# INTELLIGENT AI BASED EDR

by

Sonia Azeem

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Master of Science

(Cyber Security)

at the

NATIONAL UNIVERSITY OF SCIENCES & TECHNOLOGY, PAKISTAN

May 2023

# Intelligent AI based EDR

By

**Sonia Azeem**

00000326426

Supervisor

**Prof. Dr. Arshad Aziz**

Department of Cyber Security

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science

In

Pakistan Navy Engineering College, National University of Sciences and Technology Karachi, Pakistan.

May 2023

# Copyright Notice

To my affectionate mother, supportive father and my siblings. I dedicate my this thesis to all of you for having faith in me.

# Certificate of Originality

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, nor material which to are substantial extent has been accepted for the award of any degree or diploma at Department of Electrical Engineering at Pakistan Navy Engineering College (PNEC) or at any other educational institute, accept where due acknowledgement has he made in the thesis. Any contribution made to the research by others, with whom I have worked at Pakistan Navy Engineering College (PNEC) or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics which has been acknowledged.

Author Name: **Sonia Azeem**

Signature:

FORM TH-4

## National University of Sciences and Technology

## MASTER'S THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by: (Student Name & Regn No.) Sonia Azeem (00000326426) Titled: Intelligent AI Based EDR be accepted in partial fulfillment of the requirements for the award of Master's degree.

## EXAMINATION COMMITTEE MEMBERS

1. Name: Prof. Dr. Fawad Ahmed  Signature: ________________

2. Name: Dr. Bilal. M. Khan  Signature: ________________

3. Name: ________________  Signature: ________________

Supervisor's name: Prof. Dr Arshad Aziz  Signature: ________________

Date: 4-4-2023

QURRATULA
Lt Cdr Pakista
HOD CvSi

Head of Department

22-6-2023
Date

## COUNTERSIGNED

Date: 22-6-2023

M IRFAN NADEEM
Captain Pakistan Navy
Deputy Commandant
PNS JAUHAR

Dean / Principal

# THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS thesis written by SONIA AZEEM Regn No. 00000326426 of NUST- PNEC (College) has been vetted by undersigned, found complete in all respects as per NUST Status/Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have been incorporated in the said thesis.

Signature: ____________________

Name of Supervisor Prof. Dr. Arshad Aziz

Dated: 22-6-2023

QURRATULAIN TI(M)
Lt Cdr Pakistan Navy
HOD CySD

Signature: HoD ____________________

Dated: 22-6-2023

Signature: (Dean/Principal): ____________________

Dated: 22-6-2023

M IRFAN NADEEM
Captain Pakistan Navy
Deputy Commandant
PNS JAUHAR

# Approval

It is certified that the contents and form of the thesis entitled **"Intelligent AI based EDR"** submitted by **Sonia Azeem** have been found satisfactory for the requirement of the degree.

Advisor: **Prof. Dr. Arshad Aziz**

Signature: __________

Date: 22-6-2023

Committee Member 1: **Prof. Dr. Fawad Ahmed**

Signature: __________

Date: 22-6-2023

Committee Member 2: **Dr. Bilal M Khan**

Signature: __________

Date: 22-6-2023

# Acknowledgments

First and foremost, I would like to express my deepest gratitude to Almighty Allah Ta’ala (Jalla Jalalahu) for his compassion and blessings that he has continuously bestowed upon me.

I would like to express my sincere gratitude to my supervisor, Dr. Arshad Aziz, for his compassionate and knowledgeable supervision. His encouragement, motivation, and assistance throughout the project made this research possible. I want to thank the members of the Guidance Committee, particularly Prof. Dr. Fawad Ahmed and Dr. Bilal M. Khan for their instructions and directions. Following, it was my loved ones’ prayers that actually made it feasible to finish this research project. I would like to extend my heartfelt gratitude to my parents and siblings for their unwavering support during my research project. Their sincere assistance and encouragement throughout the entire process are deeply appreciated. Similarly, I am sincerely thankful to my close friend, Sidra Gul, for her continuous support during my journey, especially during difficult times. Her constant presence in my life, kindness, and motivation have been of immense value to me.

# Table of Contents

Page

Chapter Page

# List of Tables

Table Page

# List of Figures

Figure Page

# Abstract

Cyberspace is now facing more aggressive, sophisticated, and occasionally real-time attacks from threats. These have pushed both researchers and practitioners to safeguard cyberspace at its most fundamental level, known as the endpoint. In most cyber-attacks, attackers target endpoints having security breaches or vulnerabilities to gain access to the internal network and carry out various malicious activities. Furthermore, phishing emails have emerged as the most common method for hackers and attackers to propagate malware. To effectively address these challenges, this study presents the development of an Intelligent AI-based EDR solution that enhances endpoint security by increasing detection, investigation, and response capabilities. The system integrated with VirusTotal and Yara for real-time malware detection and response, providing an effective solution to address the identified issue. Additionally, a Spam Filtering solution is utilized as an integral part of the EDR system to prevent spam infiltration in email inboxes, thereby halting malware propagation through malicious downloads or phishing links, the developed Spam Filtering solution achieves an impressive accuracy rate of 98.50%. The EDR solution developed in this study is compared to well-known commercial EDR vendors, demonstrating satisfactory results. This comprehensive approach effectively tackles the issue of cyber-attacks targeting endpoints and the use of phishing emails for malware propagation, with findings suggesting the developed intelligent EDR system offers enhanced visibility at endpoints and effective countermeasures against malware threats.

# Chapter 1

# Introduction

According to Cisco Annual Internet report by 2023, the number of gadgets connected to IP networks will reach more than three times the world's population or 3.6 network devices per person [1]. In the early stages of the IT industry, cyber attacks were largely directed at networks as the primary purpose of attackers and hackers was to obtain an unauthorized access to corporate networks with technology advancement network security has become an integral component in all IT organisations but now as Bring your own device (BYOD) and the Internet of Things (IoT) have recently gained popularity, which has resulted in a sharp rise in the number of individual devices linked to a company's network, security risks have new attack vectors to use when connecting laptops, tablets, cell phones, and other wireless devices to business networks. When sending or receiving messages over the network, nobody wants to be bothered or eavesdropped on. Endpoint security has thus grown in popularity among researchers working in the field of security [2].

## Four Pillars of Digital Security

Information security, cyber security, network security, and endpoint security are all interconnected concepts that play a critical role in protecting digital assets as can be seen from the Fig 1.1 Information security involves safeguarding information and information systems against unauthorized access, use, disclosure, disruption, modification, or destruction. The primary aim of information security is to maintain the confidentiality, integrity, and availability of information, thus protecting it from various threats such as cyber attacks, unauthorized disclosure, or accidental loss [3].

Cyber Security refers to a set of practices, processes, and technologies that aim to protect computer systems, networks, and digital information from unauthorized access, use, disclosure, disruption, modification, or destruction [4]. Network security focuses on securing the network infrastructure, including hardware, software, and services, to prevent unauthorized access to the network. Endpoint security focuses on securing individual devices, such as laptops, desktops, smartphones, and tablets, to prevent threats that may affect these devices. These four concepts work together to form a comprehensive approach to protecting digital assets and information from cyber threats. Effective information security measures require ongoing monitoring, testing, and updating to stay ahead of evolving threats and vulnerabilities.

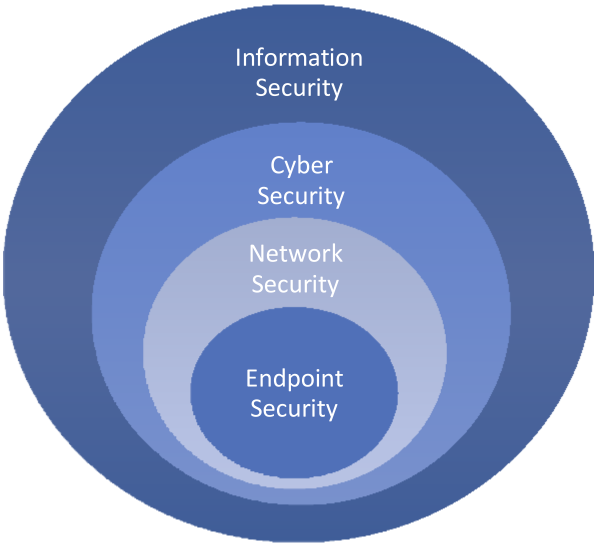


Figure 1.1: The Four Pillars of Digital Security: Information, Cyber, Network, and Endpoint

As the number of endpoints such as laptops, desktops, mobile devices, and servers in networks continue to grow, the need for effective endpoint security measures has become increasingly crucial. Endpoint devices are vulnerable to cyber-attacks, which can result in significant damage to organizations in terms of both finances and reputation. Endpoint security is essential because it offers protection not only to the endpoints themselves but also to the data and applications stored on them. With remote workforces and BYOD policies becoming increasingly prevalent, the demand for endpoint security solutions has become even more critical.

Endpoint detection and response also known as Endpoint threat detection and response is an advanced endpoint security solution with its specific features like continuous real time monitoring, threat detection and real time response since a major element of EDR is threat intelligence can provide anomaly detection and alerts, as well as remediation of a compromised internal network. Additionally, they can use machine learning to anticipate and evade the threat [5] The EDR solution's continuous, centralized record and captured all of this behavior in real time and imposed prevention so when suspicious turned malicious and escalated the alarm to authorities accordingly. This enabled them to initiate corrective action right away[2] EDR systems, as opposed to endpoint protection platforms, are made to offer incident response capabilities that let security teams find malicious objects in a 24 network, spot threats that other security technologies miss, and conduct root cause investigation [6]

Therefore, the creation of an Endpoint Detection and Response (EDR) solution that emphasizes malware detection and spam email filtering is a vital component in safeguarding endpoints against cyber threats. So, major objective of this study is to propose such an EDR solution with open-source tools that supplement endpoint security with increased detection, investigation, and response capabilities we will be able to detect malicious code, such as malware, viruses, or ransomware on endpoints, with proper remediation instructions to prevent further damage in known cases. We present an enhanced EDR solution that seamlessly integrates VirusTotal and YARA to detect and mitigate malware effectively. Our solution is an open-based platform built on open-source technologies, deployable at no cost and this characteristic makes it advantageous for resource-constrained small to medium-sized organizations, providing them with robust malware detection capabilities. Through this work, we will be able to stop malware spreading either through malicious downloading or via phishing link. Additionally, we have developed an AI-based model for spam and ham email detection, expanding our malware detection capabilities to cover email channels

## Chapter Structure

This work is divided into six chapters briefly explained below:

***Chapter 1*** begins with an introduction to the growth of cyber-attacks and provides a brief overview of the thesis's structure that how the topics will be covered in subsequent chapters.

***Chapter 2*** focuses on the significance of securing endpoints within a network, including an analysis of endpoint threats and an exploration of endpoint security technologies EDR, XDR and SIEM

***Chapter 3*** The chapter highlights the role of deep learning, machine learning, and AI in cyber-security. It provides an overview of machine learning types and explores the structure, function, and types of artificial neural networks (ANNs).

***Chapter 4*** provides an in depth review of existing literature and work that contribute to this field of security was conducted.

***Chapter 5*** provided detailed discussion on the open source technology and tools used in the project is provided, resources required in its deployment ,

***Chapter 6*** summarizes the research and puts forth suggestions for future studies.

# Chapter 2

# Endpoint Security, Threats and Technologies

## Endpoints and Endpoint security

As per Statist research, the volume of IoT-connected devices is predicted to reach 75 billion in 2025. The more well recognized endpoints include workstations, laptops, Servers, Mobile phones and Virtual machines etc. Endpoint Security is a subcategory of cyber security that protects endpoints or access points against being used as resource by bad actors or cyber attacks [7].

Cyber criminals infect endpoints with malware in order to get access to a nation's or corporation's internal systems. Malware is evolving into a more sophisticated form, such as advanced persistent threat (APT) attacks, and attacks utilizing zero day malware are also on the rise. [8]. Due to the constant evolution of security technologies, traditional endpoint protection is no longer sufficient to safeguard their IT system. [9]

## Endpoint attacks

While endpoints can improve company connection and productivity, they also pose a number of dangers and drawbacks. The following is a list of typical endpoint threats that are applicable to all businesses:

## Phishing Attacks

Phishing is a kind of social engineering assault used to steal users' data, including login details and confidential information such as credit card numbers. In an email Phishing, a user is duped into clicking a link that leads to the installation of malware. Email phishing attempts are commonly camouflaged in mass-distributed messages, such as organizational or educational email messages [7].

### Ransomware

Ransomware is a type of malware that encrypts files on a device, rendering the files and system inoperable. Ransomware is typically a strategy used to force individuals to pay a sum of money. Malicious hackers will intimidate victims by publishing payment rejection data and explaining how they will publicly disparage them as extortion [7].

### Vulnerability Exploits (Zero Day Attack)

A zero-day vulnerability is a previously unknown security flaw or software malfunction that an attacker can exploit. Zero-day exploits often target government organizations, major corporations, and individuals with access to sensitive corporate data, as well as those who use a vulnerable system, hardware devices, firmware, or IoT [7]

### Waterholes

An intentional attack that compromises a website is called a waterhole attack which is used by attackers to acquire financial benefit and to build their botnet (a network of stolen machines and computers with malware infection and controlled by a hackers). Waterhole attacks happen when an attacker discovers a weakness on a social platform, compromises it, and infects it with malware [7]

### Drive-by download

A drive-by download is a harmful code or program that is unintentionally downloaded into a computer or device. This assault is forced since the user is unaware that they have downloaded harmful program or software. Drive-by downloads are carried out by exploiting insecure programs, software products, or web browsers [7].

## Malware detection at endpoints

The term ”malicious software,” sometimes known as malware, refers to any application that engages in harmful activity. Malware can be classified into a wide range of forms, such as viruses, Trojan horses, worms, spyware, botnet malware, ransomware, and so forth. Malware is a common component of online cyber attacks, including nation-state cyber war, criminal activity, fraud, and scams. Malware has been transformed by cyber criminals into a successful business strategy. Today, malware poses a constant threat to people, businesses, and government infrastructures more than ever before [10].

Malware detection at the endpoint is critical for cyber Security since endpoints are frequently the first line of defense against malware attacks. Techniques for detecting malware can be based on behavioral analysis, signature-based detection, or machine learning. Organizations may recognize dangers and take proactive action to address them by putting these approaches into practice, which will stop malware from doing any damage.

## Top ways of Malware Spreading

Hackers employ a wide range of attack vectors to compromise endpoints, and they are constantly developing new ones. Many different reasons can lead to malware spreading few are discussed below:

## Malicious Downloads

Attackers frequently launch drive-by downloads by exploiting known flaws in legitimate websites’ software. This could be something as simple as adware, or it could be as bad as ransomware or a virus because consumers frequently do not check to see whether a file is secure before downloading it or simply do not understand what red flags to look for, such method of infection is quite popular among cyber attackers [11].

## Remote Desktop Protocol

RDP shor for remote desktop protocol is a network connection protocol that allows a user to connect towards another computer via a network connection. Automation is now used by attackers to scan the internet for PCs that are open to RDP. They will next attempt to guess a username and password in order to gain access to the distant machine [11] .

## Social Network Spam

Spam on social networks is a relatively new method of attack for hackers where users explore social media sites to view images or keep in touch with old friends and they might be unaware that the image they are about to click on might contain malware. The user is then directed to a false YouTube page where they are asked to install and download a video player plugin [11].

## Phishing Emails

One of the most popular types of cyber crime is phishing. This is mostly because practically anyone could be contacted via text, email, or direct messaging. Hackers have gotten extremely competent at designing emails that deceive employees to clicking on links or email attachments containing malicious code. Phishing emails have numerous purposes, and these goals frequently define how a phishing attack will operate. In certain circumstances, the purpose is credential theft, therefore a phishing email may appear as an email from a reputable organization and direct the victim to a fraudulent login page. In all circumstances, phishing emails utilize a combination of psychology and cunning to persuade the receiver to perform an action that the attacker desires. Some popular pretexts include problems with online accounts, parcel delivery failures, unpaid invoices, and others [11].

## Endpoint Detection and Response

The term ETDR ”Endpoint Threat Detection and Response” was originally used in July 2013 by Gartner’s Anton Chuvakin. In 2015 term ETDR updated into EDR ”Endpoint Detection and Response” term [6]

The growth in network-connected endpoints is one of the reasons of EDR adoption. Another key driver is the increased sophistication of attacks, which target network endpoints as simpler targets to enter into a network [12]. Therefore, by adopting an intelligent contextual solution, one acquires the capacity to connect the system’s various operations automatically or using artificial intelligence, making it possible to associate processes and their history, assisting in the step-by-step identification of malicious system behavior. It also assists in comprehending how detection and mitigation logic may be implemented at the contextualization level, as well as changing the quantity of events that the security team will get on the condition of their environment [12].

## EDR Features

The following section will outline and describe the essential qualities or attributes of EDR. Fig. 2.1

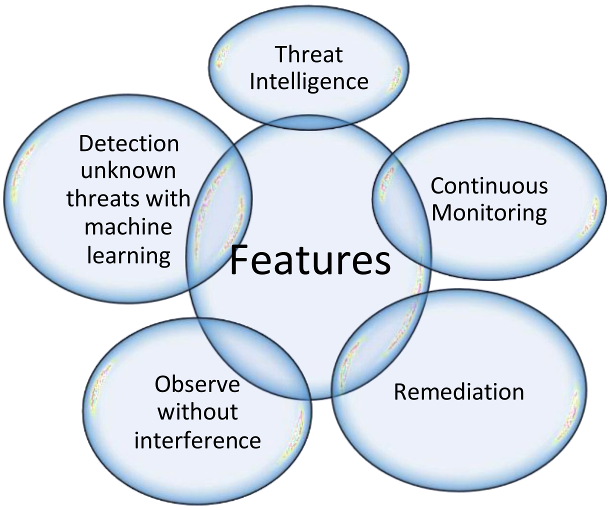

Threat Intelligence
Detection unknown threats with machine learning
Continuous Monitoring
Features
Observe without interference
Remediation

Figure 2.1: EDR features

### Threat Intelligence

Cyber Threat Intelligence (CTI), commonly referred to as threat intelligence, is the collection, analysis, and reorganization of information regarding cyber attacks that could threaten an organization. EDR is a software that offers more than just protection through its threat intelligence feature, it can alert corporations to potential dangers and threats. It may offer some threat-related information that has been gathered by the EDR server. By examining information and data regarding insider risks that have already occurred rather than projecting the latent risk, this type of intelligence will make it easier to eliminate the insider danger [5].

### Continuous Monitoring

Strengthening endpoint control is the best way to identify anomalous endpoint behavior whenever one endpoint get infected, EDR will immediately recognize that unusual behavior of that endpoint and segregate it from the network. They can manage endpoints dynamically, which implies they must continuously and automatically test endpoints. Furthermore, it also provides CPU protection to safeguard the server kernel. Visibility into actions taking place at lower system levels, such as the CPU, will be a feature of more advanced behavior-based protection. Visibility into the CPU level is useful for stopping malware, including several exploits, that tries to modify and implement modifications in memory [5].

### Remediation

The Malware's spread is stopped once aberrant endpoints have been cleaned up so user can believe that the entire system is pure and secure. However, the sophisticated virus can behave like an ink drop in clear water and can quickly spread to the opposite area and contaminate the internal network. All of the internal networks can be scanned by EDR to ensure the virus's remnants are absent [5].

## Observe without interference

So, nobody wants to load the endpoints with hefty and expensive client software any longer Therefore, that was one of the greatest downsides of anti viruses.” According to security researchers, antivirus software must be loaded on the endpoints which consumes a significant amount of endpoint resources. EDR, on the other hand,Endpoint Detection and Response (EDR) solutions are designed to consume fewer system resources compared to traditional antivirus software [5].

## Detecting unknown threats with machine learning

In the domain of computer programming, machine learning is an aspect of artificial intelligence that typically use statistical methods to enable computers to ”learn” from the data without even being explicitly programmed. The ability to recognize threats that they had never experienced before can be improved via machine learning. This capability may turn out to be one of the most important benefits of EDR in situations where threats are changeable, that’s why the majority of businesses currently prefer to employ it [5].

## XDR: The Future of EDR

In recent years, EDR has become a prevalent cyber Security technology used in many defense systems. It is employed by organizations to detect and prevent activity when an attacker gains access to an endpoint. However, this approach fails to consider the larger picture of the entire attack chain, which extends well beyond the endpoints. Extended detection and response. In addition to what occurs at the endpoint, the organization may record the entire kill chain. Comprehensive knowledge of the attack’s phases enables businesses to either stop attackers automatically or carry out human investigations and responses at each stage [10]. (XDR) is an extension of EDR, primary distinctive feature is that it gathers and connects information from the entire infrastructure, including email, endpoints, servers, cloud workloads, and networks. This allows for a comprehensive view of the environment and advanced threat context. As a result, threats can be analyzed, prioritized, pursued, and addressed to prevent data loss and security breaches [12].

### What distinguishes XDR from SIEM?

XDR is seen as a successor to SIEM technology, which is mainly focused on collecting and analyzing large amounts of log events and data. SIEM is primarily a research tool, which means that security teams need to perform extensive analysis to reach conclusions. XDR, on the other hand, prioritizes threat detection and response. Its distinguishing feature is that it automatically responds to threats, or if that is not possible, it accelerates the investigation and analysis process to improve response times.

XDR provides visibility into the entire attack life cycle and correlates it with the overall environment, from infiltration to lateral movement, cleanup, or mitigation. By contrast, implementing EDR, SIEM, and other solutions independently does not provide the strategic context and correlation necessary to assess the current threat environment effectively. XDR technology has the potential to address this gap and offer a more comprehensive and integrated solution [12].

# Chapter 3

# Major Role of Artificial Intelligence, Machine Learning, and Deep Learning in Cyber Security

## Artificial Intelligence Family :AI, ML, DL

Artificial Intelligence (AI) is a more comprehensive field that comprises various subsets, such as Natural Language Processing (NLP), Computer Vision, Robotics, and others. Its primary objective is to develop intelligent machines that can replicate human thought processes and decision-making capabilities. Where as Machine Learning (ML) is a subcategory of Artificial Intelligence (AI) aims to create algorithms and models that can learn from data and make decisions or predictions without explicit programming.[13]. Despite being a subset of AI, Machine Learning plays a critical role in powering numerous AI applications by facilitating the learning and decision-making capabilities of machines. Fig. 3.1

Meanwhile,Deep Learning (DL) is a category of Machine Learning (ML) that takes cues from the information processing patterns observed in the human brain. DL does not rely on pre-programmed rules to function, instead, it employs copious amounts of data to map input to output. DL is crafted through multiple layers of algorithms known as Artificial Neural Networks (ANNs), which offer unique interpretations of the input data presented to them[14].

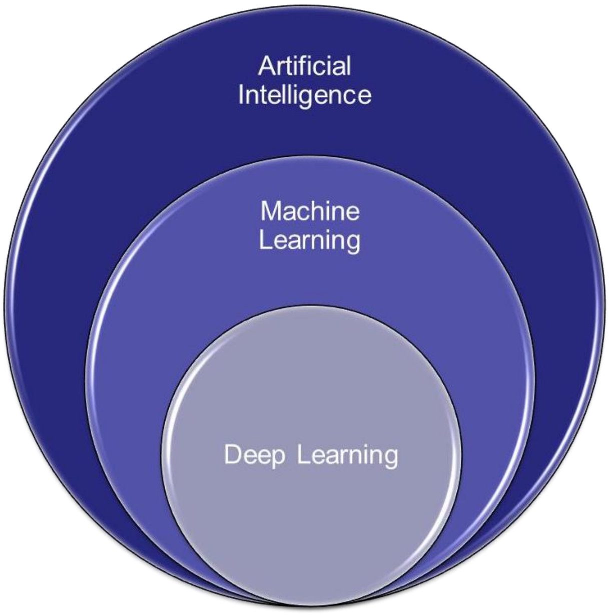


Figure 3.1: AI Family

## AI, ML and DL in Cyber Security

Artificial intelligence (AI), particular machine learning (ML) and deep learning (DL) have advanced at a rapid rate in recent years and are now poised to start having a significant impact on many facets of society and occupations, cyber security is not an exception. This growth has been attributed to improvements in computing power and algorithm development. AI-based systems, also known as cognitive systems, are assisting us in automating numerous tasks and preparing for challenges that are more complicated than most individuals are capable of tackling.

Traditional cyber security techniques may have difficulty detecting new generations of malware and cyber attacks. When they grow over time, more forceful methods are required. ML-based security solutions employ information from previous assaults to respond to more recent ones. AI systems in cyber security will also significantly increase IT staff members' productivity, which is another important benefit [15]

## AI importance in EDR

As the number of endpoints grows, so does the volume of data that must be safeguarded. Without sophisticated security automation that can intelligently anticipate and prevent assaults, humans just cannot function. Each data source has unique traits, patterns, and ways of being interpreted. Because of this, it may be challenging for security analysts to find malicious threats when they are only looking at the raw data. IT professionals' top goal is automating endpoint identification and reaction solutions. The demand for human automation is growing, particularly in recognizing potential false positives and false negatives that analysis may yield. Machine learning techniques have improved over time, making it possible to find interesting events that appear to be outside of patterns and to better identify dangers across a variety of domains. [9].

Machine learning algorithms can enhance Endpoint Detection and Response (EDR) solutions to swiftly and precisely detect and counter threats, minimizing the number of false alarms and enhancing overall security measures. Moreover, machine learning can enable EDR solutions to become better equipped to handle novel and evolving threats by drawing insights from previous events to enhance future response capabilities [9].

Every hour, 3 million new malware samples are discovered, and attackers have become adept at utilizing file-less, ransomware, and crypto currency malware. Additionally, a new type of intelligent malware has emerged that can evade endpoint detection systems and use AI to modify its signature, regulate its behavior, create lures, self-propagate, and deliver other malware strategically, while minimizing its impact. However, traditional malware detection systems are struggling to keep pace with these evolving threats, and there is a pressing need for a significant advancement in Cyber Security measures. Incorporating AI-based technologies into EDR systems is necessary to improve their capability in detecting and responding to the increasingly complex and advanced forms of malware. By analyzing large volumes of data, AI can identify patterns that indicate malicious behavior, resulting in faster and more accurate detection of malware. This integration of AI into EDR systems can significantly enhance Cyber Security measures against the constantly changing and expanding threat of malware.[16]

## Major Categories

Machine learning may generally be broken down into majorly three areas. supervised learning, unsupervised learning and reinforcement learning are among the categories. Below Fig. 3.2 depicts more information.

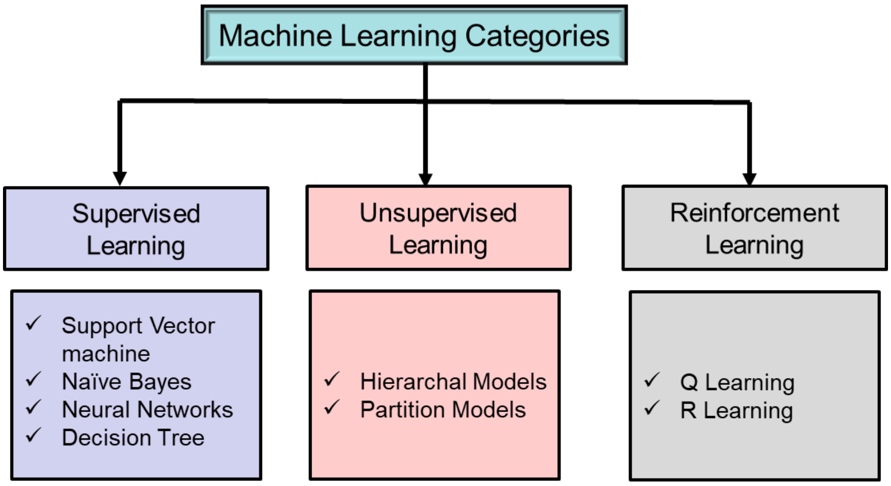


Figure 3.2: Machine Learning categories

## Supervised Learning

In supervised learning, input data is coupled with corresponding output labels, and a model is trained on this labelled data. For new input data, the model learns to anticipate the output label [9]. The purpose of supervised learning is to predict the value of one or more output variables based on the value of an input variable vector x. The output variable might be either a discrete or continuous variable (regression problem) (classification problem).

## Unsupervised Learning

The unsupervised learning training database comprises only of a collection of input vectors x. Several tasks can be addressed through unsupervised learning. Clustering or cluster analysis are the two most popular techniques for unsupervised learning [9]

## Reinforcement learning

The reinforcement learning (RL) paradigm enables agents to learn by experimenting with different behaviors and modifying their behavior using just evaluative feedback, known as the reward. The agent wants to perform at its best over the long term. As a result, the agent assesses the effects of its actions on the future in addition to the current reward. The two key characteristics of RL are delayed rewards and trial-and-error learning [9]

### Artifical Neural Networks

Artificial Neural Networks are collections of basic processing units that are connected to one another in order to exchange information through a significant number of weighted connections. Each unit receives input from its neighbors as well as outside sources, calculates the results, and then transmits the results to its neighbors. Also provided is the means for adjusting the weights of the connections. Any classification-based machine learning problem can be solved using neural networks. [17]. Every artificial neural network's fundamental building element is an artificial neuron, which is a straightforward statistical model. Fig. 3.3

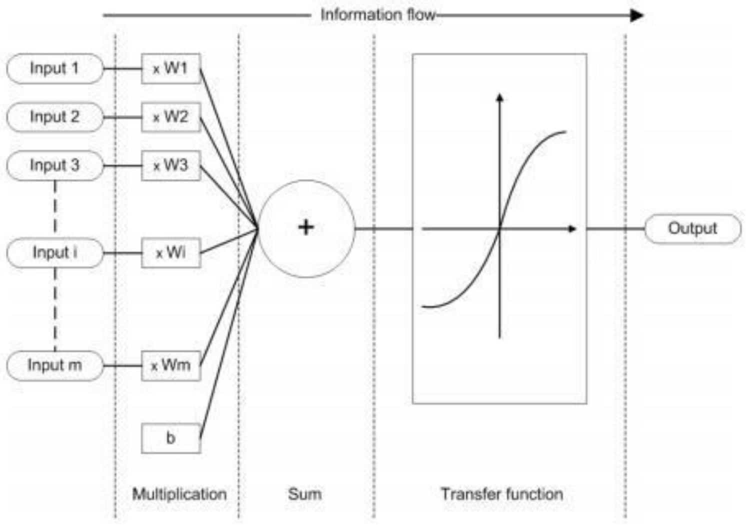


Figure 3.3: Working principle of an artificial neuron

Multiplication, summation, and activation are three straightforward sets of principles that make up this paradigm. The inputs are weighted at the entry of the artificial neuron, which implies that each input value is multiplied by a unique weight. A function called sum that adds up all weighted inputs and bias is located in the center of an artificial neuron. The total of the previously weighted inputs and bias exiting the artificial neuron passes via the activation function, also known as the transfer function .[18]

To create the output for that neuron, the incoming signals, referred to as inputs, are first added together, multiplied by the connection weights (adjusted), and then transmitted via a transfer function. The sigmoid function is the most often used transfer function, while the activation function is the weighted sum of the neuron's inputs [19] Fig. 3.4

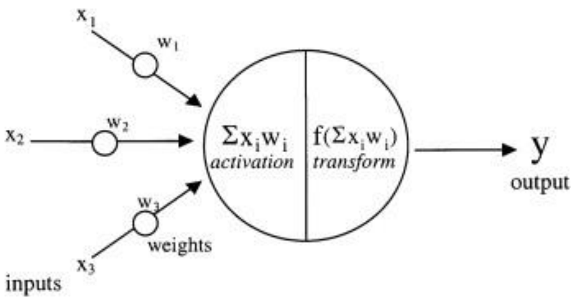


Figure 3.4: Model of an artifical neuron

There are commonly three different types of units, According to [20]

1. **Input unit** This unit is an input unit that receives signals from external sources.

2. **Output Unit** Data is transmitted outside of the network using the output unit.

3. **Hidden Unit** Inside the network, this unit receives and transmits signals.

The functioning of the artificial neural network is significantly influenced by how the neurons are linked to one another. Artificial neurons can receive either excitatory or inhibitory inputs, much as ”actual” neurones. Excitatory inputs drive the following neuron's summing mechanism to add, whereas inhibitory inputs lead it to subtract. Other neurons in the same layer can also be inhibited by a neuron. The term for this is lateral inhibition [19]

## Types

ANNs have found applications in a wide range of fields, including computer vision, natural language processing, and speech recognition. There are several types of ANNs that are commonly used, each with its own strengths and weaknesses [21]

## FFNN

One of the most common types of ANNs is the feedforward neural network (FFNN). A Feed forward neural network has three layers: an input layer of neurons, a hidden layer, and an output layer. FFNNs have exactly one input neuron for each measurement, so that there are as many input neurons as there are patterns in the data set. As many neurons make up the output layer as there are classes in the data collection. With the weights of all the connections between the neurons, one may classify a pattern by feeding its measurements as input to the input neurons and then propagating the output signals down the layers until the output signals of the output neurons are achieved. Each output neuron is assigned to a certain class. A pattern is classified by the output neuron that generates the highest output signal [22].

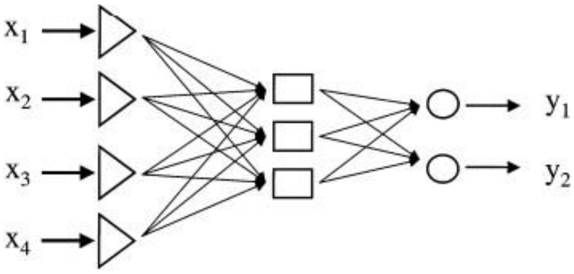


Figure 3.5: Feedforward network

## CNN

Convolutional neural networks (CNNs) are a type of artificial neural network (ANN) that have proven to be highly effective in analyzing image and video data. These networks use convolutional layers to extract patterns and features at multiple levels of complexity, making them particularly well-suited for applications such as facial recognition and object detection. As described in [21] CNNs learn and extract features from images using hierarchical feature learning, achieved through the application of convolutional filters to the image. These filters detect various features, such as edges, corners, and curves.

CNNs offer significant advantages over traditional image recognition algorithms. Firstly, they can learn and extract more complex and abstract features than traditional methods, which rely on manually crafted features. Additionally, they can leverage the spatial relationships between pixels in an image, which is particularly useful for object detection or segmentation tasks

### RNN

The Recurrent Neural Network operates by looping the output of a layer back to its input, enabling it to predict the outcome of the layer. Each neuron in the network functions as a memory cell, retaining information that can be used in subsequent time steps. This process is known as the recurrent neural network process, in which data can be stored and reused for later use. Error correction is used to enhance the accuracy of the prediction by modifying the prediction output. The learning rate determines how quickly the network can accurately predict the desired outcome. Recurrent Neural Networks have a variety of applications, including converting text to speech, and are designed for supervised learning without a requirement for a teaching signal [21]. A recurrent neural network is a type of neural network where there is at least one cycle in the connection graph. The most commonly used RNN is currently the one based on Long Short-Term Memory (LSTM) [23]

### GANs

Generative adversarial networks (GANs) are an artificial neural network (ANN) architecture that aims to produce new data that is similar to the original training dataset. GANs consist of two distinct networks: a generator network that produces new data, and a discriminator network that is responsible for distinguishing between the real and generated data. These two networks are trained concurrently in a game-like scenario, with the generator attempting to deceive the discriminator, and the discriminator trying to accurately classify the data [21].

### Autoencoders

Autoencoders, on the other hand, are ANNs that are designed to compress data by encoding it into a lower-dimensional space and then decoding it back into its original format. This approach makes autoencoders useful for a wide range of tasks, including image denoising, anomaly detection, and data compression [21].

# Chapter 4

# Literature Review, Problem Statement and Objective of Thesis

## Literature Review

There are various studies to detect malicious activities at endpoints. The author of [24] focus on creating a comprehensive system that employs the ELK stack to identify and analyze cyber-attacks both in the network and host. They succeed in a series of Mandiant-based APT-specific attacks, with each stage carried out using methods from the MITRE ATT&CK matrix. They integrate Elasticsearch and VirusTotal to determine whether the files associated with security incidents are infected with malware. Additionally, add geolocation data to network logs using Logstash's GeoIP processor. To look for signs of compromise in real-time, they also combine Elasticsearch with MISP. Furthermore, the system searches through the network traffic for anomalies using Elasticsearch's Machine Learning techniques (based on the Packet beat logs) the solution proposed in [24] may not be able to respond to cyber attacks in real-time. Additionally, used limited set of operating systems in their environment, This means that the solution may not be effective or applicable to all operating systems, which could limit its practicality and generalization.

The tools Sysmon and auditd are used in our suggested system for fine-grained logging, and Elastic Beats is used to forward the generated logs (Winlogbeat, Filebeat). A great job was done by the authors of paper[25] they present the idea of tactical provenance graphs (TPGs), which analyze causal relationships between threat alarms generated by EDR instead of encoding low-level system event dependencies. Analysts can quickly investigate multi-stage threats with the help of TPGs' concise visualization. To address EDR's false detection problem, they develop a threat scoring approach that measures risk depending on the temporal ordering of individual threat alerts with in TPG. Rather than keeping bulky system logs, they kept a skeletal graph that can provide likability between existing and future threat alarms. They use the Symantec EDR software in a corporate environment to assess Rap Sheet, their system. Results indicate that their method can rank malicious TPGs higher than TPGs that are only raising false alarms. Additionally, skeletal graph reduces log retention's long-term load by as much as 87The limitation of their work was that their log reduction technique may not always be able to maintain the ability to produce precise TPGs from the skeleton graph for future warnings if some attack activity does not enable the underlying EDR to raise an alert.

Reference [26] combined open source software framework such as GRR (Google Rapid Response) and osquery. The suggested EDR system's APT coverage is examined using MITRE's Adversarial Tactics, Techniques, and Common Approach. The criteria for detecting security incidents were developed by organising the information and query statement needs needed to detect MITRE ATT&CKs. The total of 381 query statements then examined to determine whether they met the definitions of detection and coverage.

The evaluation results suggest that APT strategies with high levels of threat detection employing non-customized osquery setups account for 28.5% of all detections, which is smaller compared to the other response levels. customized settings are needed to get the most of GRR and osquery-based open source EDR. They consider only Linux based operating system as an endpoint and also limited threat intelligence were in their proposed solution

Na-Eun Park et al. [27] designed an endpoint detection system using Google rapid response (GRR) and Elasticsearch Auditbeat, an extra logging component In the EDR context, they provide a quick detection and response technique for the early detection of the APT assault process. The efficiency of the suggested approach was examined by The detection coverage was examined using an attack environment created with an open-source APT assault emulator. States were made up depending on the MITER ATT&CK to build a detection criteria in the APT attack stage utilizing the emulator. Then, during observing and collecting log data with GRR and Auditbeat, whenever a combination of state transitions matching a specified ruleset was observed, the attack recognition and response activities could be carried out, resulting in a quick detection of the APT offensive chain. Their limitation was that they consider only windows Operating system in their experiments and chances of false posiivity rate also high

Chanwoong Hwang et al. [1] present a method for detecting previously unknown cyber attacks by examining event logs, even when typical antivirus technologies have failed. The method, which is based on semi-supervised learning, combines AutoEncoder and 1D CNN. A dataset gathered over a month in a real-world business endpoint setting was used by the authors to evaluate their methodology. They found 37 unknown threats, 26 of which were later determined to be malicious by VirusTotal.

The author of this paper .[28] employed natural language processing techniques to tackle the issue of spam email detection. To achieve this, a range of machine learning algorithms such as Naive Bayes, K-Nearest Neighbors, SVM, Logistic regression, Decision tree, and Random forest were utilized. Through training these algorithms on a ready made dataset from kaggle, the study observed remarkable accuracy levels, with Logistic regression and Naive Bayes surpassing others and achieving up to 99% accuracy in detecting spam emails. These findings present promising opportunities for developing an intelligent and robust spam detection classifier. By potentially combining multiple algorithms or incorporating advanced filtering methods, it is possible to enhance the efficacy and sophistication of spam detection systems even further. A limitation of their study [28] was their dependence on a pre-existing dataset sourced from Kaggle instead of generating a real-time dataset. This choice may impact the applicability of their findings as the characteristics of spam emails can evolve over time, potentially limiting the generalizability of their results.

Furthermore, their research had another limitation in that it did not incorporate the analysis of phishing or malicious links. Since these types of emails are commonly encountered in spam campaigns, neglecting them could undermine the effectiveness of the spam detection model in real-world scenarios.

Most of the existing studies consider limited endpoints in their experiments or consider limited operating system and limited threat detection and intelligence platforms were integrated, However, with the growing complexity of endpoint attacks and the increasing likelihood that they will spread from endpoints to cloud environments, it is important to develop an EDR solution that can detect and respond to advanced threats across a broader range of endpoints, including those in dockerized and cloud environments. This type of EDR solution would enable organizations to detect and respond to sophisticated attacks in real-time and help prevent future attacks.

## Problem Statement

As per literature survey we can conclude that still there is a lack of visibility into endpoint-level events which is a major gap in effective endpoint detection and response. Endpoint detection and response systems should deliver real time deep visibility into end user and process activities, traffic patterns, as well as other events occurring from the endpoints. Secondly current countermeasures need improvement in incident response factor. Furthermore, at some point there is a greater emphasis on network level detection, with less emphasis on endpoint level detection and real time analysis. Since Phishing emails are the most typical way for attackers or hackers to propagate malware. Therefore, its is time to need such technology that stops malware spreading through any endpoint with enhanced detection and real time response capability

Another significant limitation in the current scenario is the cost of EDR solutions. The development of a true open source solution with practically all functionality is both necessary and difficult. Indeed, constructing such a comprehensive infrastructure is so demanding (in terms of difficulties, resources, and time) that businesses simply charge a fee or offer a subscription. As a result, there is no true economic solution for all wants and requirements.

## Objective

The major objective of my thesis is to propose such an AI-based EDR solution with open-source tools that supplement endpoint security with increased detection, investigation, and response capabilities we will be able to detect malicious code, such as malware, viruses, or ransomware on endpoints, with proper remediation instructions to prevent further damage in known cases. Furthermore, through this work, we will be able to stop malware spreading either through malicious downloading or via phishing link

# Chapter 5

# Proposed Methodology

The following chapter presents the proposed methodology and the software tools that were utilized to achieve the desired solution for this thesis. The methodology describes the step-by-step approach used to achieve the research objectives, while the software tools were carefully selected based on their capabilities to support the methodology.

## Proposed Methodology

The process begins with the collection of data from diverse endpoints, including servers, desktops, laptops, virtual machines, and cloud instances. This collected data is subsequently transmitted to a centralized management server, referred to as Wazuh manager, where it is analyzed using open-source threat intelligence platforms and AI-based detection techniques. The management server utilizes threat intelligence sources such as VirusTotal , yara and MITRE Attack Framework to enhance its ability to detect security threats and generate alert.

The integration of compliance requirements and threat intelligence sources provides valuable context for security analytics, making it easier to identify and respond to potential security incidents. Following the analysis, the data is represented visually via a web-based user interface. The ”Active Response” module is then activated, which triggers automated actions such as blocking network connections, stopping running processes, or deleting malicious files in response to potential threats. This approach is depicted graphically in Fig. 5.1

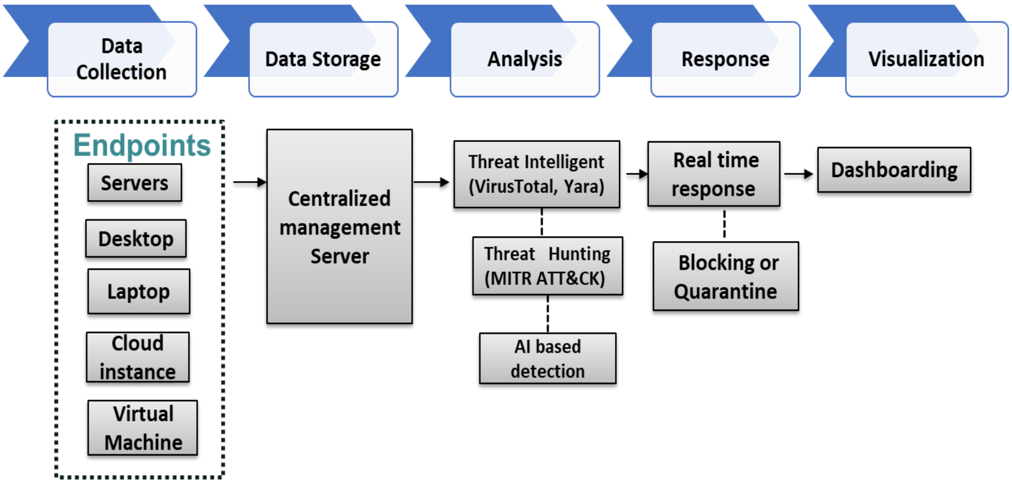


Figure 5.1: Proposed Solution

## Implementation platform

The diagram below 5.2 shows the open-source software programs that were utilized to implement our EDR (Endpoint Detection and Response) system based on AI intelligence. We used wazuh as the main backbone solution and rspamd for email filtering at an endpoint. As Wazuh architecture relies on agents that run on monitored endpoints, that forward security data to a centralized server, the central server decodes and examines the incoming information and passes the results along to the Wazuh indexer for indexing and storage. Further, we integrate wazuh with VirusTotal and yara, Once VirusTotal identifies a file as a threat, Wazuh is configured to trigger an active response to remove the file from the system. We have integrated Mailcow, Rspamd, and Redis to create a dependable and efficient email hosting platform. Mailcow forms the core of the email server, with features such as email delivery, a webmail interface, and support for encryption and authentication protocols. Rspamd, as a spam filtering solution, identifies and blocks spam messages, and Redis, an in-memory data structure store, helps to improve email delivery speed.

Our integrated solution provides a secure and adaptable email hosting platform suitable for organizations of all sizes. The integration of Mailcow, Rspamd, and Redis delivers a highly effective email server solution that offers fast and reliable email delivery and filtering.

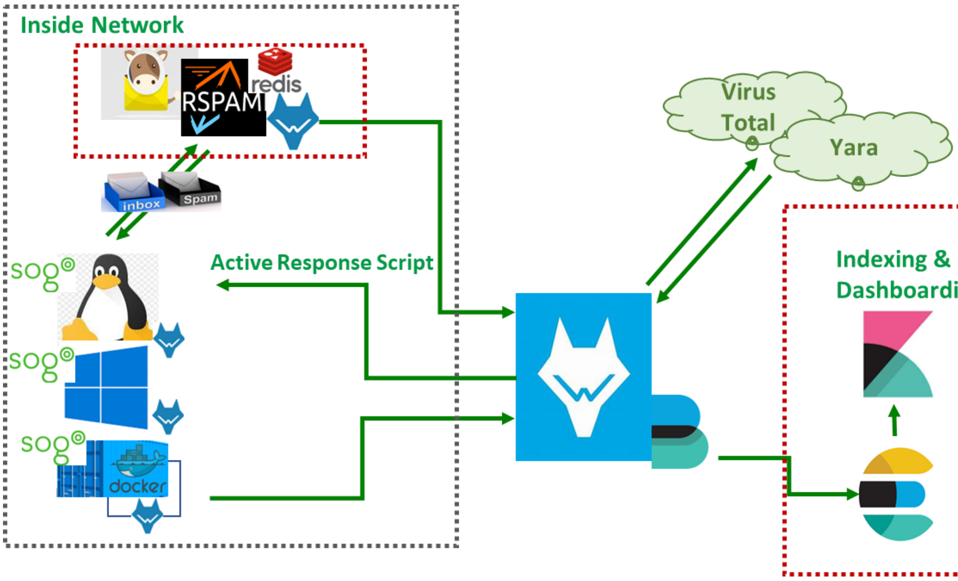


Figure 5.2: Implementation platform of proposed Architecture

## Components Description

This section will provide a detailed description and explanation of all the components utilized in the development of the product.

## Endpoint Agent

The primary component is the agent software that runs on each endpoint, collecting security data and sending it to a centralized server for analysis. We installed wazuh agent on ubuntu 20 machine and windows 10 for data collection. The Wazuh agent has quite a modular structure in which several components handle their own responsibilities such as monitoring the system files, analyzing log messages, gathering inventory data, analyzing system settings, looking for malware, and so on.

## Central Management Server

Once data has been collected from endpoints by agents that will be forwarded to central server. The Wazuh central server is in responsible of evaluating data from the agents and generating alerts when risks or anomalies are found. When configured as a cluster, a single dedicated server can assess data from a large number of agents and scale horizontally. The agents are also managed by the server, which is used to remotely configure and upgrade them as needed.

Additionally, it’s intended to remotely control the agents’ settings and stay updated on their performance. To enhance its detection abilities, the Wazuh server makes use of threat intelligence sources.

## Indexing and Dashboarding

For indexing and dash-boarding we utilized Elastic Stack which is an integrated collection of well-known open-source log management solutions, such as Elasticsearch, Kibana, File beat, and others

**Elasticsearch** Elasticsearch is the stack’s Database Management System (DBMS) a full-text search and analysis engine with great scalability. Elasticsearch is distributed, which implies that the data indices are divided into shards, with each shard having zero or maybe more replicas [29]

**Logstash** Logstash is an event processing platform that runs in real time. In other terms, it can collect data from several sources, process it via a filter or by adding metadata, and then transport it to one or much more locations. It can analyze any kind of event, particularly those in JSON and XML format [29]. Where as Filebeat is a simple forwarder that transports logs across a network, typically to Elasticsearch. It is used by the Wazuh server to send Elasticsearch events and alarms. It reads the Wazuh analysis engine’s output and sends events in real time over an encrypted connection [30]

**Kibana** A versatile and user-friendly web interface for data mining, visualization, and analysis. It runs on top of an Elasticsearch cluster’s indexed content. This is also used to administer Wazuh configuration as well as to check its performance [30]

## Email Server

This section describes the architecture of our email server, We installed an email server on Ubuntu 20.04 that is a Docker-based email server that is built on Dovecot, Postfix, and other accessible software and has a modern online UI for management. The integrated UI provides for separate domain administrator and mailbox user access as well as administrative operations on your mail server instance. Provisioning and managing sophisticated email apps is simple. Mailcow: Dockerized includes several containers connected by a single bridging chel. One application is represented by each container. Fig. 5.3 below depicts mail server architecture

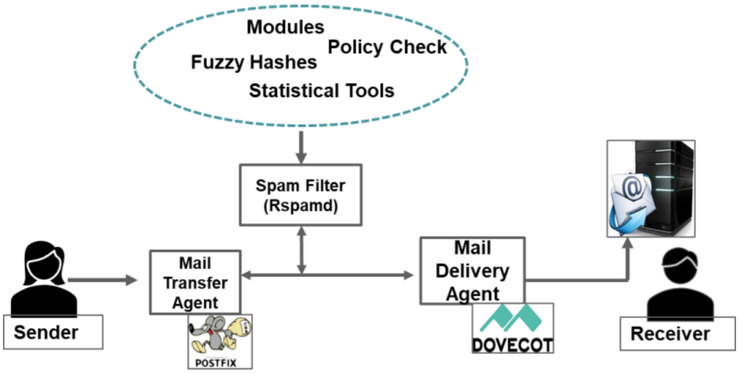


Figure 5.3: Email Server Architecture

For the function of MTA we used Postfix, which is in charge of routing email to its intended destination. If the mailbox for the recipient domain is on our server, Postfix will accept the message for delivery. In every other case, Postfix will forward the message to the mail server of the receiver. [31]

A message must first pass through the rspamd milter before it can be transmitted (you sent out an email towards someone) or permitted for transmission (someone sent user an email). (The term ”milter” just refers to a Postfix mail filter.) Rspamd does some spam filtering on incoming mail and may mark the message as spam or reject it entirely. The DKIM signature is all that rspamd handles for outgoing mail

When an email is accepted and passes the rspamd filter, it is sent to Dovecot, the Mail Delivery Agent (MDA), which receives the email from Postfix and stores it in the user’s mailbox. [31]

### In Memory Database: Redis

To store and retrieve spam filtering-related data, including blacklists, whitelists, and statistical data used for machine learning-based filtering, Rspamd employs Redis as a backend. Temporal information needed for message processing, like message IDs and metadata, is also stored in it. Redis speeds up retrieval of frequently visited data by storing it in memory, which lowers the system's total response time.

### Webmail Interface: SOGO

SOG0 is used as the end-user webmail and calendar interface. It provides a full-featured webmail experience for users to manage their email accounts, such as composing and sending messages, organizing email folders, and searching for specific messages. Instead of utilising a different email client or calendar programme, clients can access their email accounts and calendars using a web interface offered by SOGo.

### Spam Filter

Rspamd is an open-source spam filtering system that provides a number of capabilities to protect email servers and users against spam, phishing, malware, and other email-borne risks. Rspamd suggests a decision for the MTA to do with regard to the message, such as passing, rejecting, or adding a header, based on this spam value and the user's choices. Further detailed description about rspamd is explained in next section

## Key Features and Capabilities of Spam Filtering Solution

Several helpful features are offered by Rspamd, which is built to handle hundreds of messages every second. It offers a number of crucial features, some of which are:

1. Bayesian filtering: Rspamd uses a Bayesian filter to classify emails as either spam or ham based on their content. The filter works by analyzing the frequency of words and phrases in the email and comparing them to a set of known spam and ham messages.

2. Fuzzy matching: Rspamd performs a fuzzy message matching using the shingles method. This probabilistic method filters spam or ham messages by looking for patterns in word chains (similar to the shingles algorithm). For this algorithm, we employ a set of hash functions, including mumhash, siphash, and others, as well as 3-grams (trigrams). Currently, rspamd utilities 32 hashes for shingles.

3. Neural networks: Rspamd uses neural networks to classify emails based on their content. The neural network is trained on a set of known spam and ham messages, and can adapt to new patterns of spam over time.

4. Greylisting and rate limiting: Rspamd uses greylisting and rate limiting to stop spam and other email-borne dangers from inundating the email server and using up system resources.

5. DKIM and DMARC support: To confirm the validity of email communications and prevent spoofing and phishing attempts, Rspamd supports the DomainKeys Identified Mail (DKIM) and Domain-based Message Authentication, Reporting and Conformance (DMARC) authentication protocols.

6. URL and reputation checks: Rspamd discovers and block potentially harmful links by comparing URLs found in email messages to various reputation databases.

7. Customizable rules and actions: Rspamd enables users to design unique rules and procedures that are tailored to their own requirements and preferences.

8. Web interface and statistics: Rspamd offers a web interface that enables users to administer email filtering, track email traffic, and view filtering performance statistics.

9. Integration with popular email servers: Rspamd offers seamless email filtering and security by integrating with well-liked email servers like Postfix, Exim, and Dovecot.

## Spam and Ham Email Trained model

We started by collecting a dataset of 1000 spam and 1000 ham emails, of which 600 of both We collected by ourself by generating fake emails with phishing URLs, and the rest of the emails I collected from a publicly available dataset, to train and test my machine learning model. To prepare the email data for use in my model, We applied several preprocessing steps. We applied preprocessing steps to extract URLs from the email text using regular expressions and calculated features such as URL length and number of special characters, storing them in new columns of my Data Frame to prepare the data for use in my machine learning model. After preprocessing the data, We split the dataset into training and testing sets using an 80/20 split, with 80% of the data used for training and 20% for testing. A detailed depiction of the entire process involved in our model training can be observed in Fig. 5.4

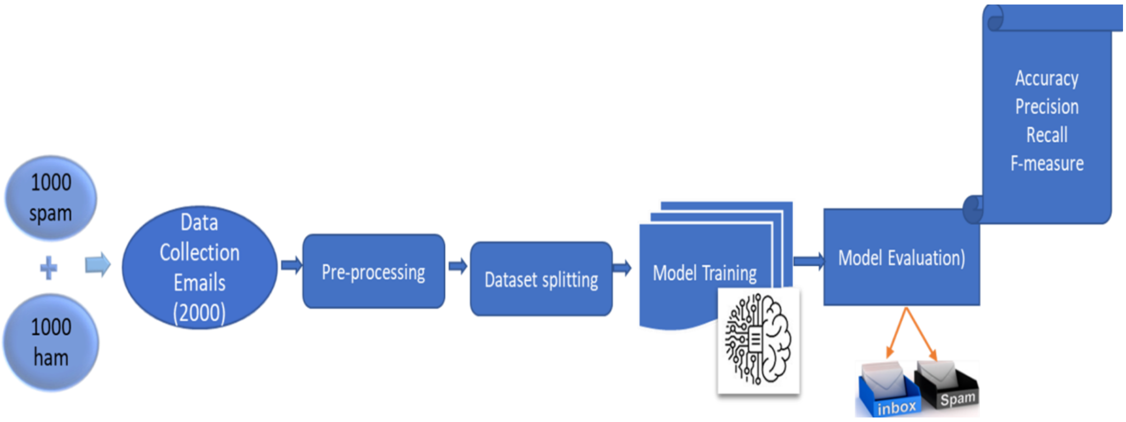


Figure 5.4: Steps of AI Model Training

We then trained my machine learning model using the training set of emails, using an artificial neural network (ANN) with 3 layers. The activation functions used were ReLU for the first two layers and sigmoid for the third layer. We measured the accuracy of the model using cross-entropy loss and used ADAM, a powerful optimization algorithm, with a batch size of 32, 10 epochs.

Finally, we evaluated the performance of my machine learning model on the testing set of emails. we measured the model’s accuracy using metrics such as recall, precision, and F1 score. This process allowed me to determine the effectiveness of my machine learning model in accurately identifying spam and ham emails. Overall, this approach using ANN with ReLU and sigmoid activation functions showed high accuracy and proved to be a powerful tool in email classification.

## ROC curve

The below figure 5.5 depicts our trained model's ROC curve diagram, which is a visual depiction of how a binary classifier performs. It plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at different classification thresholds, indicating how accurately the classifier can differentiate between positive and negative classes. When applied to an email classifier, the ROC curve would illustrate its ability to differentiate between spam and non-spam emails. By plotting the classifier's predictions as a series of points on the curve, the FPR is shown on the x-axis and the TPR on the y-axis.

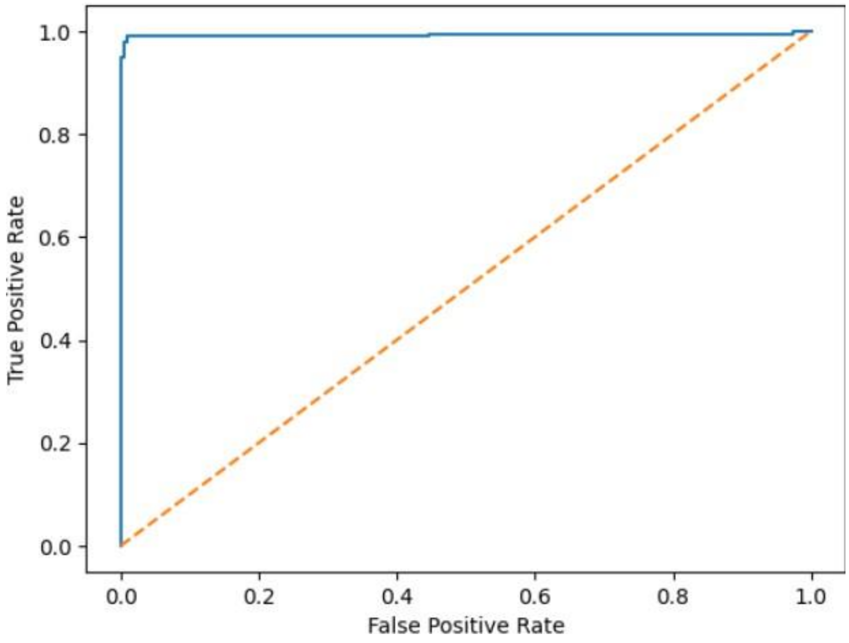


Figure 5.5: ROC Curve

## Confusion Matrix

A confusion matrix is a valuable method for evaluating the performance of a classification model. It enables a comparison between the model's predictions and the actual outcomes. The matrix consists of four distinct categories: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). TP represents the correct positive predictions, FP represents the incorrect positive predictions, TN represents the correct negative predictions, and FN represents the incorrect negative predictions. By analyzing the values within these categories, we can calculate key metrics such as accuracy, precision, recall, and F1-score. This analysis provides essential insights into the model's classification accuracy and potential areas for improvement.

Table 5.1: Confusion Matrix

| | **Predicted Negative** | **Predicted Positive** |
|---|---|---|
| **Actual Negative** | 197 (TN) | 2 (FP) |
| **Actual Positive** | 4 (FN) | 197 (TP) |

The analysis of the presented table 5.1 confirms the model's exceptional accuracy in classifying spam and ham emails. With 197 spam emails correctly identified as true positives and only 2 false positives, the model demonstrates precise discrimination. It exhibits a low rate of false negatives 4, accurately identifying 197 non-spam emails as true negatives. These results solidify the model's effectiveness in distinguishing between spam and ham emails. Visual representation of Confusion matrix is displayed in Figure below Fig. 5.6

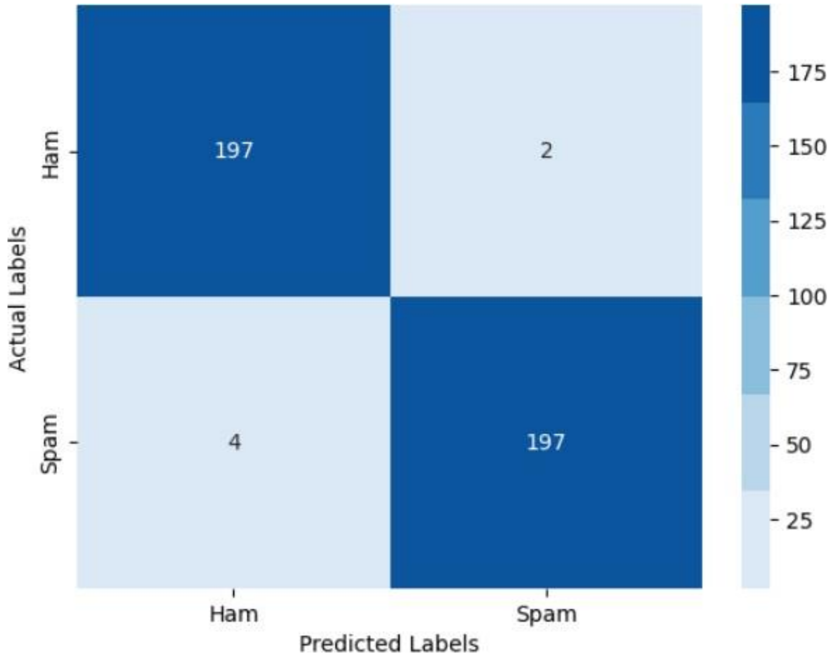


Figure 5.6: Confusion Matrix

## Performance Metrics

The effective utilization of Confusion Matrix results is crucial for evaluating the email detection system using metrics like Accuracy, Precision, Recall, and F1 score. Accuracy measures the overall correctness of email classification, while Precision focuses on accurately identifying relevant emails and minimizing false positives. Recall measures the system's ability to capture all relevant emails and minimize false negatives. The F1 score combines Precision and Recall to provide a balanced assessment of the system's performance.

The corresponding metric values, derived from the Confusion Matrix, are concisely displayed in a summary Table 5.2 enabling accurate assessment of the email detection system's efficacy

Table 5.2: Performance Metrics

| **Accuracy** | **Precision** | **Recall** | **F1 Score** |
|---|---|---|---|
| 0.9850 | 0.9899 | 0.9801 | 0.9850 |

Overall, the evaluation of the email detection system's performance using these metrics and the Confusion Matrix provides valuable insights for decision-making, underscoring their significance in effectively distinguishing between spam and ham emails.

## Attack and Detection Experiments

### Virtual Machines Specifications

We needed to build up a test environment in order to analyze our proposed EDR. We have successfully installed the Wazuh manager on an Ubuntu 20.04-based machine. Additionally, we have deployed three Wazuh agents to collect endpoint data from one Windows 10 machine and two Ubuntu 20.04-based machines. Furthermore, we have designated one of the Ubuntu 20.04 machines as the host for our Dockerized email server, Mailcow.

The virtual machines' specifications are shown in table 5.3

Table 5.3: Virtual Machine's Specification

| Components | Operating System | RAM | **Hard Disk** | **Processors** | **Network Adapter** |
|---|---|---|---|---|---|
| Wazuh Manager | Ubuntu 20.04 | 8GB | 500GB | 4 | NAT |
| Endpoint Agent 1 | Ubuntu20.04 | 4GB | 250GB | 4 | NAT |
| Endpoint Agent 2 | Windows 10 | 4GB | 250GB | 4 | NAT |
| Email Server | Ubuntu 20.04 | 4GB | 500GB | 4 | NAT |

### Use Case1: Malware Detection and response via Virustotal

VirusTotal is a comprehensive platform that utilizes a variety of antivirus solutions and an online analyzing engine to scan files and URLs for different types of malicious content, such as viruses, worms, rootkits, trojans, and other dangerous content. It is also capable of detecting false positives. Integrating this tool with our FIM engine gives an easy way to scan the files being monitored for harmful content. When VirusTotal recognizes a file as a threat, Wazuh is set to launch an immediate reaction to remove it from the system.

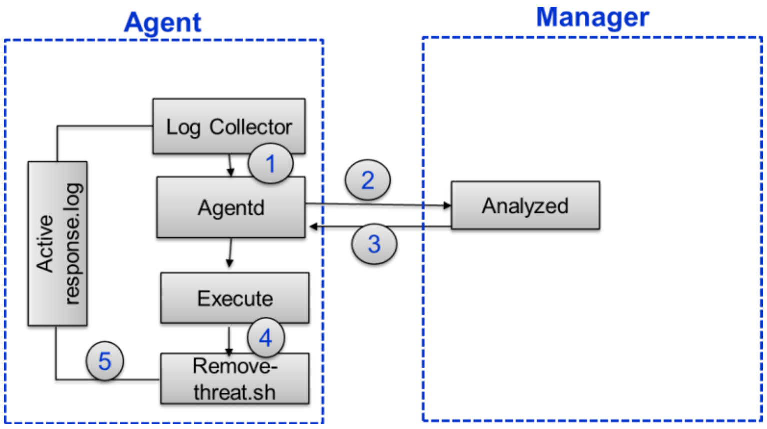


Figure 5.7: Malware detection via VirusTotal

The VirusTotal API is used in this integration to detect malicious information within the files analyzed by File Integrity Monitoring module. The Fig. 5.7 shows how this integration works:

1. The Wazuh agent employs the File Integrity Monitoring (FIM) module to continuously monitor the file system. This module scans monitored folders for file additions, updates, or deletions, while also preserving the hash values of these files. It promptly generates alerts as soon as any modifications are detected.

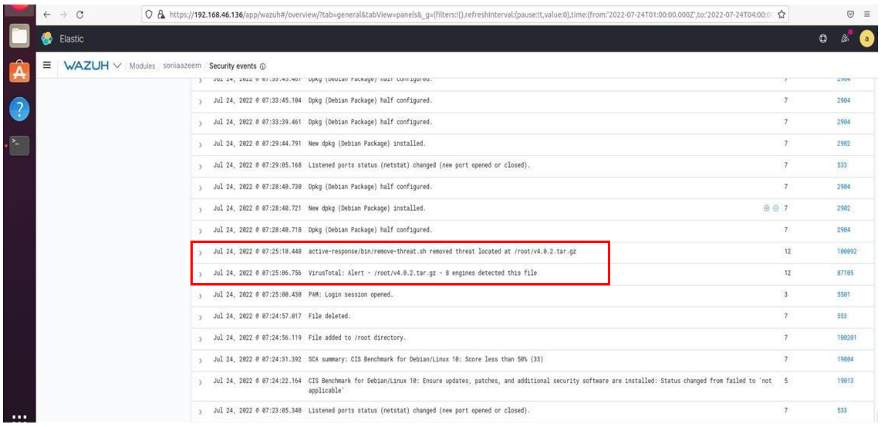
2. When a file change is detected, the agent sends a FIM event to the Wazuh manager for further analysis.

3. Wazuh Manage then does a comparison between the extracted hash and the data in the Virustotal database by sending an HTTP Request message to the VirusTotal database by using VirusTotal API.

4. Upon detecting the existence of malware through the use of virtual scanning technology, the Wazuh manager triggers a command to be sent to the relevant agent.

5. Next, The agent subsequently executes a bash script that is specifically designed to eliminate the identified malware from the impacted endpoint.

On victim endpoint with Operating System Ubuntu 20.04 we downloaded malicious file from ”https://dasmalwerk.eu/” website and detect it on the other wazuh manager side. The diagram presented below Fig. 5.8 shows an alert for malware detection via VirusTotal.


https://192.168.46.136/app/wazuh#/overview/?tab=general&tabView=panels&_g=(filters:!(),refreshInterval:(pause:!t,value:0),time:(from:'2022-07-24T01:00:00.000Z',to:'2022-07-24T04:00:0
Elastic
WAZUH
Modules
soniaazeem
Security events

```
Jul 24, 2022 @ 07:33:45.104  Dpkg (Debian Package) half configured.                                  7   2904
Jul 24, 2022 @ 07:33:39.461  Dpkg (Debian Package) half configured.                                  7   2904
Jul 24, 2022 @ 07:29:44.791  New dpkg (Debian Package) installed.                                    7   2902
Jul 24, 2022 @ 07:29:05.168  Listened ports status (netstat) changed (new port opened or closed).    7   533
Jul 24, 2022 @ 07:28:40.730  Dpkg (Debian Package) half configured.                                  7   2904
Jul 24, 2022 @ 07:28:40.721  New dpkg (Debian Package) installed.                                    7   2902
Jul 24, 2022 @ 07:28:40.718  Dpkg (Debian Package) half configured.                                  7   2904
Jul 24, 2022 @ 07:25:10.448  active-response/bin/remove-threat.sh removed threat located at /root/v4.0.2.tar.gz   12   100092
Jul 24, 2022 @ 07:25:06.756  VirusTotal: Alert - /root/v4.0.2.tar.gz - 8 engines detected this file  12   87105
Jul 24, 2022 @ 07:25:00.430  PAM: Login session opened.                                              3   5501
Jul 24, 2022 @ 07:24:57.017  File deleted.                                                           7   553
Jul 24, 2022 @ 07:24:56.119  File added to /root directory.                                          7   100201
Jul 24, 2022 @ 07:24:31.392  SCA summary: CIS Benchmark for Debian/Linux 10: Score less than 50% (33)   7   19004
Jul 24, 2022 @ 07:24:22.164  CIS Benchmark for Debian/Linux 10: Ensure updates, patches, and additional security software are installed: Status changed from failed to 'not   5   19013
                             applicable'
Jul 24, 2022 @ 07:23:05.340  Listened ports status (netstat) changed (new port opened or closed).    7   533
```

Figure 5.8: Virustotal Alert

## Use Case2: Malware detection via YARA

The detection of malware utilizing YARA and Wazuh is a multi-step process that includes a number of stages. The diagram below 5.9 illustrates all the steps involved, and each of these steps explained below

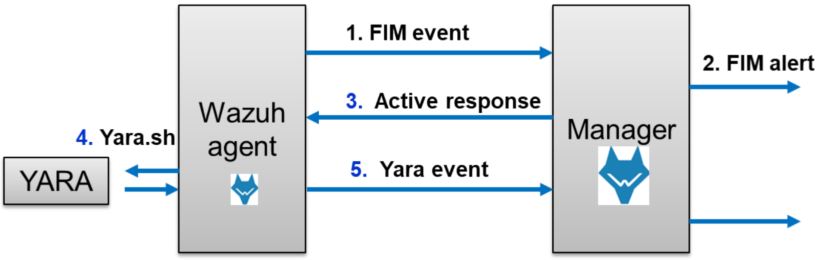


Figure 5.9: Malware detection via Yara

1. The Wazuh agent continuously monitors the file system using the File Integrity Monitoring (FIM) module. When a file change is detected, the agent sends a FIM event to the Wazuh manager.

2. The Wazuh manager receives the FIM event from the agent and analyzes it to see if it matches a pre-configured FIM rule. If the FIM event matches a rule, the manager triggers a FIM alert.

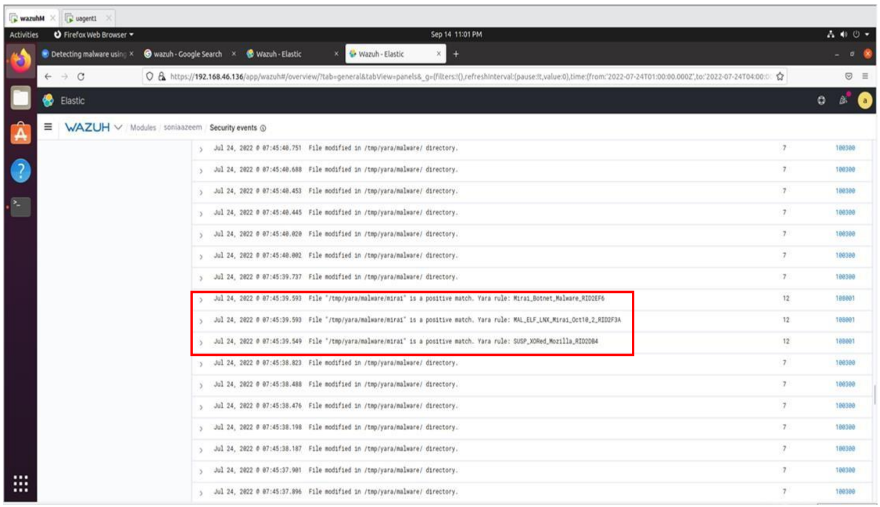
3. The Wazuh manager then initiates an active response module to respond to the FIM alert. The active response module sends a command to the Wazuh agent to execute a specific script.

4. The Wazuh agent receives the command from the Wazuh manager and executes a YARA script that matches the signature of the detected file. The YARA script contains a set of rules that describe the characteristics of the malware.

5. If the YARA script matches the signature of the detected file, the Wazuh agent sends a YARA event alert to the Wazuh manager. The YARA event alert contains information about the matched YARA rule, the path of the detected file, and other relevant information.

6. The Wazuh manager receives the YARA event alert from the agent and analyzes it to determine if it matches a pre-configured YARA rule. If the YARA event alert matches a rule, the manager triggers a YARA alert.

In order to demonstrate the effectiveness of Yara-based malware detection, we conducted a test in which we deliberately installed malware on an Ubuntu-based endpoint. The Yara alert resulting from the test is depicted in the Fig. 5.10 below.



https://192.168.46.136/app/wazuh#/overview/?tab=general&tabView=panels&_g=(filters:!(),refreshInterval:(pause:!t,value:0),time:(from:'2022-07-24T01:00:00.000Z',to:'2022-07-24T04:00:0

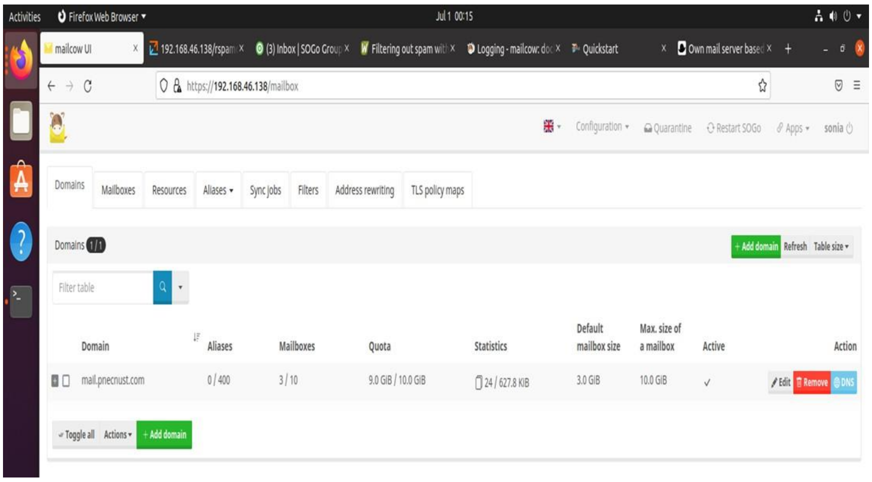
```
Jul 24, 2022 @ 07:45:40.751  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:40.688  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:40.453  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:40.445  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:40.020  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:40.002  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:39.737  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:39.593  File "/tmp/yara/malware/mirai" is a positive match. Yara rule: Mirai_Botnet_Malware_RID2EF6      12  108001
Jul 24, 2022 @ 07:45:39.593  File "/tmp/yara/malware/mirai" is a positive match. Yara rule: MAL_ELF_LNX_Mirai_Oct10_2_RID2F3A  12  108001
Jul 24, 2022 @ 07:45:39.549  File "/tmp/yara/malware/mirai" is a positive match. Yara rule: SUSP_XORed_Mozilla_RID2DB4        12  108001
Jul 24, 2022 @ 07:45:38.823  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:38.488  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:38.476  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:38.198  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:38.187  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:37.901  File modified in /tmp/yara/malware/ directory.                                      7   100300
Jul 24, 2022 @ 07:45:37.896  File modified in /tmp/yara/malware/ directory.                                      7   100300
```

Figure 5.10: Yara Alert

## Use Case3: Spam email Filtering

1. Firstly we installed Mailcow (dockerized mail server) on Ubuntu 20.04

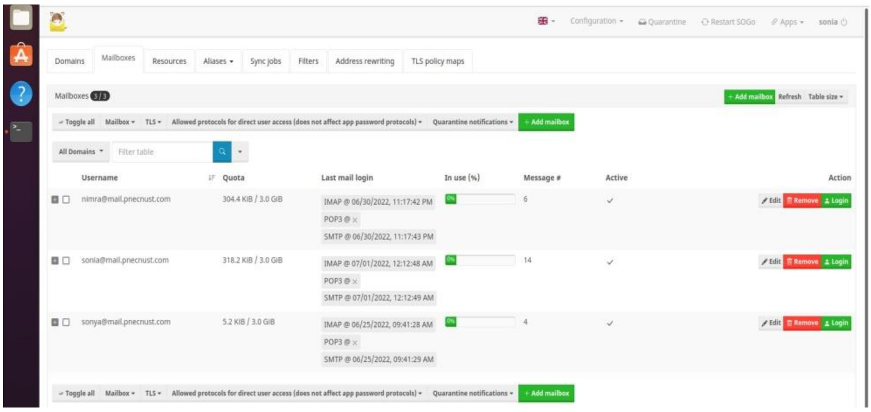
2. Add domain ”mail.pnecnust.com”



| | Domain | Aliases | Mailboxes | Quota | Statistics | Default mailbox size | Max. size of a mailbox | Active | Action |
|---|---|---|---|---|---|---|---|---|---|
| ☐ | mail.pnecnust.com | 0 / 400 | 3 / 10 | 9.0 GiB / 10.0 GiB | 24 / 627.8 KiB | 3.0 GiB | 10.0 GiB | ✓ | Edit Remove DNS |

Figure 5.11: Adding domains

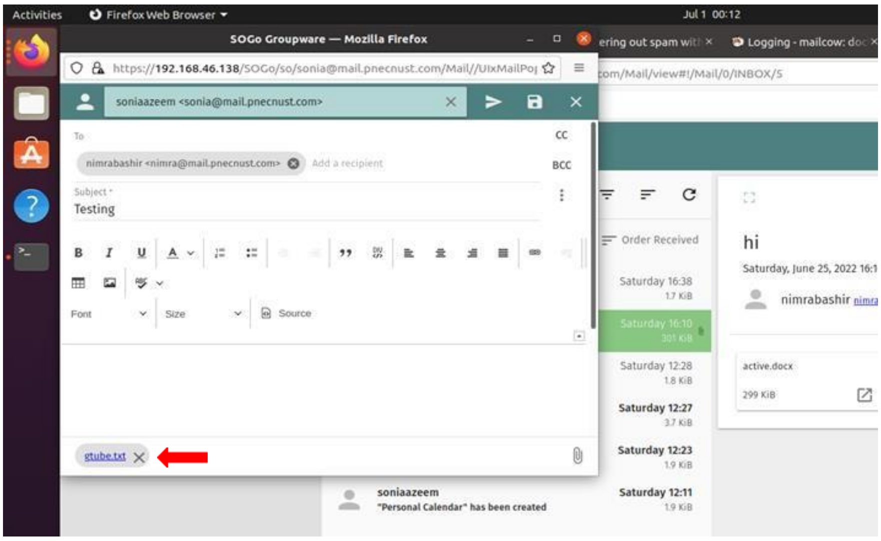
3. Add mailboxes, as per requirement and restart SOGo



| | Username | Quota | Last mail login | In use (%) | Message # | Active | Action |
|---|---|---|---|---|---|---|---|
| ☐ | nimra@mail.pnecnust.com | 304.4 KiB / 3.0 GiB | IMAP @ 06/30/2022, 11:17:42 PM<br>POP3 @ x<br>SMTP @ 06/30/2022, 11:17:43 PM | 0% | 6 | ✓ | Edit Remove Login |
| ☐ | sonia@mail.pnecnust.com | 318.2 KiB / 3.0 GiB | IMAP @ 07/01/2022, 12:12:48 AM<br>POP3 @ x<br>SMTP @ 07/01/2022, 12:12:49 AM | 0% | 14 | ✓ | Edit Remove Login |
| ☐ | sonya@mail.pnecnust.com | 5.2 KiB / 3.0 GiB | IMAP @ 06/25/2022, 09:41:28 AM<br>POP3 @ x<br>SMTP @ 06/25/2022, 09:41:29 AM | 0% | 4 | ✓ | Edit Remove Login |

Figure 5.12: Adding mailboxes

4. Access one mailbox from one system and send malicious attachment to another email address



https://192.168.46.138/SOGo/so/sonia@mail.pnecnust.com/Mail//UIxMailPo

com/Mail/view#!/Mail/0/INBOX/5

soniaazeem <sonia@mail.pnecnust.com>

To: nimrabashir <nimra@mail.pnecnust.com>

Subject *: Testing

gtube.txt

| | |
|---|---|
| Saturday 16:38 | 1.7 KiB |
| Saturday 16:10 | 301 KiB |
| Saturday 12:28 | 1.8 KiB |
| Saturday 12:27 | 3.7 KiB |
| Saturday 12:23 | 1.9 KiB |
| Saturday 12:11 | 1.9 KiB |

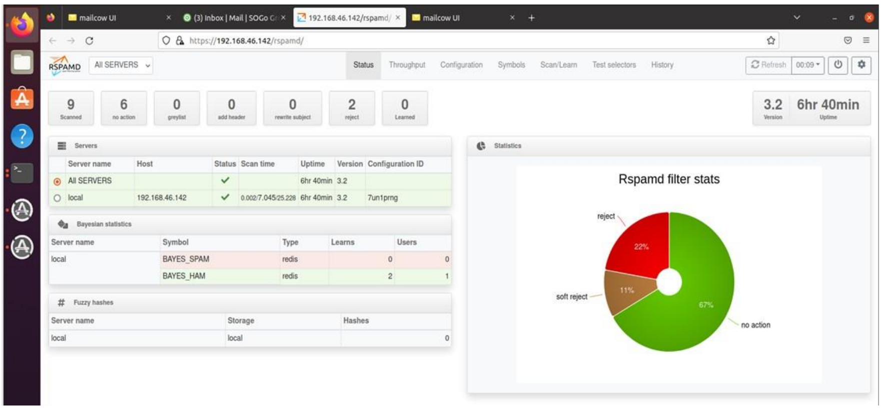
soniaazeem
"Personal Calendar" has been created

Saturday, June 25, 2022 16:1
nimrabashir nimra

active.docx
299 KiB

Figure 5.13: Sending malicious attachment

5. It can be seen from screenshots that rspamd stopped mail to reach at another end because of malicious attachment



https://192.168.46.142/rspamd/

| 9 | 6 | 0 | 0 | 0 | 2 | 0 |
|---|---|---|---|---|---|---|
| Scanned | no action | greylist | add header | rewrite subject | reject | Learned |

| 3.2 | 6hr 40min |
|---|---|
| Version | Uptime |

Servers

| | Server name | Host | Status | Scan time | Uptime | Version | Configuration ID |
|---|---|---|---|---|---|---|---|
| ● | All SERVERS | | ✓ | | 6hr 40min | 3.2 | |
| ○ | local | 192.168.46.142 | ✓ | 0.002/7.045/25.228 | 6hr 40min | 3.2 | 7un1prng |

Bayesian statistics

| Server name | Symbol | Type | Learns | Users |
|---|---|---|---|---|
| local | BAYES_SPAM | redis | 0 | 0 |
| | BAYES_HAM | redis | 2 | 1 |

Fuzzy hashes

| Server name | Storage | Hashes |
|---|---|---|
| local | local | 0 |

Figure 5.14: Spam filter detecting malicious attachment and rejecting

# Chapter 6

# Results and Findings

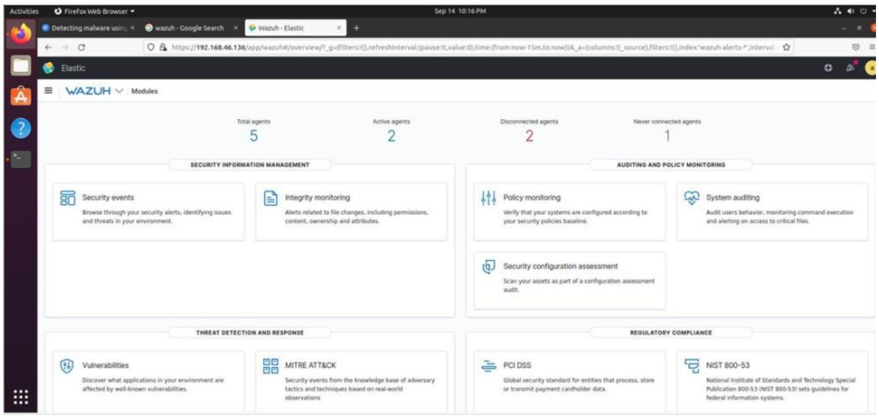
In the final chapter, we will summarize our research findings and highlight the features of our EDR (Endpoint Detection and Response) solution along with screenshots. We will also compare our results with commercial products and relevant literature to showcase our contributions and advancements.

## EDR Features

In this following section, we present a comprehensive list of the EDR features that our solution has successfully achieved.

### Modules Display

Users can view the current state of the platform's modules, such as the agents, manager, and rules. Users can also edit the configuration options for each module



Activities
Firefox Web Browser
Sep 14 10:16 PM
Detecting malware using
wazuh - Google Search
Wazuh - Elastic
https://192.168.46.136/app/wazuh#/overview/?_g=(filters:!(),refreshInterval:(pause:!t,value:0),time:(from:now-15m,to:now))&_a=(columns:!(_source),filters:!(),index:'wazuh-alerts-*',interval
Elastic
WAZUH
Modules

Total agents 5

Active agents 2

Disconnected agents 2

Never connected agents 1

SECURITY INFORMATION MANAGEMENT

Security events
Browse through your security alerts, identifying issues and threats in your environment.

Integrity monitoring
Alerts related to file changes, including permissions, content, ownership and attributes.

AUDITING AND POLICY MONITORING

Policy monitoring
Verify that your systems are configured according to your security policies baseline.

System auditing
Audit users behavior, monitoring command execution and alerting on access to critical files.

Security configuration assessment
Scan your assets as part of a configuration assessment audit.

THREAT DETECTION AND RESPONSE

Vulnerabilities
Discover what applications in your environment are affected by well-known vulnerabilities.

MITRE ATT&CK
Security events from the knowledge base of adversary tactics and techniques based on real-world observations

REGULATORY COMPLIANCE

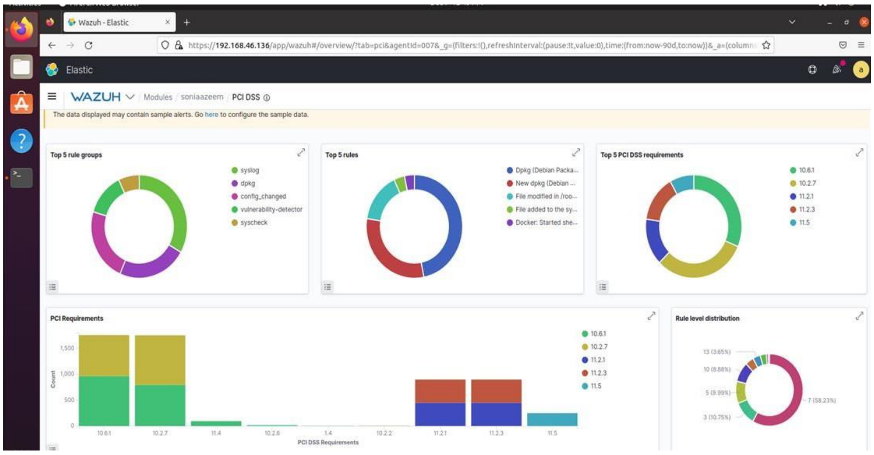
PCI DSS
Global security standard for entities that process, store or transmit payment cardholder data.

NIST 800-53
National Institute of Standards and Technology Special Publication 800-53 (NIST 800-53) sets guidelines for federal information systems.

Figure 6.1: Modules Dashboard

## Compliance management

Compliance module provides automatic checks for compliance with industry standards and regulations such as PCIDSS, HIPPA and CIS

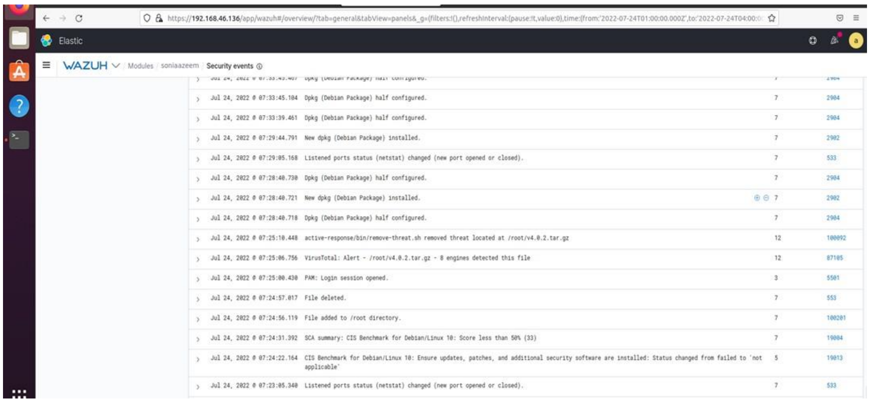


The data displayed may contain sample alerts. Go here to configure the sample data.

Figure 6.2: Compliance management

## Real-time log monitoring

Log monitoring in real time enables users to identify and address security threats as soon as they arise. This function shortens the time needed for incident response and lessens the effect of security issues.

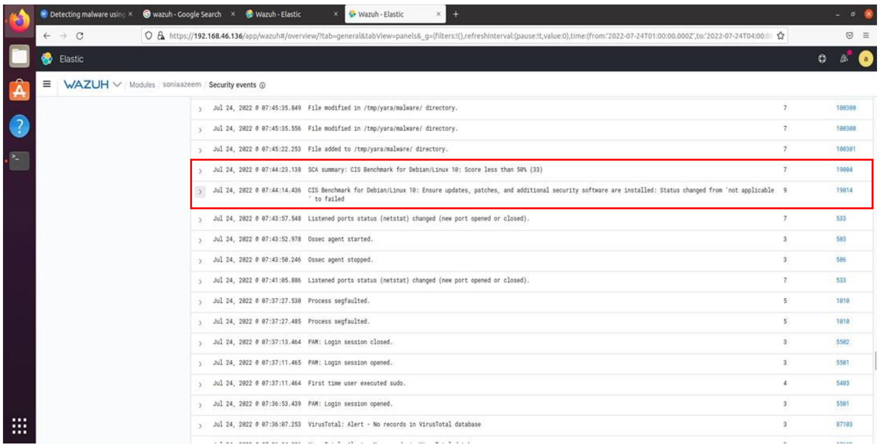

https://192.168.46.136/app/wazuh#/overview/?tab=general&tabView=panels&_g=(filters:!(),refreshInterval:(pause:!t,value:0),time:(from:'2022-07-24T01:00:00.000Z',to:'2022-07-24T04:00:0
Elastic
WAZUH
Modules
soniaazeem
Security events

```
Jul 24, 2022 @ 07:33:45.104  Dpkg (Debian Package) half configured.                                    7   2904
Jul 24, 2022 @ 07:33:39.461  Dpkg (Debian Package) half configured.                                    7   2904
Jul 24, 2022 @ 07:29:44.791  New dpkg (Debian Package) installed.                                      7   2902
Jul 24, 2022 @ 07:29:05.168  Listened ports status (netstat) changed (new port opened or closed).      7   533
Jul 24, 2022 @ 07:28:40.730  Dpkg (Debian Package) half configured.                                    7   2904
Jul 24, 2022 @ 07:28:40.721  New dpkg (Debian Package) installed.                                      7   2902
Jul 24, 2022 @ 07:28:40.718  Dpkg (Debian Package) half configured.                                    7   2904
Jul 24, 2022 @ 07:25:10.448  active-response/bin/remove-threat.sh removed threat located at /root/v4.0.2.tar.gz   12   100092
Jul 24, 2022 @ 07:25:06.756  VirusTotal: Alert - /root/v4.0.2.tar.gz - 8 engines detected this file    12   87105
Jul 24, 2022 @ 07:25:00.430  PAM: Login session opened.                                                3   5501
Jul 24, 2022 @ 07:24:57.017  File deleted.                                                             7   553
Jul 24, 2022 @ 07:24:56.119  File added to /root directory.                                            7   100201
Jul 24, 2022 @ 07:24:31.392  SCA summary: CIS Benchmark for Debian/Linux 10: Score less than 50% (33)  7   19004
Jul 24, 2022 @ 07:24:22.164  CIS Benchmark for Debian/Linux 10: Ensure updates, patches, and additional security software are installed: Status changed from failed to 'not   5   19013
                             applicable'
Jul 24, 2022 @ 07:23:05.340  Listened ports status (netstat) changed (new port opened or closed).      7   533
```

Figure 6.3: Real-time log monitoring

## Configuration Assessment

Configuration assessment modules checks for system misconfigurations and vulnerabilities like open ports , weak passwords and generate reports to notify users


Detecting malware using
wazuh - Google Search
Wazuh - Elastic
Wazuh - Elastic
https://192.168.46.136/app/wazuh#/overview/?tab=general&tabView=panels&_g=(filters:!(),refreshInterval:(pause:!t,value:0),time:(from:'2022-07-24T01:00:00.000Z',to:'2022-07-24T04:00:0
Elastic
WAZUH
Modules
soniaazeem
Security events

```
Jul 24, 2022 @ 07:45:35.849  File modified in /tmp/yara/malware/ directory.                            7   100300
Jul 24, 2022 @ 07:45:35.556  File modified in /tmp/yara/malware/ directory.                            7   100300
Jul 24, 2022 @ 07:45:22.253  File added to /tmp/yara/malware/ directory.                               7   100301
Jul 24, 2022 @ 07:44:23.138  SCA summary: CIS Benchmark for Debian/Linux 10: Score less than 50% (33)  7   19004
Jul 24, 2022 @ 07:44:14.436  CIS Benchmark for Debian/Linux 10: Ensure updates, patches, and additional security software are installed: Status changed from 'not applicable   9   19014
                             ' to failed
Jul 24, 2022 @ 07:43:57.548  Listened ports status (netstat) changed (new port opened or closed).      7   533
Jul 24, 2022 @ 07:43:52.978  Ossec agent started.                                                      3   503
Jul 24, 2022 @ 07:43:50.246  Ossec agent stopped.                                                      3   506
Jul 24, 2022 @ 07:41:05.886  Listened ports status (netstat) changed (new port opened or closed).      7   533
Jul 24, 2022 @ 07:37:27.530  Process segfaulted.                                                       5   1010
Jul 24, 2022 @ 07:37:27.485  Process segfaulted.                                                       5   1010
Jul 24, 2022 @ 07:37:13.464  PAM: Login session closed.                                                3   5502
Jul 24, 2022 @ 07:37:11.465  PAM: Login session opened.                                                3   5501
Jul 24, 2022 @ 07:37:11.464  First time user executed sudo.                                            4   5403
Jul 24, 2022 @ 07:36:53.439  PAM: Login session opened.                                                3   5501
Jul 24, 2022 @ 07:36:07.253  VirusTotal: Alert - No records in VirusTotal database                     3   87103
```

Figure 6.4: Configuration Assessment

## Customize dashboard

Customize dashboard allowing users to create their own custom dashboards with the information that is most relevant to them.

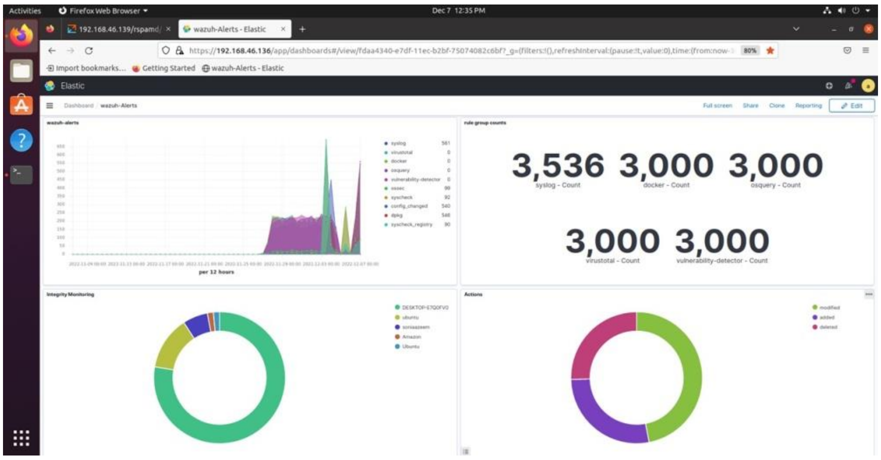


Figure 6.5: Customize Dashboard

## MITRE ATT&CK Framework

The Mitre ATT&CK framework has been integrated to provide users with a thorough understanding of security issues. This integration involves matching alarms produced by security agents to the appropriate Mitre ATT&CK tactics and approaches. Users can use this feature to understand the gravity and repercussions of a security threat and take the necessary countermeasures.

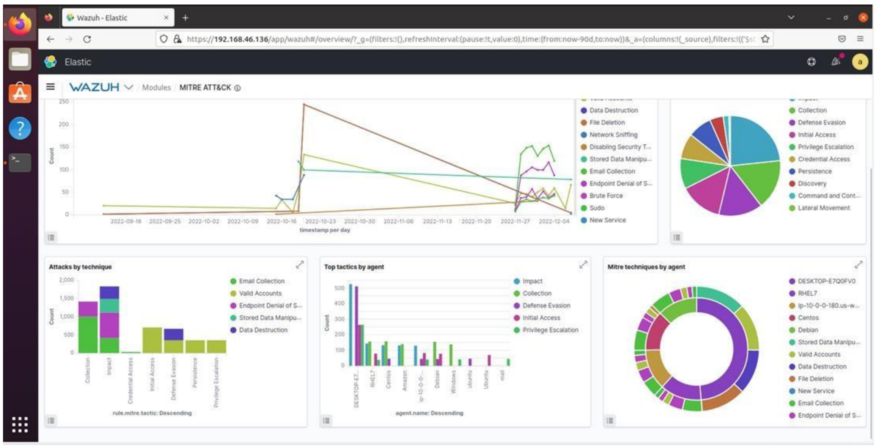


Figure 6.6: MITRE Framework Dashboard

## Comparative Analysis of Literature Papers

Table 6.1: Key Findings from Literature Review

| Reference Paper | Detection and Response Method / Implementation softwares | Limitations |
|---|---|---|
| [26] | GRR (Google Rapid Response) and osquery as open source software frameworks. | Customized settings are necessary to optimize the functionality. |
| [25] | Propose Tactical Provenance Graphs (TPGs) to analyze causal relationships between EDR-generated threat alarms | the limitation lies in the log reduction technique's potential inability to produce false alarms |
| [24] | Utilize ELK stack Integrate Elasticsearch with Virustotal, MISP and add geolocation data using Logstash GeoIP Processor | Their solution does not contain incident response capability and considered limited Operating systems (OS) |
| [1] | Combines AutoEncoder and 1D CNN using semi-supervised learning to detect previously unknown cyber attacks | Limited experimental evaluation and limited threat intelligence platform integration |
| [27] | Used Google Rapid Response (GRR) , Elasticsearch and Auditbeat, | focusing only on Windows operating systems in their experiments, and there is a potential for a high false positivity rate. |
| [28] | Natural language processing (NLP) was applied for spam email detection, utilizing machine learning algorithms including Naive Bayes, K-Nearest Neighbors, SVM, Logistic regression, Decision tree, and Random forest. Logistic regression and NB reaches upto highest accuracy 99 | They collected ready made dataset from kaggle therefore , malicious phishing links were also not considered. |
| Our Product | Our proposed EDR collects data from extensive endpoints , Integrated with Virustotal, Yara , MISP and provide spam filtering capability | Future work is to integrate AI model with spam filtering solution and add more XDR capabilities |

The Table above summarizes our literature review, highlighting the gaps in previous research and showcasing the achievements we have made in our work. Our research successfully led to the development of a robust open-source tool known as Endpoint Detection and Response (EDR). This versatile tool effectively collects data from various operating systems, including Windows, Linux, laptops, servers, and Docker environments, providing comprehensive coverage. Integration with renowned threat intelligence platforms, such as Virustotal, Yara, and MISP, enhances our EDR's malware detection capabilities, enabling swift identification and removal of harmful files that pose a threat to network security.

Additionally, our EDR solution excels in combatting malware, ensuring the protection of critical systems and data. Incorporating advanced features, including XDR capabilities, fortifies our EDR and actively prevents malware infiltration. This enhanced effectiveness is particularly valuable in safeguarding against increasingly sophisticated attacks orchestrated by skilled hackers. In addition to these achievements, we have trained an AI model that accurately classifies spam and legitimate emails with an outstanding accuracy rate of 98.50%. The precision, recall, and F1 score values of the model were also exceptionally high, at 98.99%, 98.01%, and 98.50% respectively, surpassing previous accomplishments in this field. To continuously improve our system, it is essential to regularly update the trained model to adapt to relevant environments and include updated email samples. This proactive approach plays a crucial role in countering the increasingly sophisticated attacks devised by skilled hackers.

## Feature Comparison with Commercial Products

The presented table provides a feature-based comparison between our product and leading commercial products, with the last column showing the features of our product. We have sourced this comparison from the Forrester Wave™ report. The report can be accessed through this link [32]. The evaluations are based on a scale of 0 to 5, with 0 indicating weak and 5 indicating strong performance. Overall, the comparison aim is to highlight the superior qualities of our product in comparison to other commercial developments in the markets

Table 6.2: Feature comparison with commercial products

| S. No | Features | Bitdefender (Gravity Zone) | CrowdStrike (Falcon Insight) | VMware Carbon Black (Carbon Black Cloud) | Microsoft (Microsoft Defender for Endpoint) | Polo Alto Networks (Cortex XDR) | Our Product |
|---|---|---|---|---|---|---|---|
| 1 | Supported systems | 5 | 3 | 5 | 3 | 3 | 5 |
| 2 | Endpoint management | 3 | 3 | 3 | 5 | 3 | 5 |
| 3 | Detection Capabilities | 3 | 5 | 3 | 3 | 5 | 3 |
| 4 | Investigation Capabilities | 5 | 5 | 1 | 5 | 3 | 4 |
| 5 | Response Capabilities | 3 | 5 | 3 | 3 | 3 | 4 |
| 6 | Threat Hunting Capabilities | 1 | 5 | 3 | 3 | 3 | 4 |
| 8 | User experience | 3 | 5 | 3 | 5 | 3 | 4 |
| 9 | Machine Learning Capabilities | 3 | 5 | 3 | 5 | 3 | 3 |
| 10 | Extended Capabilities | 3 | 3 | 1 | 5 | 5 | 3 |

## Conclusion and Future work

In conclusion, our EDR product is a cost effective solution suitable for small to medium-sized organizations. The product offers realtime malware identification and response using Artificial Intelligence (AI), focusing on tools such as VirusTotal and Yara. Additionally, it performs regulatory compliance checks, security configuration assessments, and MITRE ATT&CK based threat hunting. We have also integrated advanced email spam filtering using AI algorithms. This method allows for more accurate evaluation of email content, identification of patterns, and detection of behaviors frequently found in spam emails. Given the strong market competition, our product presents an excellent option for businesses in need of an efficient and budget friendly EDR solution.

Looking ahead, our future research efforts involve integrating a trained AI model with our existing email detection classifier, Rspamd, to enhance system performance and accuracy through advanced algorithms. We will continuously refine the AI model using relevant email data and establish seamless communication mechanisms. Additionally, we aim to expand the capabilities of our EDR solution to the XDR Level, integrating additional security tools and data sources such as network traffic analysis and identity and access management. This enhancement will broaden the scope of protection and improve overall functionality. Furthermore, our plans include enhancing our malware detection capabilities by incorporating more comprehensive feeds, ensuring precise and up-to-date threat detection. By pursuing these avenues of research, we are committed to continuously improving the effectiveness and efficiency of our EDR solution to meet the evolving needs of organizations in an ever-changing cybersecurity landscape.

# Bibliography

[1] C. Hwang, D. Kim, and T. L. and, “Semi-supervised based unknown attack detection in edr environment,” *KSII Transactions on Internet and Information Systems*, vol. 14, pp. 4909–4926, December 2020. 1, 4.1, 6.1

[2] S. Chandel, S. Yu, T. Yitian, Z. Zhili, and H. Yusheng, “Endpoint protection: Measuring the effectiveness of remediation technologies and methodologies for insider threat,” in *2019 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*, pp. 81–89, 2019. 1, 1.1

[3] C. Lábodi and P. Michelberger, “Necessity or challenge-information security for small and medium enterprises.,” *Annals of the University of Petrosani Economics*, vol. 10, no. 3, 2010. 1.1

[4] R. Von Solms and J. Van Niekerk, “From information security to cyber security,” *computers & security*, vol. 38, pp. 97–102, 2013. 1.1

[5] S. Chandel, S. Yu, T. Yitian, Z. Zhili, and H. Yusheng, “Endpoint protection: Measuring the effectiveness of remediation technologies and methodologies for insider threat,” in *2019 international conference on cyber-enabled distributed computing and knowledge discovery (cyberc)*, pp. 81–89, IEEE, 2019. 1.1, 2.5.1.1, 2.5.1.2, 2.5.1.3, 2.5.1.4, 2.5.1.5

[6] T. Liggett, *Evolution of endpoint detection and response platforms*. PhD thesis, Utica College, 2018. 1.1, 2.5

[7] A. Kamruzzaman, S. Ismat, J. C. Brickley, A. Liu, and K. Thakur, "A comprehensive review of endpoint security: Threats and defenses," in *2022 International Conference on Cyber Warfare and Security (ICCWS)*, pp. 1–7, IEEE, 2022. 2.1, 2.2.1, 2.2.2, 2.2.3, 2.2.4, 2.2.5

[8] S. Kim, C. Hwang, and T. Lee, "Anomaly based unknown intrusion detection in endpoint environments," *Electronics*, vol. 9, no. 6, p. 1022, 2020. 2.1

[9] N. N. A. Sjarif, S. Chuprat, M. N. Mahrin, N. A. Ahmad, A. Ariffin, F. M. Senan, N. A. Zamani, and A. Saupi, "Endpoint detection and response: Why use machine learning?," in *2019 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 283–288, 2019. 2.1, 3.3, 3.4.1, 3.4.2, 3.4.3

[10] A. Arfeen, S. Ahmed, M. A. Khan, and S. F. A. Jafri, "Endpoint detection & response: A malware identification solution," in *2021 International Conference on Cyber Warfare and Security (ICCWS)*, pp. 1–8, IEEE, 2021. 2.3, 2.6

[11] S. Brown, "The top 4 ways malware is spread." https://www.snaptechit.com/article/the-top-4-ways-malware-is-spread-2/. 2.4.1, 2.4.2, 2.4.3, 2.4.4

[12] P. R. Brandao and J. Nunes, "Extended detection and response," 2.5, 2.6, 2.7

[13] P. Domingos, "A few useful things to know about machine learning," *Communications of the ACM*, vol. 55, no. 10, pp. 78–87, 2012. 3.1

[14] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamar´ıa, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions," *Journal of big Data*, vol. 8, pp. 1–74, 2021. 3.1

[15] B. Geluvaraj, P. Satwik, and T. Ashok Kumar, "The future of cybersecurity: Major role of artificial intelligence, machine learning, and deep learning in cyberspace," in *International Conference on Computer Networks and Communication Technologies: ICCNCT 2018*, pp. 739–747, Springer, 2019. 3.2

[16] H. Kaur and R. Tiwari, “Endpoint detection and response using machine learning,” in *Journal of Physics: Conference Series*, vol. 2062, p. 012013, IOP Publishing, 2021. 3.3

[17] E. G. Dada, J. S. Bassi, H. Chiroma, A. O. Adetunmbi, O. E. Ajibuwa, *et al.*, “Machine learning for email spam filtering: review, approaches and open research problems,” *Heliyon*, vol. 5, no. 6, p. e01802, 2019. 3.5

[18] A. Krenker, J. Besˇter, and A. Kos, “Introduction to the artificial neural networks,” *Artificial Neural Networks: Methodological Advances and Biomedical Applications. InTech*, pp. 1–18, 2011. 3.5

[19] S. Agatonovic-Kustrin and R. Beresford, “Basic concepts of artificial neural network (ann) modeling and its application in pharmaceutical research,” *Journal of pharmaceutical and biomedical analysis*, vol. 22, no. 5, pp. 717–727, 2000. 3.5, 3.5

[20] A. Chandra, M. Suaib, D. Beg, *et al.*, “Web spam classification using supervised artificial neural network algorithms,” *arXiv preprint arXiv:1502.03581*, 2015. 3.5

[21] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *nature*, vol. 521, no. 7553, pp. 436–444, 2015. 3.5.1, 3.5.1.2, 3.5.1.3, 3.5.1.4, 3.5.1.5

[22] H. Ramchoun, Y. Ghanou, M. Ettaouil, and M. A. Janati Idrissi, “Multilayer perceptron: Architecture optimization and training,” 2016. 3.5.1.1

[23] M. A. Ferrag, L. Maglaras, H. Janicke, and R. Smith, “Deep learning techniques for cyber security intrusion detection: A detailed analysis,” in *6th International Symposium for ICS & SCADA Cyber Security Research 2019 6*, pp. 126–136, 2019. 3.5.1.3

[24] R. Stoleriu, A. Puncioiu, and I. Bica, “Cyber attacks detection using open source elk stack,” in *2021 13th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pp. 1–6, 2021. 4.1, 6.1

[25] W. U. Hassan, A. Bates, and D. Marino, “Tactical provenance analysis for endpoint detection and response systems,” in *2020 IEEE Symposium on Security and Privacy (SP)*, pp. 1172–1189, 2020. 4.1, 6.1

[26] S.-H. Park, S.-W. Yun, S.-E. Jeon, N.-E. Park, H.-Y. Shim, Y.-R. Lee, S.-J. Lee, T.-R. Park, N.-Y. Shin, M.-J. Kang, *et al.*, “Performance evaluation of open-source endpoint detection and response combining google rapid response and osquery for threat detection,” *IEEE Access*, vol. 10, pp. 20259–20269, 2022. 4.1, 6.1

[27] N.-E. Park, Y.-R. Lee, S. Joo, S.-Y. Kim, S.-H. Kim, J.-Y. Park, S.-Y. Kim, and I.-G. Lee, “Performance evaluation of a fast and efficient intrusion detection framework for advanced persistent threat-based cyberattacks,” *Computers and Electrical Engineering*, vol. 105, p. 108548, 2023. 4.1, 6.1

[28] Y. Kontsewaya, E. Antonov, and A. Artamonov, “Evaluating the effectiveness of machine learning methods for spam detection,” *Procedia Computer Science*, vol. 190, pp. 479–486, 2021. 4.1, 6.1

[29] A. Bolla, *Threat Hunting driven by Cyber Threat Intelligence*. PhD thesis, POLITECNICO DI TORINO, 2022. 5.3.3

[30] S. Stanković, S. Gajin, and R. Petrović, “A review of wazuh tool capabilities for detecting attacks based on log analysis,” 5.3.3

[31] S. Brown, “How to run your own mail server.” https://www.c0ffee.net/blog/ mail-server-guide, Aug 2017. 5.3.4

[32] A. Mellen, “The forrester wave™: Endpoint detection and response providers, q2 2022.” https://www.forrester.com/report/ the-forrester-wave-tm-endpoint-detection-and-response-providers-q2-2022/ RES176332, Apr 2022. 6.3