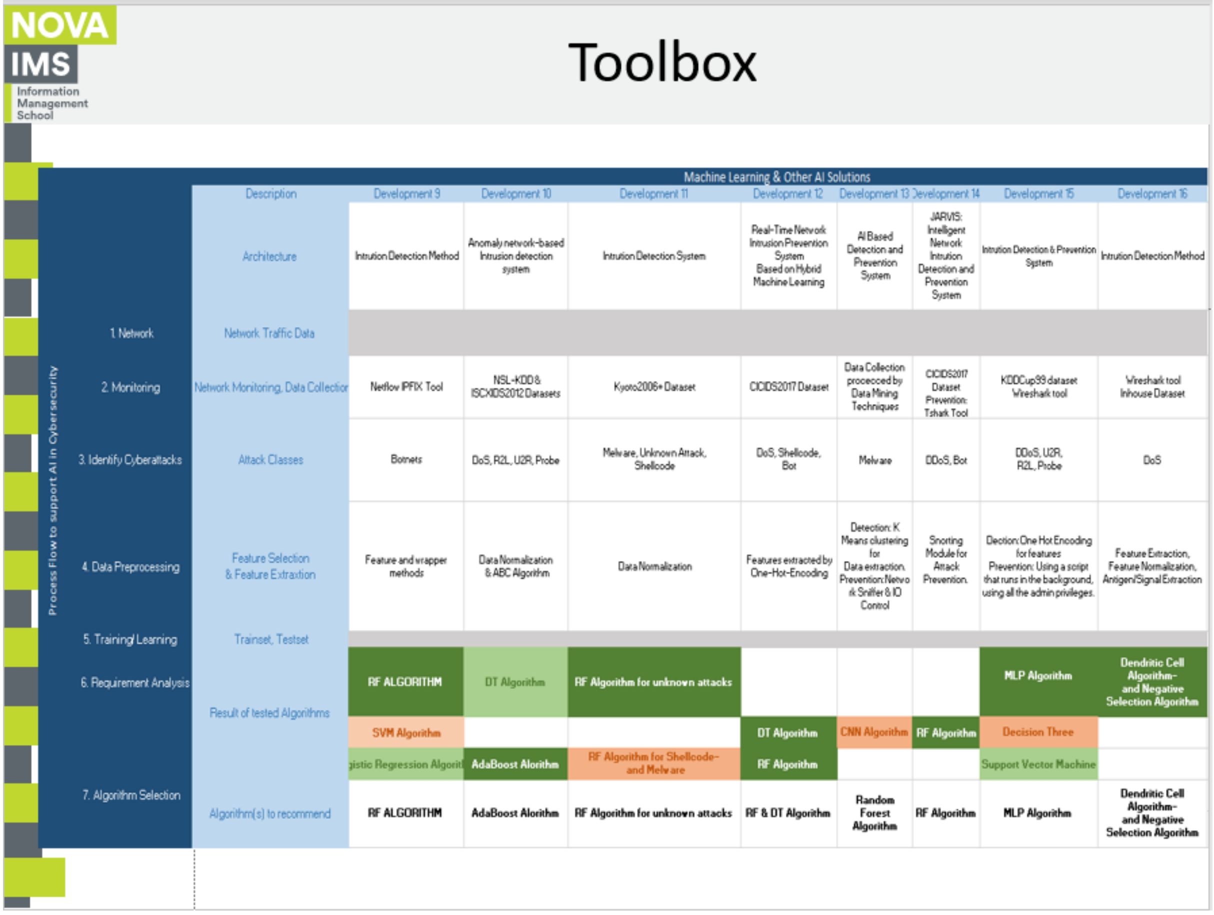
Figure 1- Proposal



NOVA
IMS
Information Management School

Toolbox

Process Flow to support AI in Cybersecurity		Developments Machine Learning & Other AI Solutions							
	Description	Development 1	Development 2	Development 3	Development 4	Development 5	Development 6	Development 7	Development 8
	Architecture	Intrution Detection System	Hybrid Traffic-User Behaviour Detection Model	IOT Intrution Detection	Intrution Detection System	Early Melware prediction System	Intrution Detection System	Intrution Detection System	Intrution Detection System
1. Network	Network Traffic Data								
2. Monitoring	Network Monitoring, Data Collection	Raw Data Extraction Cuckoo Sandbox	CICIDS2017 Dataset	KDDCup-99 Dataset	Raw Data extraction, TShark tool	Raw data Extraction Cuckoo Sandbox	Cuckoo Sandbox	KK Cup 99 dataset	Knowledge Discovery dataset.
3. Identify Cyberattacks	Attack Classes	Known and Unknown-Melvare Families	ALDDoS	ALDDoS	Melvare and DDoS	Melvare, Ransomewhere	DoS,DDoS, R2L, U2R, Probe	DoS, U2R, Probe, R2L	DoS, U2R, Probe, R2L
4. Data Preprocessing	Feature Selection & Feature Extraxtion	N/A	Feature Extraction for traffic features & IP Level Node. Deep Feature Selection	(CNNs) - for Feature Extraction. Transient search optimization (TSO) for Feature Selection.	Feature Vector Representation	Data Normalization	Fisher Score feature-selection method	Feature grouping based on Linear correlation coefficient (FGLCC).	N/A
5. Training/ Learning	Trainset, Testset								
6. Requirement Analysis	Result of tested algorithms	DNN Algorithm (PReLU) DNN Algorithm (reLU)	Traffic Based KNN BPNN Algorithm	PSO Algorithm TSODE Algorithm TSO Algorithm	DNN Algorithm	RNN Algorithm Native Bayes Algorithm GBDT Algorithm	KNN Algorithm DT Algorithm SVM Algorithm	DT Algorithm	RF Algorithm AdaBoost Algorithm Native Bayas Algorithm
7. Algorithm Selection	Algorithm(s) to recommend	DNN Algorithm (PREeLU)	BPNN Algorithm	TSODE Algorithm	DNN Algorithm	GBDT Algorithm	Fisher Score + DT Algorithm	FGLCC-CFA Algorithm + Decision Three	RF Algorithm, AdaBoost Algorithm, Native Bayas Algorithm

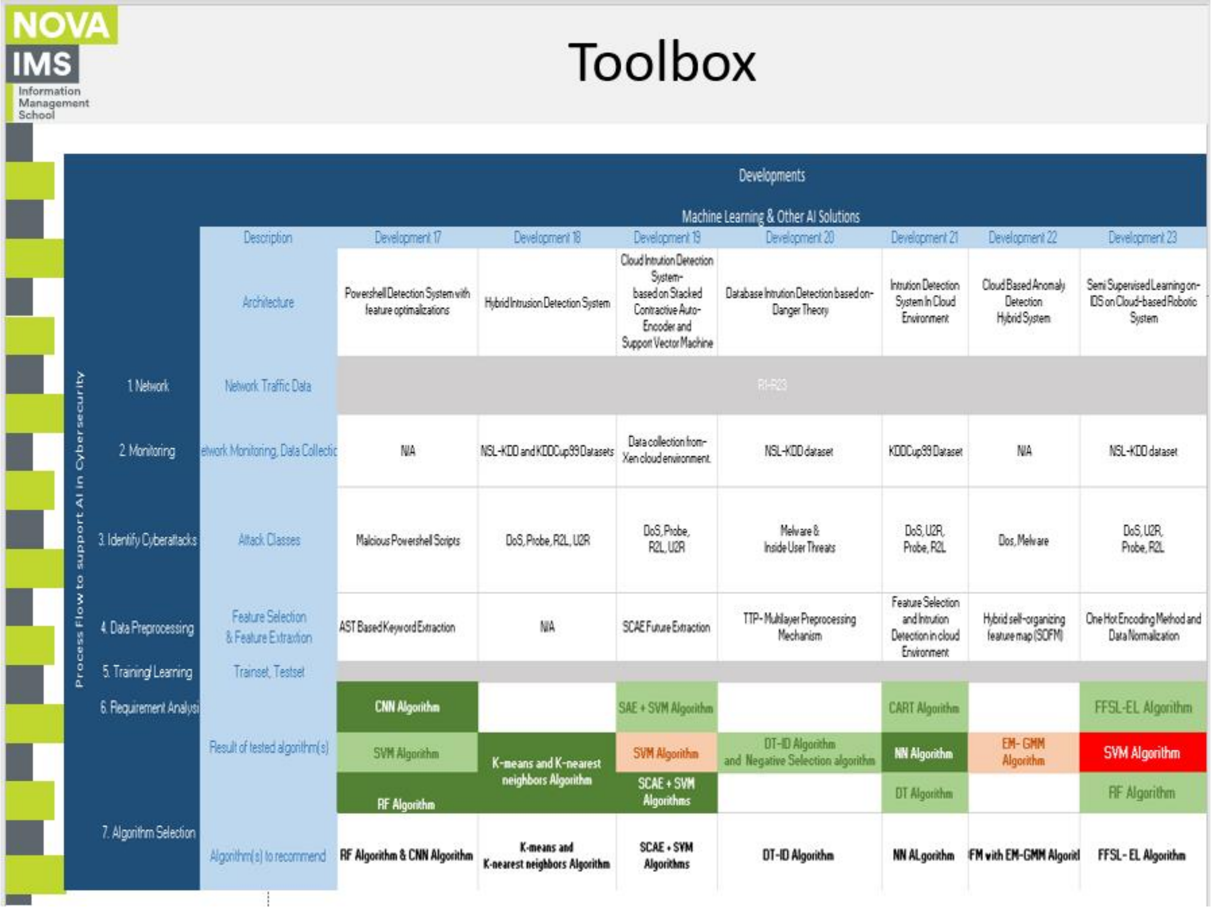


NOVA IMS
Information Management School

Toolbox

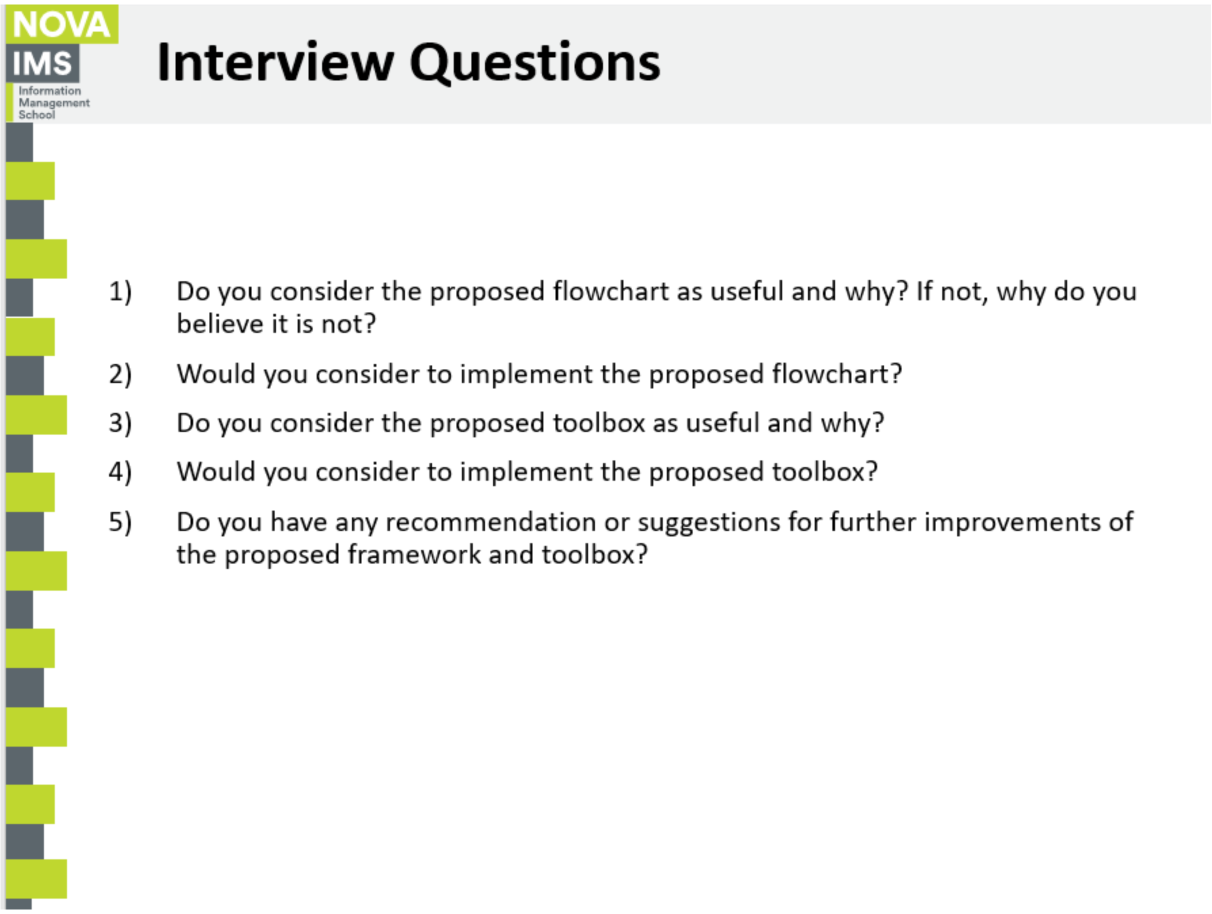
Process Flow to support AI in Cybersecurity

		Machine Learning & Other AI Solutions							
	Description	Development 9	Development 10	Development 11	Development 12	Development 13	Development 14	Development 15	Development 16
	Architecture	Intrution Detection Method	Anomaly network-based Intrusion detection system	Intrution Detection System	Real-Time Network Intrusion Prevention System Based on Hybrid Machine Learning	AI Based Detection and Prevention System	JARVIS: Intelligent Network Intrution Detection and Prevention System	Intrution Detection & Prevention System	Intrution Detection Method
1. Network	Network Traffic Data								
2. Monitoring	Network Monitoring, Data Collectior	Netflow IPFIX Tool	NSL-KDD & ISCXIDS2012 Datasets	Kyoto2006+ Dataset	CICIDS2017 Dataset	Data Collection proceced by Data Mining Techniques	CICIDS2017 Dataset Prevention: Tshark Tool	KDDCup99 dataset Wireshark tool	Wireshark tool Inhouse Dataset
3. Identify Cyberattacks	Attack Classes	Botnets	DoS, R2L, U2R, Probe	Melware, Unknown Attack, Shellcode	DoS, Shellcode, Bot	Melware	DDoS, Bot	DDoS, U2R, R2L, Probe	DoS
4. Data Preprocessing	Feature Selection & Feature Extraxtion	Feature and wrapper methods	Data Normalization & ABC Algorithm	Data Normalization	Features extracted by One-Hot-Encoding	Detection: K Means clustering for Data extraction. Prevention: Netwo rk Sniffer & IO Control	Snorting Module for Attack Prevention.	Dection: One Hot Encoding for features Prevention: Using a script that runs in the background, using all the admin privileges.	Feature Extraction, Feature Normalization, Antigen/Signal Extraction
5. Training/ Learning	Trainset, Testset								
6. Requirement Analysis	Result of tested Algorithms	RF ALGORITHM	DT Algorithm	RF Algorithm for unknown attacks				MLP Algorithm	Dendritic Cell Algorithm- and Negative Selection Algorithm
		SVM Algorithm			DT Algorithm	CNN Algorithm	RF Algorithm	Decision Three	
		gistic Regression Algorit	AdaBoost Alorithm	RF Algorithm for Shellcode- and Melware	RF Algorithm			Support Vector Machine	
7. Algorithm Selection	Algorithm(s) to recommend	RF ALGORITHM	AdaBoost Alorithm	RF Algorithm for unknown attacks	RF & DT Algorithm	Random Forest Algorithm	RF Algorithm	MLP Algorithm	Dendritic Cell Algorithm- and Negative Selection Algorithm



Toolbox

Process Flow to support AI in Cybersecurity		Developments Machine Learning & Other AI Solutions						
	Description	Development 17	Development 18	Development 19	Development 20	Development 21	Development 22	Development 23
	Architecture	Powershell Detection System with feature optimalizations	Hybrid Intrusion Detection System	Cloud Intrution Detection System- based on Stacked Contractive Auto-Encoder and Support Vector Machine	Database Intrution Detection based on-Danger Theory	Intrution Detection System In Cloud Environment	Cloud Based Anomaly Detection Hybrid System	Semi Supervised Learning on-IDS on Cloud-based Robotic System
1. Network	Network Traffic Data	R1-R23						
2. Monitoring	etwork Monitoring, Data Collectic	N/A	NSL-KDD and KDDCup99 Datasets	Data collection from-Xen cloud environment.	NSL-KDD dataset	KDDCup99 Dataset	N/A	NSL-KDD dataset
3. Identify Cyberattacks	Attack Classes	Malcious Powershell Scripts	DoS, Probe, R2L, U2R	DoS, Probe, R2L, U2R	Melware & Inside User Threats	DoS, U2R, Probe, R2L	Dos, Melware	DoS, U2R, Probe, R2L
4. Data Preprocessing	Feature Selection & Feature Extraction	AST Based Keyword Extraction	N/A	SCAE Future Extraction	TTP- Multilayer Preprocessing Mechanism	Feature Selection and Intrution Detection in cloud Environment	Hybrid self-organizing feature map (SOFM)	One Hot Encoding Method and Data Normalization
5. Training/ Learning	Trainset, Testset							
6. Requirement Analysi	Result of tested algorithm(s)	CNN Algorithm		SAE + SVM Algorithm		CART Algorithm		FFSL-EL Algorithm
		SVM Algorithm	K-means and K-nearest neighbors Algorithm	SVM Algorithm	DT-ID Algorithm and Negative Selection algorithm	NN Algorithm	EM- GMM Algorithm	SVM Algorithm
		RF Algorithm		SCAE + SVM Algorithms		DT Algorithm		RF Algorithm
7. Algorithm Selection	Algorithm(s) to recommend	RF Algorithm & CNN Algorithm	K-means and K-nearest neighbors Algorithm	SCAE + SVM Algorithms	DT-ID Algorithm	NN ALgorithm	FM with EM-GMM Algoritl	FFSL - EL Algorithm



Interview Questions

1) Do you consider the proposed flowchart as useful and why? If not, why do you believe it is not?

2) Would you consider to implement the proposed flowchart?

3) Do you consider the proposed toolbox as useful and why?

4) Would you consider to implement the proposed toolbox?

5) Do you have any recommendation or suggestions for further improvements of the proposed framework and toolbox?

Thank you for your time and expertise!

Address: Campus de Campolide, 1070-312 Lisboa, Portugal
Phone: +351 213 828 610 Fax: +351 213 828 611

Acreditações e Certificações

UNIGIS
A3ES
iSchools
eduniversal
official statistics
ABET
Computing Accreditation Commission
USGIF

Instituto Superior de Estatística e Gestão da Informação
Universidade Nova de Lisboa