Leveraging Machine Learning and AI to Combat Modern Cyber Threats

Arman Zeidan Bin Mohd Riza , Lee Jennsen , Priscelia Anggani , Arifin Islam Rafeen , Payet Nathaniella Jacqueline Ruth , Dhaavita Sookun , Vashist Sookun , Nur Aliyah Zafirah Binti M Yusri , Lai Jhoon Sern , Lisa Luximon , Mohamed Lulwa Omer , Siva Raja Sindiramutty *

Posted Date: 6 January 2025

doi: 10.20944/preprints202501.0360.v1

Keywords: Artificial Intelligence; Cybersecurity; Threat Detection; Machine Learning; Automated Security Systems

 doi:10.20944/preprints202501.0360.v1

Article

Leveraging Machine Learning and AI to Combat Modern Cyber Threats

Arman Zeidan Bin Mohd Riza, Lee Jennsen, Priscelia Anggani, Arifin Islam Rafeen, Payet Nathaniella Jacqueline Ruth, Dhaavita Sookun, Vashist Sookun, Nur Aliyah Zafirah Binti M Yusri, Lai Jhoon Sern, Lisa Luximon, Mohamed Lulwa Omer and Siva Raja Sindiramutty *

* Correspondence: magan.shiva91@gmail.com (S.R.S.)

Abstract: The modern world of technology focuses on the use of Artificial Intelligence (AI). Artificial Intelligence (AI) is revolutionary and can be considered the powerhouse for the recent advancements seen in the world of technology. The recent rise of sophisticated technological threats is also advancing since Artificial Intelligence (AI) resources can be manipulated for exploited and unethical use which concerns security in terms of both, cyberspace, and physical environments. The key factor here is to discuss the optimisation for the use of Artificial Intelligence (AI) in Cybersecurity. Artificial intelligence (AI) provides a disruptive approach by increasing the efficiency, accuracy, and adaptability of security systems. This abstract examines the integration of AI-based security protocols to combat growing threats. AI systems, such as machine learning and deep learning, can examine large volumes of data in real-time, detecting patterns and abnormalities that humans may overlook. By automating threat detection, AI can respond to situations more quickly and precisely, lowering the likelihood of breaches and cutting response time. Furthermore, AI-driven systems can self-learn, enhancing their threat detection skills as new data and attack vectors emerge. This versatility makes AI a great tool for tackling both known and undiscovered dangers, including cyber-attacks such as phishing, malware, and ransomware. This report focuses on the use of Artificial Intelligence in Cybersecurity. It acknowledges the capability and countermeasures for more advanced Cybersecurity threats and how it differs from the traditional method. Not only does it applaud the use of Artificial Intelligence (AI) in Cybersecurity, but it also provides necessary drawbacks and constructive criticism about the literacy of its uses.

Keywords: artificial intelligence; cybersecurity; threat detection; machine learning; automated security systems

1. Background

Artificial Intelligence (AI) security refers to the techniques and practices that are employed by organizations to prevent their Security Systems from being manipulated or tampered with. There is a need to secure them fully, particularly in the context of 21st-century time-sensitive realms where AI is starting to play a role. Sectors such as finance, healthcare and critical infrastructure. Artificial intelligence is poised to revolutionize as data analysis, decision-making and process automation tools are advanced. However, that also brings in security vulnerabilities which can be utilized by the wrongdoer. A door lock as complex security technology in AI, cannot be underestimated. Because they rely on large amounts of data and complex algorithms, AI systems are susceptible to cyber-attacks. Without adequate security in general, artificial intelligence (AI) can lead to data manipulation, where models are altered for malicious purposes (such as deepfakes), or even the misuse of advanced AI algorithms for automated phishing assaults.

AI systems present distinct obstacles in terms of security because of their complex and large-scale nature. Since AI models require large volumes of data for training and collection, using these

 doi:10.20944/preprints202501.0360.v1

large data sets introduces the potential risk that they will be breached by an attacker. It has been statistically proven that cyberattacks are now more frequent and costly. It was statistically proven that the global average cost to remediate a data breach in 2023 was USD 4.45 million, a 15% increase over three years, IBM. (2023) *Cost of a Data Breach Report 2023*. A second major vulnerability is that of data poisoning, where trainers hack into the training data to provoke arbitrary AI models' behaviours or bias results. Similarly, images with small perturbations added to them could be used to fool machine-learning classifiers, leading the AI systems in sectors such as self-driving cars or medical diagnostics astray — a dangerous possibility. Furthermore, AI models can be threatened with model inversion attacks to obtain personal data of the training data (Alpcan et al. 2017; Ananna et al., 2023), exposing a privacy issue. Overcoming these challenges requires advanced approaches like differential privacy and federated learning, protecting AI models from external threats or violations of confidentiality yet maintaining their performance. The increased presence of AI in cybersecurity underlines the relevance of using these implementations for recognizing and resisting emerging threats, in contrast, an attempt has also been made to ensure that their integrity remains untampered with so as not to become yet another tool exploited by malicious actors.

Its ability to improve cybersecurity defences has considerable advantages to corporations as well as society-wide. And with the capability to analyze real-time data at massive scales, AI can detect emerging cyber threats much quicker than traditional systems. Using models from machine learning helps AI to recognize patterns and anomalies that could be an indication of some hidden or previously unknown threats, for example, zero-day vulnerabilities. Because they target vulnerabilities that haven't even made it onto a long list of issues for developers and legacy systems, these threats can be more dangerous (Azam, Dulloo, Majeed, Wan, Xin, Tajwar, et al., 2023). Moreover, AI adds a new dimension to cyber-resilience where the latter can work on integrated advanced analytics and decision-making capabilities through adaptive defences. AI can, for instance, score security alerts in terms of threat severity to ensure that organizations are addressing the most severe threats and minimizing damage. This automation not only reduces response times but also shows that their AI tools have a self-healing capacity: implementing patches or blocking malicious activities on their own before they lead to widespread disruption.

The use of AI-driven cybersecurity systems, on the other hand, markedly decreases much reliance as far as manual monitoring and response are concerned (a factor reproductively helpful from a point-of-view of operational efficiency). AI can take over repetitive security duties such as network surveillance and log study. AI also strengthens security by detecting vulnerabilities in systems, allowing humans to focus on higher-level issues, rather than repetitive security tasks. Additionally, automation means that there is less of a factor for human error (a common vulnerability in secure systems) and increases system reliability. AI handles advancements in behavioural analytics which allows it to keep a record of behavioural patterns to be able to segment normal user activity but this is also handled by AI as well and detects outliers often indicating stolen credentials or insider threats (Azam, Dulloo, Majeed, Wan, Xin, & Sindiramutty, 2023). As a result, organizations can drive improvements in security whether they are on the server side or even at internal levels – all without having any effect on user experience. AI-powered authentication methods such as biometrics, for instance, deliver zero-friction and secure access that automatically adjusts to user behaviour in time.

In addition, AI aids in regulatory compliance by using automation to monitor data processing and ensure that all is done according to legal standards (GDPR, HIPAA). This kind of automation works to ensure that organisations will always be compliant which prevents them from steep fines and maintain customer trust. Given the scale and complexity of today's IT environments, it is also necessary for AI to be able to orchestrate security mechanisms reliably. Because of the increased complexity and number of systems used by organizations as they grow, AI's ability to work well with other existing cybersecurity tools such as Security Information & Event Management (SIEM) platforms is also very important. This delivers real-time threat intelligence that enables full and automated defence in depth across networks, endpoints and cloud services (Azam, Tajwar, Mayhialagan, Davis, Yik, Ali, et al., 2023). The increased AI adoption across the board reveals a

 doi:10.20944/preprints202501.0360.v1

critical necessity to secure it fast. The security of AI systems protects proprietary information and can build confidence in infrastructure that is only possible with artificial intelligence. Improving AI security bolsters trust in digital systems and underpins the secure development of AI technologies for future use cases.

2. Discussion on How Security-Related Technology Works

i. Component

a) Predictive Analysis

A common feature included in most security systems nowadays, it uses algorithms and data techniques to be able to predict and forecast any potential cyber threats by predicting and analysing patterns of past attempts and preventing them before they occur.

To integrate this into AI security, AI predictive simulators and machine learning models must have data on different types of threat intelligence and must be up to date on all of the latest cyber threats and their behaviours through sources such as statistical algorithms network traffic, logs or behaviour(Azam, Tan, Pin, Syahmi, Qian, Jingyan, et al., 2023). They must also use data from other security sources such as firewalls, WPAs and antivirus to understand every type of defence used against cyber threats. This means the model must be able to always monitor network traffic, check user activity and logs, etc. This can simulate a cyberattack so that responders to the cyberattack know how to effectively deal with it.

This analysis is essential to any security system as its ability and insights can allow companies to gain the advantage on cyber criminals at an early stage, since AI is more effective in going through threat and behavioural data more accurately than manually, it allows the predictive model to be more secure and can help companies improve their awareness of cyberattacks as they can pre-emptively anticipate threats and dangers before they can take any real advantage on the company.

b) User Authentication

This is a traditional feature included for most devices that have constantly been upgraded and advanced as time goes on, this ensures access to only authorised users with different levels of access can access certain devices, accounts or data, this can include biometric authentication, identification, passwords and much more.

To integrate this, AI learning machines must be input with data on any user with authorised access to be able to match their data with the knowledge that is in their database. This AI will store data on different characteristics of certain users such as their fingerprint, facial data or even user behaviours and it will use classification algorithms to verify the biometric information belongs to the correct user (Hussain et al., 2024). This allows for restricted access to certain individuals; therefore, the required data will be difficult to replicate or forge and there is less risk of human error as it can replace passwords which are vulnerable to human error, therefore AI can ensure an easy process of verifying users by analysing specific authentication behaviours.

c) Secure Software Development

When in the process of developing software, there will tend to be unpredicted issues and human error which can lead to cyber threats, therefore regular security testing, using secure protocols and threat modelling as part of the development of the system before the system is fully published to ensure there is less risk of cyber threats in the early stages of the system life cycle.

This can be integrated by using AI tools to prevent any code vulnerabilities by scanning it during the early stage of development. Threat modelling using AI identifies patterns and behaviours, allowing it to analyse the code and predict attacks and this can be done while the system is in development and not even implemented. Identifying threats early in the system development cycle allows it to have better longevity and reduces flaws before fully implemented.

 doi:10.20944/preprints202501.0360.v1

d) Machine Learning

Traditional detection systems introduced are slow, inefficient and no longer effective at dealing with new and modern threats. Albeit still capable of dealing with any current or traditional threat. Security systems must always be one step ahead of any potential and future threats. As a result, the current system requires an Intelligent and effective system to deal with consistent and relentless intrusion attempts. Artificial intelligence's capability to undergo machine learning is a perfect example of how current or future security systems can deal with this ever-evolving blight.

The need for security systems to evolve is also due to the demand for more sophisticated and adaptable security solutions driven by AI and ML has arisen because of the shortcomings of conventional cybersecurity techniques as well as the growing volume, velocity, and variety of cyber threats (Jun et al., 2024). AI-powered cyber security solutions can provide several significant advantages by utilizing ML algorithms' capacity to learn from enormous volumes of data and get better over time. Katiyar et al. (2024)

Machine learning has become more and more popular. It is now an essential component of the industry, having left the labs. Machine learning is used by Google, Microsoft, and Facebook to enhance user experience, link users with apps, and create new personal relationships. Machine learning is also used in the field of cybersecurity. It has the power to drastically alter the dynamics of cyberspace. Attackers feast on sensitive data and establish beachheads for upcoming strikes. After using up all the data on the network, they move on to other networks whose access they have obtained via the breach and offer network access as a service. This issue can be addressed with novel machine-learning approaches. It is more dependable, lowers costs, and boosts efficiency. Kaur and Tiwari (2021)

e) Advance Threat Detection

Firstly, utilizing previously undiscovered attack patterns and adaptively learning and improving from network traffic data, this method of approach can improve NIDS. To divide new dangers according to their attributes, supervised learning algorithms, including decision trees, SVMs, or neural networks, can be trained on labelled datasets comprising malicious and legitimate network traffic. Avoiding the usage of labelled data, unsupervised learning techniques like clustering or anomaly detection may and could be utilized to find odd patterns or differences from typical network activity.

These techniques are especially effective in identifying and detecting any insider threats or new attacks that could not have the same signatures or footprints. Deep learning models can learn intricate representations of network traffic patterns as they have also shown promise in network intrusion detection (Manchuri et al., 2024). Examples of these models are autoencoders and recurrent neural networks. Autoencoders, for instance, may be taught to reconstruct typical network traffic, and abnormalities can be identified by observing departures from the learnt reconstruction. Katiyar et al. (2024)

f) Endpoint Security

Researchers pave the way for the focus on User Behaviours Analytics (UBA). When developing a security system against any external threats, defending parties must also consider the plausibility of internal breaches. Most current organizations are unfortunately unaware of any active hostile threats in their very own domain or network. Albeit internal threats may cause even more severe damages. The standard security protocol is completely vulnerable to such a rule of engagement, with endpoint security this threat may be mitigated but not completely taken down. Due to the nature of the challenge in detecting this threat.

A.Chuvakin used the term endpoint detection and response (EDR) in 2013. EDR is an endpoint security system that aims to be sufficiently capable of providing real-time protection. It works in a way where data from endpoints is gathered by endpoint detection systems, which are then stored

 doi:10.20944/preprints202501.0360.v1

and processed in a central database. Afterwards, in real-time, these gathered data are connected to identify any abnormalities or anomalies in the host's behaviour. As a result, EDR systems warn users and emergency response teams about online risks. Kaur and Tiwari (2021)

g) Vulnerability Management

Standard security vulnerability management is conducted via security experts, these roles could be substituted by AI to reduce the need for manpower. With Deep learning approaches, the results of its use have been most desired in assessing vulnerabilities(Ravichandran et al., 2024). This task includes not limited to vulnerability severity evaluation, code clone detection and vulnerability detection. The API tested by previous researchers are graph neural network (GNN), transformer and the widely used Large Language Model (LLM). The report from more than 90% of previous research has dictated that the result from the use of previously named models has completely outclassed the standard static analysers.

The main factor in understanding security flaws is more on to create awareness. Each employee or staff must be trained as needed and suggested by the flaw detected by the AI system. When staff are trained and tasked accordingly it will almost completely demonstrate effectiveness in preventing security problems. AI can also be used to smartly repair vulnerabilities in the system. It will suggest appropriate approaches to fortifying vulnerabilities in our system. These AI models will greatly aid cyber security practitioners in dealing with their tasks as AI is trained with millions or even billions of texts including codes and tokens. Cooperation with AI can help security experts mitigate human limitations on a broader scale. Wan et al. (2024)

ii. Process

In today's evolving world cybersecurity has become essential as the digital landscape advances. Security must be robust and swift to prevent any unwanted threats. Artificial intelligence has been developed to protect web-based activities such as business data including online transactions as well as authentication passwords. The rise in remote work since COVID-19 necessitates abundant security measures to defend against threats. The network is crucial for all of this to function. AI security enhances threat detection which surpasses traditional methods by integrating with SDN. Software-defined networking (SDN) merges general networking devices with a software-based controller providing centralized network management to system operators. The controller directs networking functions by delivering instructions to forwarding devices that integrate traditional networking.

SDN programmability enables the extension of AI-based protection functionality without the need for dedicated hardware devices offering a robust and diversified solution compared to traditional approaches. SDN can be used for anomaly detection and threat mitigation. Reliable, secure, and effectively monitored networking is achieved through a programmable protection system capable of identifying anomalies including heavy traffic with fully automated detection as provided by SDN. Ahmed and Kim (2017) stated that statistics are automatically collected on all SDN-enabled devices for anomaly detection, offering unique benefits in terms of traffic statistics. This approach eliminates the need for specific devices for sampling-based detection reducing costs and simplifying network topology. Possessing statistical data allows for classifying data flows and applying filtering behaviours such as blocking traffic. Faustino et al (2016) developed ATLANTIC which is a framework combining information theory and machine learning to calculate deviations in flow table entropy and automate mitigation. Data collection on networking devices can be challenging as it is aggregated and visible only at lower layers. This issue is mitigated by forwarding traffic through the SDN controller. Qiumei developed an SDN-based firewall that uses supervised machine learning to classify traffic then achieving 96.79% detection accuracy with an average latency of 0.2 ms.

The inherent nature of SDN in software processing renders it slower when compared to traditional hardware-accelerated networking devices. The capabilities of SDN go beyond mere software processing leveraging proactive flow insertion and utilizing hardware-accelerated tables for data storage to mitigate these delays (Seng et al., 2024). Numerous researchers have demonstrated

 doi:10.20944/preprints202501.0360.v1

this by implementing firewall rules directly on networking devices achieving a 23-fold performance improvement over conventional software-based firewalls. This advancement showcases SDN's potential to transcend its initial limitations. SDN controllers can further eradicate constraints related to limited capacity and insufficient networking including optimizing traffic management and enabling automated threat detection and mitigation when integrated with AI-based protection functionalities. This synergy ensures improved performance making networks more resilient and adaptive against evolving security challenges. The proactive features of SDN coupled with AI-driven enhancements create a robust and diversified solution that seamlessly integrates innovative security measures (Sindiramutty et al., 2024). This amalgamation ensures that networks are protected and optimized for performance and efficiency, marking a significant step forward in the evolution of network security. It makes networks more reliable efficient and resilient than ever before.

The implementation of AI in security contexts has assisted in transformative capabilities for predictive analysis, real-time threat detection, and adaptive response strategies. Leveraging a combination of supervised learning, unsupervised learning, and reinforcement learning (RL) techniques, AI security technologies enhance cybersecurity systems by continuously learning from new data, identifying patterns indicative of threats, and making data-driven decisions to secure networks and systems. However, the intricate processes underpinning these AI technologies are accompanied by significant technical challenges, necessitating a thorough examination of both their strengths and limitations.

a) Data Collection and Preprocessing

AI security technologies initiate their operations with extensive data collection, which encompasses diverse sources such as:

- **Network Traffic: Monitoring data packets for unusual patterns.**
- **Endpoint Behavior: Gathering information on user activity and system performance.**
- **Threat Intelligence: Compiling data from known threats, vulnerabilities, and exploits (Tan et al., 2020).**

The raw data undergoes preprocessing to enhance its quality and usability. This includes techniques such as normalization, dimensionality reduction, and feature extraction (Abolfathi et al., 2021).

b) Model Training

The core of AI security systems lies in the training of machine learning models. Two primary paradigms—supervised learning and unsupervised learning—are employed in cybersecurity contexts:

- **Supervised Learning: In supervised learning, models are trained on labelled datasets to predict outcomes based on input features (Nguyen et al., 2018).**
- **Unsupervised Learning: Unsupervised learning models analyze unlabeled data to identify patterns or anomalies (Kim et al., 2019).**

c) Pattern Recognition with Historical Data for Predictive Threat Detection

AI models trained on large datasets are invaluable in learning patterns that signal potential security threats. Supervised learning, especially using Deep Neural Networks (DNNs), is often employed in this regard. DNNs can analyze extensive logs of network traffic to detect deviations that may signal an intrusion (Li et al., 2019). By observing features like packet headers, IP addresses, and data volumes, DNNs establish a baseline of normal activity, flagging deviations that may signify malicious attempts such as DDoS attacks or unauthorized access to restricted resources. DNNs detect anomalies by comparing observed patterns against learned patterns, often through a loss function that penalizes deviations from expected network behaviour (Alzhrani et al., 2016). This approach

 doi:10.20944/preprints202501.0360.v1

enhances intrusion detection systems (IDS) by alerting cybersecurity teams to anomalies, thereby enabling proactive responses to potential threats.

d) Real-Time Cyber-Attack Detection and Response with Reinforcement Learning

Real-time detection of cyber-attacks is one of the most dynamic applications of AI in security. Reinforcement Learning (RL), a paradigm in which algorithms learn from experience by interacting with their environment, has shown promise in adaptive threat response systems. RL algorithms can detect and respond to advanced persistent threats (APTs) by learning optimal defence actions based on historical data (Gualberto et al., 2020). Here, the RL agent observes network activity and identifies potential threats by balancing the trade-off between exploration (detecting new threats) and exploitation (taking the known best action).

e) Anomaly Detection with Unsupervised Machine Learning

While supervised methods require labelled data, unsupervised learning can detect anomalies without predefined labels, making it invaluable for discovering novel threats. Clustering algorithms such as K-means and Autoencoders are widely utilized in anomaly detection applications. Autoencoders, a form of neural network designed to learn compressed representations of data, can highlight atypical user behaviour by analyzing the reconstruction error between observed and expected patterns. In practice, this approach has been applied in endpoint protection systems to detect irregular user activity, such as unauthorized access attempts. The K-means algorithm, a popular clustering technique, partitions data into clusters to uncover hidden patterns(Sindiramutty, Jhanjhi, Tan, Khan, Shah, & Manchuri, 2024). For anomaly detection, K-means minimizes the within-cluster sum of squares to categorize data points based on similarity. Anomalies are detected when a point lies far from any cluster centroid, often indicating malicious activity or insider threats (Wu et al., 2021). This process empowers AI-driven Security Information and Event Management (SIEM) tools, improving their capacity to identify potential breaches without needing extensive labelled data.

f) Malware Classification Using Supervised Learning and Generative Models

The classification and identification of malware are critical aspects of AI-based cybersecurity. Supervised learning algorithms, including Support Vector Machines (SVM) and Random Forests (RF), categorize malware based on attributes extracted from executable files. RF classifiers rely on ensemble learning principles to evaluate multiple decision trees, improving the robustness of classification tasks. Each decision tree is generated based on subsets of features and samples, mitigating the risk of overfitting (Vinayakumar et al., 2018; Sindiramutty, Jhanjhi, Tan, Khan, Shah, Yun, et al., 2024). Moreover, Generative Adversarial Networks (GANs) have recently been leveraged for malware detection by generating realistic malware samples for model training. In this adversarial framework, the generator and discriminator networks compete, enhancing the discriminator's ability to distinguish between genuine and malicious inputs. By simulating diverse attack scenarios, GANs contribute to enhanced malware classifiers that adapt to emerging threats.

g) Leveraging NLP in Phishing Detection and Email Security

Natural Language Processing (NLP) is applied in phishing detection systems to parse and interpret email content, detecting phishing attempts based on linguistic and semantic patterns. Techniques like Bag-of-Words (BoW) and word embeddings enable systems to evaluate textual features and assess the likelihood of phishing attempts. Suspicious language or unusual sentence structure in emails is flagged for further review. BoW represents text as a collection of word frequencies, while embeddings transform words into vector spaces to capture contextual similarities (Alotaibi et al., 2021). An NLP-based phishing detection model can assess email authenticity by calculating cosine similarity scores between legitimate and suspicious emails, flagging cases where similarity exceeds a defined threshold.

 doi:10.20944/preprints202501.0360.v1

iii. Threats

Artificial intelligence is swiftly revolutionising the cybersecurity landscape by harnessing its capacity to analyse vast amounts of real-time data and uncover hidden vulnerabilities. This innovation offers sophisticated and promising solutions to combat the growing complexity of various cyber threats. Through AI, we can proactively mitigate risks and strengthen security defences in an evolving digital world. The following section outlines various threats that AI can detect and address, particularly those to which we are most vulnerable.

a) Malware Attack

Malware, which is also known as malicious software, is designed by cybercriminals to infiltrate systems with the intent of stealing data, jeopardising operations, or disrupting normal functionality. This category includes threats such as viruses, worms, Trojan viruses, spyware, adware, and ransomware (CISCO, 2024; Sindiramutty et al., 2024). Malware often tricks users into clicking links, pop-ups or installing software from untrusted sources.

A computer virus is designed to reproduce itself and spread from one file or program to another or across networks. On the other hand, Trojan horses disguise themselves as legitimate programs but cause damage upon activation. Unlike viruses, their main goal is to steal files or passwords rather than replicating themselves. Meanwhile, computer worms will replicate themselves and spread through a network.

Spyware, another form of malware, installs itself into systems to gather and transfer sensitive information without the user's knowledge. For instance, Pegasus spyware could extract data from mobile devices while users would be unaware (Greengard, 2023). Ransomware, in contrast, locks devices or an entire system, encrypts data and demands a ransom for its recovery. The infamous WannaCry ransomware which occurred on 12 May 2017 had affected over 300, 000 systems in 150 countries and a payment in Bitcoins was demanded by the attackers (Akbanov, Vassilakis and Logothetis, 2019; Sindiramutty, Tan, Shah, et al., 2024).

b) Phishing Attack Detection

Phishing involves sending fraudulent communications through email or text messages that appear to be from reputable and trusted sources. The attacker's goal is to retrieve sensitive and personal information such as usernames, passwords, credit card numbers, bank account details, or other valuable data with the intent to use or sell the stolen information. Deceptive messages are created to trick victims into revealing their data (CloudFlare, 2023). For instance, in January 2024, Darktrace detected a phishing attempt which used Dropbox, where employees of a customer received a seemingly legitimate email from a Dropbox address (Traill, 2024).

Artificial Intelligence (AI), particularly through Deep Learning (DL) and Machine Learning (ML), plays a pivotal role in enhancing malware detection by identifying suspicious behaviours in real time (Djenna et al., 2023). Companies such as Darktrace use machine learning and deep neural networks to detect similar threats (Darktrace.com, 2024). AI can also recognise patterns and features that differentiate phishing attempts, such as unusual URLs, suspicious email content, and doubtful user behaviour. ML classifiers such as Decision Trees and Random Forests accelerate the detection process by categorising data (Basit et al., 2020).

c) Denial Of Service (DOS)

Denial of Service attacks are becoming an increasingly popular cyber threat, posing substantial risks to various online systems and services. These attacks usually occur when malicious attackers overwhelm a system, network, or service with excessive traffic or requests, ultimately causing system failure or crashes that render the system incapable of processing legitimate traffic and inaccessible to legitimate users. This leads to severe consequences such as extended downtime, and significant operational disruptions, often causing considerable financial losses (CISA, 2021; Sindiramutty, Tan,

 doi:10.20944/preprints202501.0360.v1

& Wei, 2024). One good example of a major DoS attack that occurred was in October 2016 when the attack on DNS provider Dyn flooded its servers with traffic from a botnet of over 500,000 compromised IoT devices infected with Mirai malware. This overwhelmed Dyn's servers, causing widespread outages on major sites like Twitter, Amazon, Spotify, Airbnb, Paypal and Netflix (Frankenfield, 2020).

However, AI has emerged as a powerful tool offering solutions to mitigate DoS attacks. AI leverages cutting-edge algorithms that significantly enhance traffic filtering and provide real-time analysis, surpassing the capabilities of traditional methods (Luffy, 2024). Through these sophisticated filtering techniques, AI can examine certain factors such as where the traffic is coming from, how often requests are made as well as the type of data being sent and carefully assess them to verify if there are any inconsistencies and precisely distinguish between legitimate and malicious traffic offering more accuracy in detecting and mitigating DoS threats. Legitimate traffic usually follows consistent patterns, unlike malicious traffic which is typically characterised by abrupt and substantial increases in volume that deviate from normal user traffic and activity (Kaiyue, 2024).

d) Intrusion Detection

Intrusion refers to any unlawful entry or breach into a system or network to obtain sensitive information or inflict damage. With the advancement of technology and networks, the frequency and complexity of system and network intrusion have escalated significantly, necessitating the need for more advanced and robust security solutions to mitigate them. This is where AI-driven Intrusion Detection systems step in, offering a cutting-edge and adaptive mitigation approach. An Intrusion Detection System is a sophisticated cyber security tool used to detect and produce alerts in any instance of potential intrusions (Waheed et al., 2024). AI-driven IDS offers more promising results compared to traditional IDS. Through the integration of machine learning and Deep learning techniques in AI-based IDS, the detection of any malicious attacks can be detected in a short amount of time making it more effective (Insights2Techinfo,2024). The implementation of AI in Intrusion Detection Systems provides several important benefits.

Firstly, AI-based IDS can improve threat detection by performing real-time analysis of vast volumes of data from many different sources at very high speeds and identifying any irregularities or intricate patterns that conventional rule-based systems or human analysts may miss out on. This contributes to faster detection and response to potential threats compared to normal IDS systems. Additionally, an AI-based IDS uses adaptive learning to enhance its detection capabilities over time. Adaptive learning allows AI models to constantly absorb new data and information allowing AI-based IDS to effectively combat any new emerging attack vectors without the need for any manual upgrades. This continuous learning process enables the IDS to expand alongside the growing threat landscape, identifying harmful behaviour patterns, especially those that deviate from established norms. Consequently, the system's ability to identify complex threats and zero-day assault improves.

To conclude, artificial intelligence is reshaping the cybersecurity landscape more efficiently by enabling faster, more accurate threat detection and response. AI-driven solutions are critical in combating today's increasingly sophisticated cyber threats, ranging from malware detection to denial-of-service attacks. However, AI still faces challenges in tackling specific threats such as zero-day attacks, insider threats, and advanced persistent threats (APTs). Zero-day attacks exploit unknown vulnerabilities in software or hardware, making detection difficult without prior knowledge. Additionally, supervised machine learning models may miss unfamiliar patterns, and anomaly detection systems can trigger false positives for minor changes in an organisation's behaviour. Insider threats, which are either malicious or accidental, may pose a challenge to AI-based security tools (Darktrace.com, 2024). Nonetheless, as AI technology progresses, its incorporation into cybersecurity will be vital for developing resilient and robust defences in an ever-evolving digital world.

iv. Example Security-Related Technologies

 doi:10.20944/preprints202501.0360.v1

a) AI-Powered Intrusion Detection and Prevention Systems (IDPS)

Introduction

While modern cyber-attacks continue to be sophisticated, conventional measures for detection, including firewalls and antivirus, are no longer sufficient to detect APTs or zero-day attacks. AI-driven IDPS has emerged to become the next generation of technology in neutralizing such security threats. Propelled by machine learning and AI algorithms, such systems detect, analyze, and respond to malicious activities in real time by initiating the necessary responsive actions on their own as an active defence mechanism (Sharma et al., 2024; Wen et al., 2023).

What makes an IDPS special? AI-powered IDPS are unique because of their real-time autonomous defence capabilities. Most security tools depend upon predefined rules and signature-based systems, however, an AI-driven IDPS deploys machine learning algorithms to detect, trace, and respond to any threat in real-time (Hall, 2024). They can identify zero-day attacks and APTs that conventional systems cannot detect or identify. This is because of the continuous learning features of AI models, which help these systems change and adapt to emerging new threats.

Why IDPS over other AI Security Technologies? The reason for deploying AI-powered IDPS lies in its holistic and proactive role in cybersecurity. Whereas all the other AI security technologies focus on a very key area, say endpoint protection, malware detection, or intrusion defence mechanism. It performs not only the detection of anomalies but also threat classification and automated response actions. This integrated approach reduces the window of vulnerability and minimizes the impact of potential breaches (Hall, 2024).

Detailed Explanation Compared with Other AI Security Technologies

1. Endpoint Detection and Response (EDR): Focuses on detecting and responding to threats that come into endpoints such as desktops, laptops, and mobile devices. Unlike IDPS, which monitors the entire network, EDR is localized, making it a good candidate for finding threats that target the endpoints only (Sohail, 2024).
2. Security Information and Event Management (SIEM): IDPS is proactive, meaning it's always monitoring network traffic in real-time and actively preventing attacks by stopping malicious activity as it happens (Hall, 2024). SIEM, on the other hand, tends to be more focused on log analysis: aggregating and correlating events after they have occurred to help identify threats retrospectively (Kidd, 2023).
3. AI-Powered Antivirus: These are traditional antivirus powered up with AI, which detects and mitigates known threats by basing the detection logic on malware signature and behaviour pattern identification. They do not offer network-wide anomaly detection and automated response capabilities of IDPS.

b) Practical Application (Example of Real-Life Usage)

Dark Trace's Enterprise Immune System

Darktrace is among the leading cyber security solutions that apply artificial intelligence and machine learning in real-time threat detection and response. It’s well-recognized by its Enterprise Immune System, which mimics the human immune system in finding and neutralizing anomalies across networks, devices, and cloud environments (DarkTrace, n.d.). Though considered a powerful tool, like any other technology, Darktrace does have advantages and drawbacks.

Advantages of Using Darktrace

1. Unsupervised Learning in Self-Learning AI

 doi:10.20944/preprints202501.0360.v1

Adaptive Threat Detection: Darktrace AI does not require any pre-existing knowledge of threats, like signatures. Instead, it learns the normal behaviour of users, devices, and systems, hence discovering new threats and small-size anomalies that can't be found with signature-based systems (Geijsendorpher, 2024).

Customization: The AI learns what is "normal" for the environment of each organization while continuously adapting to the changing nature of the network (PeerSpot, 2024). The self-learning nature of Darktrace means that its relevance and accuracy remain intact in changing IT environments.

2. Autonomous Response with Antigena

Real-time Threat Response: Antigena could respond to the threat autonomously, without the intervention of humans. Real-time intervention has become necessary in mitigating fast-moving attacks like ransomware or distributed denial-of-service (DDoS) attacks, where mere seconds may determine containment in these types of events (PeerSpot, 2024).

Granular Response Options: Antigena provides a range of responses from slowing down malicious actions to full isolation of compromised systems, enabling minimal damage with maximum business continuity.

3. Comprehensive Network Visibility

Darktrace provides network-wide visibility, from the monitoring of all devices to systems and user connections (DarkTrace, n.d.). This holistic approach does not exclude any part of the network from its view. This has been critical in finding those threats that bypass traditional security mechanisms like firewalls or antivirus systems. DarkTrace extends its monitoring to cloud environments and IoT devices, normally targeted during attacks because of their weak security (DarkTrace, n.d.).

4. Early Threat Detection

Darktrace allows for the detection of anomalous behaviour early, which means that the potential threat can also be detected early (DarkTrace, n.d.). This early warning capability is valuable for detecting advanced persistent threats (APTs) or insider threats that are difficult to detect with conventional tools.

5. Reduced Human Intervention

With Darktrace's AI performing much of the threat detection and response in an autonomous manner, the system significantly reduces the load on security teams (Mukuna, 2024). It is especially very helpful for those organizations that struggle with a shortage of staff or even a skills gap in cybersecurity.

c) Drawbacks of Darktrace

v. False Positives

Anomaly-Based Detection: While the AI of Darktrace does a very good job of finding anomalies, the system may on occasion misinterpret normal variations in network traffic as threats, leading to false positives (Hassen, 2023). This could lead to unnecessary work on security teams and reduce the response times in case of a valid threat.

Learning Period: During its initial deployment, Darktrace requires a learning period to set up the benchmark of normal behaviour. During this phase, the system might generate more false positives as it refines its understanding of the environment.

vi. High Cost

 doi:10.20944/preprints202501.0360.v1

Premium Pricing: Darktrace is a high-end solution, and with advanced capabilities, it comes with quite a price (Khashab, 2022). This may be out of league for smaller businesses or organizations on tight budgets.

Continuous Maintenance Cost: Besides its one-time cost, the recurring cost for the maintenance and updates of Darktrace's AI model, and the potential licensing cost of the Antigena module, will be added up (Hassen, 2023).

vii. Complexity and Learning Curve

User Training: While powerful and effective, this AI-driven approach can be a little more complex to work with for some users. The company might therefore have to invest some costs in specialized training for their security teams to use all the features and interpret the data correctly (Zuniga, 2024).

Managing Alerts: The notifications from Darktrace, especially for large complex networks, are extremely difficult to manage without a quality security operation team. Even with the autonomous response capability of Antigena, human oversight and management are still necessary in order to ensure that threats are managed and prioritized accordingly (Kaushik, 2024).

viii. Limited Transparency in AI decisions

Manual Verification: Since the AI is autonomous, some organizations may yet require human intervention to verify certain decisions made by Darktrace, which is quite time-consuming (Zuniga, 2024).

ix. Dependence on Data Quality

Data Dependency: The capability of Darktrace depends on the quality and completeness of data analyzed by the technology (Darktrace, n.d.), and therefore gaps in the network coverage and low-quality data meant the AI could not detect the threats.

Incomplete visibility: Darktrace is widely visible (Khashab, 2024), but in cases where some parts of the infrastructure are not connected to the system due to cloud services or remote endpoints, it will limit visibility by Darktrace (Lethoba, 2024).

x. Integration with Existing Security Tools

Compatibility Issues: For some organizations, there could be an integration problem with their own security toolsets such as SIEM platforms or firewalls. While Darktrace has a lot of integrations, sometimes the software just does not fit well with the security ecosystem of any particular organization (Kaushik, 2024).

Reliance on Complementary Tools: Darktrace works well in complementarity with other security tools for full coverage. It simply means that organizations may still have to invest in other security solutions on top of Darktrace. Table 1 shows a Summarization of the Advantages and Disadvantages of Darktrace.

Table 1. Summarization of Advantages and Disadvantages of Darktrace.

Strengths	Weaknesses
Self-learning AI that adapts over time.	False positives can overwhelm security teams.
Autonomous threat response with Antigena.	High costs may be prohibitive for smaller businesses.
Comprehensive network-wide visibility.	Complexity and learning curves for users.

 doi:10.20944/preprints202501.0360.v1

Early detection of emerging threats.	Limited transparency in AI decision-making.
Reduces human intervention needs.	Dependence on data quality for accuracy.

3. Discussion on the Impact

i. Benefits

When security-related technologies are integrated into Artificial Intelligence (AI), not only do multiple sectors benefit from it, but new avenues to improve cybersecurity efforts are also implemented. Improved Threat Detection: One of the most significant advantages is that AI systems can assist in detecting threats better. AI can look for patterns and other signals in the tremendous number of data points it processes at a given time to help detect threats. Especially in an age when cyber threats are becoming more sophisticated and omnipresent. Traditional security systems are less efficient in handling the scale of data that is being generated, ultimately leading to slipping through threats. Darktrace, for example, is a provider that creates an "immune system" algorithm using machine learning to learn and adapt to systems as they tend to change in undesirable ways. This proactive functionality creates an additional layer of security that expedites threat discovery and maintains the quality of detection, giving companies timely insight into malicious activity to allow remediation attempts before attacks grow large enough to turn into full-fledged breaches.

Along with improved threat detection, AI can also be used to automate incident response, reducing human error until the compromise is addressed. The capacity of AI to automate typical procedures like log analysis, alert generation or initial threat assessment is especially useful for Security Operations Centers (SOCs). These skilled analysts are now able to focus on more intricate and crucial analysis, thereby increasing overall operational efficiency. For instance, IBM Watson for Cyber Security automatically analyzes security incidents and enables the team to focus on high-priority threats rather than being tied down by repeated tasks. Further, predictive analytics predicts potential security threats based on previous data and trends for an organization. A study of previous incidents and user behaviour patterns can build stronger defences and prepare better for future attacks. So tools such as Splunk help organizations to recognize their vulnerabilities in real-time and take steps against that risk before it can occur. And lastly, powerful AI technologies such as federated learning and differential privacy provide more sophisticated reinforcement of data security. How to Arm AI Models with Advanced Privacy Techniques: These techniques allow for the learning of such models from decentralized data sources in a privacy-preserving way, providing an opportunity for a more secure cybersecurity posture — one that increases security while preserving user confidentiality.

ii. Limitations

Nonetheless, AI also has many shortcomings that hinder its use in cybersecurity. One concerns the ability of the AI models to remain robust when they are subjected to adversarial perturbations. This, in turn, results in incorrect classification or prediction and further leads to security issues such as false positives or false negatives. Some of the threats that tread this behaviour are cybercriminals who try to bypass the AI mechanism and utilize the data-centric operation of AI by distorting the inputs mildly so that the model is not able to pick them up. Besides, the recent advances in generative AI also allow attackers to produce realistic personalized deepfakes for social engineering schemes aimed at users who trust familiar faces or voices. All these weaknesses emphasize the need for further model retraining and model improvement to protect the AI system against adversaries (CISO Collective, 2023; Fortinet, 2023).

 doi:10.20944/preprints202501.0360.v1

The rise in the popularity of AI technology around the world has pushed the cybersecurity skills gap even further. Even though AI solutions are designed to create relief for human resources by automating some activities, AI technology's advancement is making it necessary to have more skilled people who can manage, interpret and update these tools. The lack of knowledgeable AI Cybersecurity practitioners increases the chances of abuse and wastage of such practices. In addition, even the least management might completely render such AI tools worthless or create new security flaws which will be worse to the intended functionality of the AI tools. Addressing this skills gap entails having organizations committed to funding specific training programs with the assistance of AI-based responsive instruments taking up the more straightforward processes and experts being left with the most irksome processes (World Economic Forum, 2024; Pluralsight, 2023).

Data privacy is another major limitation, especially for large language models (LLMs) that require substantial datasets to perform optimally well. AI systems trained on such data are vulnerable to such exposure, which is potentially embarrassing to privacy violations and legal non-compliance. Laws such as the GDPR and CCPA still raise pressure on organizations, which use AI technologies to safeguard personal data in their operations. Furthermore, AI is data-intensive and therefore could be significantly prejudiced by bias in the data since biased models are likely to generate bias in their output and hence affect decision-making. To overcome such limitations strict data management practices and compliance must be observed to ensure non-disclosure of sensitive information causing data breaches and individual privacy infringement (Fortinet, 2023).

iii. Future Potentials

Artificial Intelligence offers transformative potential for the cybersecurity sector, particularly in Intrusion Detection. AI provides adaptable, proactive defences, enabling cybersecurity systems to detect, assess, and counteract evolving cyber threats in real-time. In the face of the rapidly evolving digital threat landscape, this cutting-edge technology enhances resilience, creating a flexible, dynamic defence against complex digital threats.

AI offers future potential in various ways including:

a) Predictive Analysis for Threat Forecasting

Predictive analysis involves using statistical algorithms and machine learning techniques to analyse historical data and identify patterns for future events (Adebola Folorunso 1, 2024). By analysing previous attacks and current system behaviours, future cyber-attacks can be forecasted and prevented. Machine Learning models can evaluate vast datasets faster and more deeply than human analysts. This comprehensive analysis enables organisations to understand the context and sophistication of the attack vectors, enhancing their strategic response (Maryam Roshanaei, July 2024). However, manually monitoring and analysing this data is a daunting task for human analysts or hacker forums and it becomes harder to signify an impending attack (Vegesna, July 2024). AI can identify anomalies and patterns by adopting unconventional data sources such as the dark web to predict and learn about cyber threats. The predictive analysis of AI allows organisations to bolster their preparedness and strengthen their intrusion and detection systems against any cyber threat.

Table 2. Comparison of pre and post-AI/ML integration cybersecurity metrics, (Maryam Rohshanaei, July 2024).

Metric	Before AI/ML Integration	After AI/ML Integration	Improvement (%)
Average Detection Time	48 hours	3 hours	93.75
False Positive Rate	20%	5%	75
Threat Response Time	24 hours	1 hour	95.83
Number of Undetected Attacks	50 per year	15 per year	70

 doi:10.20944/preprints202501.0360.v1

b) AI-Powered Privacy Safeguards and Regulatory Compliance

AI plays a crucial role in ensuring that organisations comply with regulatory frameworks such as the General Data Protection Regulations (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA). Organisations are required to ensure that there is stringent control over how personal data is handled. AI can therefore analyse the organisation's data and ensure that there is compliance with the set of regulations (Javeria, 2024). Real-time monitoring of data by AI helps mitigate the risks of data breaches or unsuspicious data tracking which reduces the potential financial repercussions from non-compliance and penalties (GDPR, 2018). Table 2 shows a Comparison of pre and post-AI/ML integration cybersecurity metrics.

c) Quantum-Enhanced Threat Detection

Quantum Computing represents a paradigm shift in computational capabilities, with unprecedented processing power that could revolutionise cybersecurity. Merging Quantum computing and AI will fortify cybersecurity measures such as quantum-resistant encryption algorithms that can withstand potential cyber-attacks such as ransomware (Javeria, 2024). AI and quantum computers will improve traditional encryption and solve complex cryptographic challenges (Egbuna, July 2021). As of 2023, most organisations had low Quantum-safe Readiness, with an average score of just 21 out of 100 (Ray Harishankar, 14 May 2024. AI can eventually help in predicting quantum threats, automating the adoption of new encryption methods and enhancing real-time security. Moreover, Quantum computing can process vast amounts of data at an unprecedented speed that will allow threats and vulnerabilities to be analysed rapidly (Adebola Folorunso 1, 2024). By leveraging AI and quantum computing, organisations can secure their data against the forthcoming quantum decryption revolution.

d) Automatic and Autonomous Cyber Defense

Artificial Intelligence is set to contribute significantly to the automation and agility of cybersecurity to make systems more adaptive and responsive to emerging and sophisticated threats. Additionally, AI can also transform cybersecurity by replacing conventional approaches that mostly rely on human interventions and static rules. One promising future capability is Autonomous cyber defence, which has the potential to help frontline cyber defenders improve threat identification, incident response, and risk mitigation at an unprecedented scale and speed. Autonomous Cyber Defense refers to systems that protect organisations and users by strengthening systems, managing networks and endpoints, detecting threats, and responding to and recovering from intrusions without needing direct human intervention. These systems autonomously devise and execute safe, proportionate, and effective actions to achieve goals based on observation, knowledge, as well as an understanding of the environment. By deploying AI-enabled autonomous agents, systems can make real-time decisions and take proactive measures in network security, effectively mitigating risks before they escalate. Autonomous cyber defence complements existing human-centric cybersecurity methods by leveraging machine learning algorithms. This allows swift and accurate responses to evolving threats offering a scalable and reliable layer of defence that fortifies the overall security landscape (Centre for Emerging Technology and Security, 2024)

e) Self-Learning and Adaptive Security Systems

Self-learning and Adaptive Security Systems represent the next frontier in AI-driven cybersecurity, promising a transformative shift from traditional reactive defences to a dynamic, proactive strategy. These intelligent systems have the potential to revolutionise cybersecurity by continuously evolving to anticipate and thwart sophisticated cyber threats and opening doors to more robust and flexible digital defensive environments. They constantly analyse vast amounts of historical and real-time data, adapting to evolving cyber threats with unprecedented precision. By learning from every interaction and incident, self-learning AI systems improve detection accuracy,

 doi:10.20944/preprints202501.0360.v1

identifying previously unseen attack vectors and adjusting to new cyber threats as they emerge. (Singhal et al., 2020). Additionally, self-learning AI systems reduce false positives, thus allowing security professionals to concentrate solely on real threats rather than manually wading through volumes of pointless alerts. Unlike conventional security approaches that rely on preset guidelines or signatures, Self-learning AI systems leverage machine learning algorithms to analyse large datasets and discern patterns that indicate malicious activities. These algorithms can classify malware activities and detect abnormal activities by utilising indications of compromise (IOCs) (MixMode, 2022). Self-learning and Adaptive Security Systems provide highly scalable, responsive defences that satisfy the demands of an interconnected, rapidly changing digital world through their ongoing self-improvement. By incorporating self-learning AI, these systems create a robust, proactive security framework that adapts on its own, effectively outperforming and eliminating any threats. The outcome is a dynamically responsive defence, expertly designed to tackle the intricate challenges of an ever-evolving cybersecurity landscape, equipping organisations with resilient protection tailored to the complexities of today's digital era.

f) AI-Driven Biometric and Behavioural Authentication

AI-driven biometric and behavioural authentication represents an advanced security technique that combines individuals' physical characteristics and behavioural patterns to verify their identity. This technology offers promising potential as future advances in AI improve user verification's ease and security. It offers an extremely secure, multi-layered approach that continuously adapts to subtle behavioural changes, enhancing both accuracy and precision. As artificial intelligence continues to evolve, AI-driven biometric and Behavioural Authentication is set to transform digital security, reducing the risk of unauthorised access and minimising identity fraud (Sharma et al., 2021). By seamlessly integrating machine learning algorithms with real-time data, this technology not only secures sensitive information but also improves accessibility and trust, minimising false positives and reducing the risk of identity fraud while offering users a smooth experience. Industries adopting this technology can benefit from robust, frictionless security systems that dynamically adjust to users' unique characteristics, offering a smooth and reliable verification process. Ai-driven authentication has the potential to revolutionise security standards as it develops, providing both state-of-the-art security and user comfort.

AI is revolutionising cybersecurity by enabling organisations to stay ahead of threats through predictive analysis. When combined with quantum computing, AI's ability to handle vast amounts of data is unparalleled. It fosters stronger and more adaptive security. Innovations such as self-learning systems and behavioural-based authentication will further fortify defences against complex cyber threats. Additionally, autonomous operations enable rapid responses without human intervention. Together, these advancements play a pivotal role in technology by moulding a more robust, resilient and sophisticated approach to cybersecurity and it allows organisations to protect their data and be prepared for emerging cyber threats.

4. Discussion on Security Countermeasures

i. Security Countermeasures

While discussing security countermeasures for AI-driven cybersecurity solutions such as Darktrace and Intrusion Detection and Prevention Systems (IDPS) a detailed analysis of the measures that need to be taken to address new challenges and threats faced by these systems must be presented. This part will define the standard countermeasures that are currently deployed in modern security architectures with a brief commentary on their effectiveness.

ii. Multi-Layered Defence Strategy

 doi:10.20944/preprints202501.0360.v1

Countermeasure: The defence-in-depth strategy ensures that multiple layers of different security controls, including firewalls, endpoint protection, network monitoring, and threat intelligence, work together in a manner that provides a consistent security posture. This will further help reduce the risk of failure at any one point and complement AI-driven IDPS capabilities by feeding in multiple sources of data for threat correlation (Fruhlinger, 2022; Shah et al., 2022).

Analysis: While multi-layered defence gives extensive coverage and improved accuracy in detection, managing these multiple systems may get complicated and resource-consuming (What is Multi-layered security? - Comprehensive Cybersecurity, no date). Integration with automated tools like Darktrace's Antigena simplifies response at different layers of defence.

iii. Regular Threat Intelligence Updates

Countermeasure: The use of threat intelligence feeds will help the AI models in the IDPS systems and other security solutions to identify current and emerging threats. This is very critical in the detection of novel attacks that may not immediately be identified through behaviour analysis (What is Cyber Threat Intelligence? 2024).

Analysis: Integration of real-time threat intelligence dramatically improves the proactive nature of AI-based systems. However, these feeds should be vetted for accuracy and reliability to obviate false positives and misdetections (Wolf Eye Shell, 2023).

iv. Continuous AI Model Training and Tuning

Countermeasure: This involves updating and training AI algorithms regularly to make them adjust to new network behaviours and attack methods (Santos, 2023; Nayyar et al., 2021) for them to be effective in facing zero-day vulnerabilities and advanced persistent threats (APTs).

Analysis: Continuous training is important in maintaining a high detection rate but is resource-intensive in computation and specialized expertise. This process, therefore, should be balanced to prevent biased or low-quality data from losing the accuracy of detection.

v. Reducing False Positives with Advanced Analytics

Countermeasure: Advanced data analytics and machine learning can be put into practice to enhance anomaly detection; that way, false positives are reduced (Stanham, 2023; Lim et al., 2019). Advanced algorithms that enable contextual analysis can differentiate between harmless anomalies and actual threats (Palo Alto Networks, n.d.).

Analysis: This approach reduces the workload of security teams and improves the reliability of automated responses, but such accuracy can sometimes be challenging and require complex model calibration (Monica S, 2024).

vi. Human Oversight and Response Teams

Countermeasure: It is important that AI-driven solutions, like Darktrace's IDPS, are supported by a professional Security Operations Center (SOC) team to interpret the findings of AI, investigate incidents in detail and respond to alerts that require human judgment (Legrand, 2022).

Analysis: Human oversight acts as a safeguard regarding deficiencies in AI, such as handling ambiguous cases or analyzing multi-stage attacks (Alvero, 2021; Kumar et al., 2021). However, adding this to the process brings operational costs and demands continuous training to keep them proficient with any evolving toolset.

vii. Incident Response Planning and Simulation

Countermeasure: Incident response drills and creating playbooks in regular intervals help organizations be prepared for real-time attacks. These simulations improve the response times and coordination among various security teams (Wlosinski, 2022; Jhanjhi et al., 2021).

 doi:10.20944/preprints202501.0360.v1

Analysis: Best practices of strategies of responses put teams in a ready state and familiarize them with the roles they would need to assume in the event of an incident (CISA, 2024). However, organizations are required to update these playbooks to reflect the latest threat landscape.

viii. Security of Integration with Other Systems

Countermeasure: Secured integration of AI-driven IDPS with other critical infrastructures, like SIEMs and firewalls, essentially creates an ecosystem of seamless defence that leverages data from various sources for better threat analysis (Gonsalves, 2023; Humayun et al., 2022).

Analysis: Proper integration increases the strength of AI-driven systems and enhances data correlation. However, compatibility issues can hinder these efforts especially when dealing with legacy systems (Bhutada, 2024).

ix. Enhanced Data Encryption and Access Controls

Countermeasure: Strong encryption mechanisms are needed to ensure data is properly secured at rest and during transfer. This reduces the likelihood of a data breach that could potentially be used by attackers in an attempt to bypass detection (Adams, 2022; Gopi et al., 2021). RBAC and MFA ensure that only authorized personnel can access sensitive data and system controls (Bowman and Tunks, 2024).

Analysis: All these measures protect against insider threats and unauthorized access to critical information and strengthen the security posture. However, they introduce layers of complexity that need regular maintenance and updating of policies (Reasonlabs.com, 2023; Gouda et al., 2022).

x. Analysis of Effectiveness and Challenges

Effectiveness: Most certainly, the combination of these countermeasures will provide a robust defence, especially when AI-powered systems like Darktrace are employed in real-time monitoring of network activities. The autonomous response capabilities of tools like Antigena by Darktrace can help to limit damages during fast-moving cyber incidents (Darktrace, 2024).

Challenges: The fact that such countermeasures may be limited by high false positive incidents, data dependencies, and potential constraints linked to resource availability. Further, AI models, by nature, require significant investment in training and continuous learning to keep up with new threat vectors and network behaviour.

 doi:10.20944/preprints202501.0360.v1

5. Proposed Countermeasures

The current security procedures for AI are insufficient to protect business data. As we have mentioned on Multi-Factor Authentication or known as MFA needs two or more methods to verify to access the systems to complicate attackers' attempts and be successful in penetrating the systems. However, it is seen enough that the current security of authentication has experienced getting breached by harmful attackers. To fortify these systems, biometric authentication must be enhanced. One viable approach is to standardize biometric authentication in AI security by incorporating dynamic signatures and facial data into deep learning-powered multimodal biometric authentication. It has potential future advances for increasing online security through dynamic signatures and facial data. In today's fast-paced environment, adapting to modern technologies is crucial. Systems based on 3D sensors and scanners can be excessively expensive and difficult to implement due to limited access to specialized technology. The innovative approach of combining both stable and fluid signature information with facial recognition recorded data by standard computer webcams provides a viable solution. This technology can be utilized on any smartphone or computer equipped with a simple camera to enable the collection of users' facial and signature information.

Advanced hierarchical learning models such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and Temporal Convolutional Networks (TCN) develop a sturdy various category sorting framework that significantly enhances biometric authentication performance (Fatima-Tuz-Zahra et al., 2020). It emphasizes the feasibility of integrating both constant and evolving biometric data from commonly easily obtainable resources to construct an optimized efficiency recognition system. CNNs are ideal for tasks involving images whereas LSTMs and GRUs proficient in capturing temporal relationships within sequential data and TCNs are adept at modelling extensive dependencies within time series data. The study focused on precisely obtaining hand joint coordinates using MediaPipe Hands (MPH) which is a highly efficient Python tool known for its proficiency in tracking hands and fingers.

This innovation captured signatures traced in the air before the screen enabling remote verification and demonstrating enhanced identification success when combining in-air signatures and facial photos. The data collected using a simple camera and MPH provided comprehensive and accurate three-dimensional hand tracking that worked well on both desktop and mobile computers. The CNN method is implemented in the network model for facial and 2D signature images due to its robust architecture for image tasks. LSTM networks handle signature signals while GRU networks address missing gradient issues, improve information flow, and train faster due to their streamlined architecture compared to LSTM. TCNs excel at sequence modelling unlike LSTM networks as they can capture a broader range of interactions with fewer parameters and computational resources compared to traditional convolutional architectures. The built-in model succeeds in modernizing security protocols across industries and making it beneficial for organizations seeking simple yet effective security solutions.

Every built-in model surely has its lacking and since the world become more modernized can cause potential drawbacks to existing. The possible attacks pointed out by Pattnayak (2023) are attack detection advances such as Generative Adversarial Networks (GANs) can generate hyper-realistic simulated fingerprints and faces to bypass detectors including high-profile biometric hacking or a lack of agency can erode consumer trust. The centralized biometric databases can be targeted by theft causing the likeliness to get intruded and creating attraction towards attackers to do malicious hacking and identity stolen when exclusive control centralization of biometrics in immense proprietary datasets which breaches could unveil millions.

There are also breaches of data and unauthorized access which is a risky problem where sensitive data including financial records, property rights, and personal information can be at risk of illegal access if it is not appropriately protected or secured (Dogra et al., 2021). Unauthorized access can still occur due to poor management or human mistakes although data is protected by encryption and access restrictions. To solve this problem and decrease the possibility of data breaches during

 doi:10.20944/preprints202501.0360.v1

transfers or while data at dormant, strong encryption techniques are required such as Access Pattern Monitoring with User Behaviour Analytics (UBA) which could be a great solution to avoid data breaches and unauthorized access.

UBA systems can identify unusual patterns of activity including suspicious data transfers, unauthorized access attempts, and odd login timings by tracking and evaluating user behaviour. Businesses can more rapidly and efficiently recognize and address security threats and other hazards by utilizing user behaviour analytics software. Financial institutions, e-commerce platforms, and business networks are just a few environments where UBA might be applied. Recognizing client challenges and problems can enhance user experience and improve company security (Chesti et al., 2020). For example, UBA systems can identify unusual login times, strange data transfers, and illegal access attempts by tracking and evaluating user behaviour. Microsoft has implemented Access Pattern Monitoring with UBA as part of its Identity Protection service, which has been in use since 2016. Microsoft added User Behaviour Analytics to Azure Active Directory which is a tool that looks for odd login patterns and access from unfamiliar devices or locations, thereby indicating potential credential breaches using AI and machine learning. Azure's UBA features notify security teams and implement measures like Multi-Factor Authentication (MFA) for high-risk sessions.

Then moving to another security area, the rise of quantum computing presents a transformative challenge to existing encryption protocols, particularly those employed in AI systems. Classical cryptographic schemes such as RSA, ECC (Elliptic Curve Cryptography), and Diffie-Hellman key exchange rely heavily on mathematical problems like integer factorization and discrete logarithms. These problems, though computationally infeasible for classical systems, can be solved efficiently by quantum algorithms such as Shor's algorithm. These vulnerabilities expose critical AI processes, including encrypted model sharing, secure data transmission, and authentication protocols, to severe risks such as model extraction attacks, training data breaches, and compromised authentication mechanisms. (Fernández-Caramés, 2019; Brohi et al., 2020) As a result, the security foundation of these cryptosystems is rendered obsolete in the post-quantum era.

To counteract these vulnerabilities, quantum-resistant encryption (QRE), also known as post-quantum cryptography (PQC), emerges as a critical countermeasure. QRE employs cryptographic algorithms that are secure against both classical and quantum attacks. By implementing these algorithms, AI systems can sustain robust defences against quantum-enabled threats.

QRE leverages mathematical problems resistant to known quantum algorithms, ensuring long-term cryptographic security. Prominent QRE techniques include:

1. Lattice-Based Cryptography: Algorithms like Learning with Errors (LWE) and Ring-LWE rely on the hardness of lattice problems, such as the Shortest Vector Problem (SVP). (Amirkhanova, Iavich, & Mamyrbayev, 2024; Babbar et al., 2021). These problems remain computationally intractable even for quantum computers, offering a foundation for encryption schemes, digital signatures, and homomorphic encryption. In AI security, lattice-based cryptography could encrypt AI model parameters and datasets, ensuring confidentiality even against quantum attacks.
2. Hash-Based Cryptography: Merkle tree-based schemes enable secure digital signatures using collision-resistant hash functions. These signatures are highly efficient and quantum-resistant, making them ideal for verifying model authenticity and integrity in distributed AI systems. (Farooq, Altaf, Iqbal, Bautista Thompson, Ramírez Vargas, de la Torre Díez, & Ashraf, 2023)
3. Code-Based Cryptography: Techniques such as McEliece cryptosystems leverage the decoding problem in error-correcting codes. These schemes can be utilized to secure large-scale AI systems that rely on extensive data exchanges, such as federated learning architectures.

 doi:10.20944/preprints202501.0360.v1

4. Multivariate Quadratic Cryptography (MQC): By solving multivariate polynomial equations over finite fields, MQC-based cryptography secures lightweight AI applications deployed on resource-constrained devices. (Farooq, Altaf, Iqbal, Bautista Thompson, Ramírez Vargas, de la Torre Díez, & Ashraf, 2023; Alkinani et al., 2021)

To integrate QRE into AI systems, the following approaches are proposed:

- Employ lattice-based encryption to protect AI model parameters during the training and inference phases. This ensures confidentiality even when models traverse untrusted environments. (Sreerangapuri, 2024; Alex et al., 2022)
- Replace existing RSA-based digital signatures in biometric systems with hash-based or lattice-based schemes. This protects user data against quantum-enabled impersonation attacks. (Sreerangapuri, 2024; Alferidah & Jhanjhi, 2020)
- Lattice-based homomorphic encryption enables computations on encrypted data without decryption. AI systems using federated learning can perform collaborative model training while preserving data privacy across clients.
- Use quantum-resistant key exchange mechanisms, such as LWE-based key exchange protocols, to establish secure communication channels between AI modules and external systems.

But we should note that while QRE presents a robust solution, its adoption faces several challenges:

- Quantum-resistant algorithms typically require larger key sizes and higher computational resources than classical counterparts. For instance, McEliece cryptosystems demand significant storage for public keys, which could strain resource-constrained AI systems.
- Although QRE algorithms are improving, they are still being reviewed and standardized by organizations like NIST. It’s important to ensure they can work smoothly across different AI systems for widespread adoption.
- While QRE secures data against Shor's algorithm, emerging quantum attacks may target its underlying mathematical constructs. Continuous research into cryptographic resilience is essential.

 doi:10.20944/preprints202501.0360.v1

6. Conclusion

AI-powered solutions have proven themselves to be extremely important in the role of contributing to the world of cybersecurity as they can provide fast response and detection of cyber threats, this allows them to improve cybersecurity in responding to cyber-attacks and protecting their services from such attacks. Providing technologies such as user authentication, predictive analysis and much more, AI-powered security has changed how people see security as a whole in this modern age, not only making devices around the world more secure than ever but also preventing many attacks such as phishing, malware and potential data breaches.

However, with all these benefits, AI-powered security is still in a developing phase, as requiring a learning curve for users and needing extremely accurate data to function properly makes it more limited in its abilities. Furthermore, the high implementation cost may make it unappealing to many people out there, not to mention security concerns about a user's personal information being stored by AI may also be unappealing. Zero-day vulnerabilities and insider-planted physical threats that computers and algorithms are currently unable to analyse are also issues that AI simply cannot solve yet, however, it could solve over more time and development.

Integrating AI-powered security with more tools and developments would require lots of cost, research and time to develop, however, it is not without benefits as AI-powered security is absolutely essential for many applications, websites and digital landscapes, therefore more development would make it the ultimate tool for preventing any cyber-attack or data breach no matter how strong or critical it is.

In conclusion, while AI-powered security is a powerful tool as of right now, it still has its limitations, having privacy concerns and threats that a computer is unable to immediately analyse. However, given more time and development and allowing it to grow and expand alongside the growing digital landscape in the current era, AI-powered cybersecurity will play a pivotal role in developing new technologies and ensuring a safer future.

References

1. Adams, M. (2022). Understanding the Advanced Encryption Standard: Enhancing Your Data Security. [online] Businesstechweekly.com. Available at: https://www.businesstechweekly.com/cybersecurity/data-security/advanced-encryption-standard/.
2. Adebola Folorunso, Temitope Adewumi, Adeola Adewa, Roy Okonkwo and Tayo Nathaniel Olawumi (2024). Impact of AI on cybersecurity and security compliance. *Global Journal of Engineering and Technology Advances*, 21(1), pp.167–184. https://doi.org/10.30574/gjeta.2024.21.1.0193.
3. Akbanov, M., Vassilakis, V.G. and Logothetis, M.D. (2019). WannaCry Ransomware: Analysis of Infection, Persistence, Recovery Prevention and Propagation Mechanisms. *Journal of Telecommunications and Information Technology*, [online] 1(1), pp.113–124. https://doi.org/10.26636/jtit.2019.130218.
4. Alex, S. A., Jhanjhi, N., Humayun, M., Ibrahim, A. O., & Abulfaraj, A. W. (2022). Deep LSTM Model for Diabetes Prediction with Class Balancing by SMOTE. *Electronics*, *11*(17), 2737. https://doi.org/10.3390/electronics11172737
5. Alferidah, D. K., & Jhanjhi, N. (2020). Cybersecurity Impact over Bigdata and IoT Growth. *2020 International Conference on Computational Intelligence (ICCI)*. https://doi.org/10.1109/icci51257.2020.9247722
6. Alkinani, M. H., Almazroi, A. A., Jhanjhi, N., & Khan, N. A. (2021). 5G and IoT Based Reporting and Accident Detection (RAD) System to Deliver First Aid Box Using Unmanned Aerial Vehicle. *Sensors*, *21*(20), 6905. https://doi.org/10.3390/s21206905
7. Allen, J. (2023). The Human and AI Partnership: Collaborating for Enhanced Cybersecurity. [online] Available at: https://www.isaca.org/resources/news-and-trends/industry-news/2023/the-human-and-ai-partnership-collaborating-for-enhanced-cybersecurity.
8. Alvero, K. M. (2021). What Role do Humans Play in Ensuring Cybersecurity? [online] Available at: https://www.isaca.org/resources/isaca-journal/issues/2021/volume-5/what-role-do-humans-play-in-ensuring-cybersecurity.

 doi:10.20944/preprints202501.0360.v1

9. Ananna, F. F., Nowreen, R., Jahwari, S. S. R. A., Costa, E. A., Angeline, L., & Sindiramutty, S. R. (2023). Analysing Influential factors in student academic achievement: Prediction modelling and insight. *International Journal of Emerging Multidisciplinaries Computer Science & Artificial Intelligence*, *2*(1). https://doi.org/10.54938/ijemdcsai.2023.02.1.254
10. Aryaman Pattnayak (2023). *AZoAi*. [online] AZoAi. Available at: https://www.azoai.com/article/Biometric-AI-Advantages-and-Challenges.aspx?form=MG0AV3.
11. Azam, H., Dulloo, M. I., Majeed, M. H., Wan, J. P. H., Xin, L. T., & Sindiramutty, S. R. (2023). Cybercrime Unmasked: Investigating cases and digital evidence. *International Journal of Emerging Multidisciplinaries Computer Science & Artificial Intelligence*, *2*(1). https://doi.org/10.54938/ijemdcsai.2023.02.1.255
12. Azam, H., Dulloo, M. I., Majeed, M. H., Wan, J. P. H., Xin, L. T., Tajwar, M. A., & Sindiramutty, S. R. (2023). Defending the digital Frontier: IDPS and the battle against Cyber threat. *International Journal of Emerging Multidisciplinaries Computer Science & Artificial Intelligence*, *2*(1). https://doi.org/10.54938/ijemdcsai.2023.02.1.253
13. Azam, H., Tajwar, M. A., Mayhialagan, S., Davis, A. J., Yik, C. J., Ali, D., & Sindiramutty, S. R. (2023). Innovations in Security: A study of cloud Computing and IoT. *International Journal of Emerging Multidisciplinaries Computer Science & Artificial Intelligence*, *2*(1). https://doi.org/10.54938/ijemdcsai.2023.02.1.252
14. Azam, H., Tan, M., Pin, L. T., Syahmi, M. A., Qian, A. L. W., Jingyan, H., Uddin, M. F., & Sindiramutty, S. R. (2023). Wireless Technology Security and Privacy: A Comprehensive Study. *Preprints.org*. https://doi.org/10.20944/preprints202311.0664.v1
15. Babbar, H., Rani, S., Masud, M., Verma, S., Anand, D., & Jhanjhi, N. (2021). Load balancing algorithm for migrating switches in software-defined vehicular networks. *Computers, Materials & Continua/Computers, Materials & Continua (Print)*, *67*(1), 1301–1316. https://doi.org/10.32604/cmc.2021.014627
16. Basit, A., Zafar, M., Liu, X., Javed, A.R., Jalil, Z. and Kifayat, K. (2020). A comprehensive survey of AI-enabled phishing attacks detection techniques. *Telecommunication Systems*, [online] 76(1). https://doi.org/10.1007/s11235-020-00733-2.
17. Bhudata, T. (2024) Integration Security: Safeguarding your data in connected systems. https://exalate.com/blog/integration-security/.
18. Bowman, K. and Tunks, J. (2024). Understanding Data Access Control: A Comprehensive Guide. [online] Available at: https://pathlock.com/learn/data-access-control/.
19. Brohi, S. N., Jhanjhi, N., Brohi, N. N., & Brohi, M. N. (2020). Key Applications of State-of-the-Art Technologies to Mitigate and Eliminate COVID-19.pdf. *TECHRxiv*. https://doi.org/10.36227/techrxiv.12115596.v1
20. Centre for Emerging Technology and Security. (2024). *Autonomous Cyber Defence Phase II*. [online] Available at: https://cetas.turing.ac.uk/publications/autonomous-cyber-defence-autonomous-agents.
21. Chesti, I. A., Humayun, M., Sama, N. U., & Jhanjhi, N. (2020). Evolution, Mitigation, and Prevention of Ransomware. *2020 2nd International Conference on Computer and Information Sciences (ICCIS)*. https://doi.org/10.1109/iccis49240.2020.9257708
22. CISA (2021). *Understanding Denial-of-Service Attacks*. [online] Cybersecurity and Infrastructure Security Agency. Available at: https://www.cisa.gov/news-events/news/understanding-denial-service-attacks.
23. CISA (2024). Incident Response Plan (IRP) Basics | CISA. [online] www.cisa.gov. Available at: https://www.cisa.gov/resources-tools/resources/incident-response-plan-irp-basics.
24. CISCO (2024). *What Is Malware? - Definition and Examples*. [online] Cisco. Available at: https://www.cisco.com/site/us/en/learn/topics/security/what-is-malware.html.
25. CloudFlare (2023). What is a Phishing attack? *Cloudflare*. [online] Available at: https://www.cloudflare.com/learning/access-management/phishing-attack/.
26. *Cost of a data breach 2024 | IBM* (no date). https://www.ibm.com/reports/data-breach.
27. Crouse, M. (2023). *IBM Report: Average Cost of a Data Breach Rises to $4.45 Million*. [online] TechRepublic. Available at: https://www.techrepublic.com/article/ibm-data-breach-cost/.
28. *Cybersecurity and AI: The challenges and opportunities* (2024). https://www.weforum.org/agenda/2023/06/cybersecurity-and-ai-challenges-opportunities/.

29. Dana Sairangazhykyzy Amirkhanova, Maksim Iavich and Orken Mamyrbayev (2024). Lattice-Based Post-Quantum Public Key Encryption Scheme Using ElGamal's Principles. *Cryptography*, 8(3), pp.31–31. https://doi.org/10.3390/cryptography8030031.
30. Dansimp (2023). Incident response planning. [online] learn.microsoft.com. Available at: https://learn.microsoft.com/en-us/security/operations/incident-response-planning.
31. Darktrace (2019). *Pros and Cons of Darktrace 2024 | PeerSpot*. [online] Peerspot.com. Available at: https://www.peerspot.com/products/darktrace-pros-and-cons.
32. Darktrace Enterprise PRODUCT OVERVIEW Enterprise Immune System. (n.d.). Available at: https://apexassembly.com/wp-content/uploads/2018/07/Darktrace-Enterprise_-Product-Overview.pdf.
33. Darktrace.com. (2024). *Darktrace*. [online] Available at: https://darktrace.com/cyber-ai-glossary/.
34. darktrace.com. (n.d.). The State of AI in Cybersecurity: How AI will impact the cyber threat landscape in 2024 | Darktrace Blog. [online] Available at: https://darktrace.com/blog/the-state-of-ai-in-cybersecurity-how-ai-will-impact-the-cyber-threat-landscape-in-2024.
35. darktrace.com. (n.d.). The State of AI in Cybersecurity: The Impact of AI on Cybersecurity Solutions | Darktrace Blog. [online] Available at: https://darktrace.com/blog/the-state-of-ai-in-cybersecurity-the-impact-of-ai-on-cybersecurity-solutions.
36. darktrace.com. (n.d.). The State of AI in Cybersecurity: The Impact of AI on Cybersecurity Solutions | Darktrace Blog. [online] Available at: https://darktrace.com/blog/the-state-of-ai-in-cybersecurity-the-impact-of-ai-on-cybersecurity-solutions.
37. darktrace.com. (n.d.). *Darktrace | Cyber security that learns you*. [online] Available at: https://www.darktrace.com.
38. darktrace.com. (n.d.). *Darktrace Launches Enterprise Immune System Version 4 • Darktrace Newsroom*. [online] Available at: https://darktrace.com/news/darktrace-launches-enterprise-immune-system-version-4.
39. Djenna, A., Bouridane, A., Rubab, S. and Marou, I.M. (2023). Artificial Intelligence-Based Malware Detection, Analysis, and Mitigation. Symmetry, 15(3), p.677. https://doi.org/10.3390/sym15030677.
40. Dogra, V., Singh, A., Verma, S., Kavita, N., Jhanjhi, N. Z., & Talib, M. N. (2021). Analyzing DistilBERT for Sentiment Classification of Banking Financial News. In *Lecture notes in networks and systems* (pp. 501–510). https://doi.org/10.1007/978-981-16-3153-5_53
41. E. Bertino and N. Islam, "Botnets and Internet of Things Security," in Computer, vol. 50, no. 2, pp. 76-79, Feb. 2017, doi: 10.1109/MC.2017.62.
42. Farooq, S., Altaf, A., Iqbal, F., Ernesto Bautista Thompson, Vargas, D., Isabel and Ashraf, I. (2023). Resilience Optimization of Post-Quantum Cryptography Key Encapsulation Algorithms. *Sensors*, 23(12), pp.5379–5379. https://doi.org/10.3390/s23125379.
43. Fatima-Tuz-Zahra, N., Jhanjhi, N., Brohi, S. N., Malik, N. A., & Humayun, M. (2020). Proposing a Hybrid RPL Protocol for Rank and Wormhole Attack Mitigation using Machine Learning. *2020 2nd International Conference on Computer and Information Sciences (ICCIS)*. https://doi.org/10.1109/iccis49240.2020.9257607
44. Frankenfield, J. (2020). Denial-of-Service (DoS) Attack Definition. [online] Investopedia. Available at: https://www.investopedia.com/terms/d/denial-service-attack-dos.asp.
45. Fruhlinger, J. (2022) Defense in depth explained: Layering tools and processes for better security. https://www.csoonline.com/article/573221/defense-in-depth-explained-layering-tools-and-processes-for-better-security.html.
46. General Data Protection Regulation (GDPR). (n.d.). *Fines / Penalties*. [online] Available at: https://gdpr-info.eu/issues/fines-penalties.
47. Gonsalves, B. (2023) The Power of Cybersecurity Product Technology Integrations: Strengthening our digital defenses. https://blogs.cisco.com/security/the-power-of-cybersecurity-product-technology-integrations-strengthening-our-digital-defenses.
48. Gopi, R., Sathiyamoorthi, V., Selvakumar, S., Manikandan, R., Chatterjee, P., Jhanjhi, N. Z., & Luhach, A. K. (2021). Enhanced method of ANN based model for detection of DDoS attacks on multimedia internet of things. *Multimedia Tools and Applications*, *81*(19), 26739–26757. https://doi.org/10.1007/s11042-021-10640-6
49. Gouda, W., Almurafeh, M., Humayun, M., & Jhanjhi, N. Z. (2022). Detection of COVID-19 based on chest x-rays using deep learning. *Healthcare*, *10*(2), 343. https://doi.org/10.3390/healthcare10020343

50. Greengard, S. (2023). *Pegasus (spyware) | Description, Origins, Spying, & Controversies | Britannica*. [online] www.britannica.com. Available at: https://www.britannica.com/topic/Pegasus-spyware.
51. H, S. (2024). *The Role of Intrusion Detection and Prevention Systems in Cybersecurity - More Cybersecurity*. [online] More Cybersecurity. Available at: https://morecybersecurity.com/the-role-of-intrusion-detection-and-prevention-systems-in-cybersecurity/?form=MG0AV3.
52. Humayun, M., Sujatha, R., Almuayqil, S. N., & Jhanjhi, N. Z. (2022). A Transfer Learning Approach with a Convolutional Neural Network for the Classification of Lung Carcinoma. *Healthcare*, *10*(6), 1058. https://doi.org/10.3390/healthcare10061058
53. Hussain, K., Rahmatyar, A. R., Riskhan, B., Sheikh, M. a. U., & Sindiramutty, S. R. (2024). Threats and Vulnerabilities of Wireless Networks in the Internet of Things (IoT). *2024 IEEE 1st Karachi Section Humanitarian Technology Conference (KHI-HTC)*, *2*, 1–8. https://doi.org/10.1109/khi-htc60760.2024.10482197
54. IBM (2019). uk-en_homepage. [online] Ibm.com. Available at: https://www.ibm.com.
55. IBM (2023). *What is an intrusion detection system (IDS)?* [online] www.ibm.com. Available at: https://www.ibm.com/topics/intrusion-detection-system.
56. IBM (2023). *What is an intrusion prevention system (IPS)? | IBM*. [online] www.ibm.com. Available at: https://www.ibm.com/topics/intrusion-prevention-system.
57. IBM. (n.d.). *The quantum-safe clock is ticking*. [online] Available at: https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/quantum-safe.
58. Insights2Techinfo. (2024). AI-Based Intrusion Detection Systems. [online] Available at: https://insights2techinfo.com/ai-based-intrusion-detection-systems/.
59. Jhanjhi, N., Humayun, M., & Almuayqil, S. N. (2021). Cyber security and privacy issues in industrial internet of things. *Computer Systems Science and Engineering*, *37*(3), 361–380. https://doi.org/10.32604/csse.2021.015206
60. Jun, A. Y. M., Jinu, B. A., Seng, L. K., Maharaiq, M. H. F. B. Z., Khongsuwan, W., Junn, B. T. K., Hao, A. a. W., & Sindiramutty, S. R. (2024). Exploring the Impact of Crypto-Ransomware on Critical Industries: Case Studies and Solutions. *Preprints.org*. https://doi.org/10.20944/preprints202409.1325.v1
61. Kaiyue (2024). How Does AI Help Detect and Prevent DDoS Attacks in Modern Networks - Kaiyue October 23, 2024 Accelerate, Secure & Compute Your Global Business Streaming Content Data File Downloads Take your edge, to the next level. [online] Edgenext - Accelerate, Secure & Compute Your Global Business Streaming Content Data File Downloads Take your edge, to the next level. Available at: https://www.edgenext.com/how-does-ai-help-detect-and-prevent-ddos-attacks-in-modern-networks/.
62. Katiyar, D.N., Tripathi, M.S., Kumar, M.P., Verma, M.S., Sahu, D.A.K. and Saxena, D.S. (2024). AI and Cyber-Security: Enhancing threat detection and response with machine learning. *Educational Administration: Theory and Practice*, [online] 30(4), pp.6273–6282. https://doi.org/10.53555/kuey.v30i4.2377.
63. Kaur, H. and Tiwari, R. (2021). Endpoint detection and response using machine learning. *Journal of Physics: Conference Series*, 2062(1), p.012013. https://doi.org/10.1088/1742-6596/2062/1/012013.
64. Kim, A. (2024) Enhancing your cyber defense: comparing simulations and tabletop exercises. https://www.sans.org/blog/enhancing-your-cyber-defense-comparing-simulations-and-tabletop-exercises/.
65. Kumar, M. S., Vimal, S., Jhanjhi, N., Dhanabalan, S. S., & Alhumyani, H. A. (2021). Blockchain based peer to peer communication in autonomous drone operation. *Energy Reports*, *7*, 7925–7939. https://doi.org/10.1016/j.egyr.2021.08.073
66. Legrand, J. (2022). Humans and cybersecurity— the weakest link or the best defense? [online] ISACA. Available at: https://www.isaca.org/resources/isaca-journal/issues/2022/volume-1/humans-and-cybersecurity-the-weakest-link-or-the-best-defense.
67. Li, W., Jin, J. and Lee, J.-H. (2019). Analysis of Botnet Domain Names for IoT Cybersecurity. *IEEE Access*, [online] 7, pp.94658–94665. https://doi.org/10.1109/ACCESS.2019.2927355.
68. Lim, M., Abdullah, A., Jhanjhi, N., Khan, M. K., & Supramaniam, M. (2019). Link Prediction in Time-Evolving Criminal Network with deep Reinforcement learning technique. *IEEE Access*, *7*, 184797–184807. https://doi.org/10.1109/access.2019.2958873

69. Luffy, M.D. (2024). *AI technology: Opportunities and challenges in DDoS mitigation*. [online] Opportunities and challenges of applying AI to DDoS attack mitigation. Available at: https://www.vnetwork.vn/en-US/news/ung-dung-ai-trong-chong-tan-cong-ddos/.
70. Manchuri, A., Kakera, A., Saleh, A., & Raja, S. (2024). pplication of Supervised Machine Learning Models in Biodiesel Production Research - A Short Review. *Borneo Journal of Sciences and Technology*. https://doi.org/10.35370/bjost.2024.6.1-10
71. Masoumeh Abolfathi, Zohreh Raghebi, Haadi Jafarian and Farnoush Banaei-Kashani (2021). A Scalable Role Mining Approach for Large Organizations. *Proceedings of the 2021 ACM Workshop on Security and Privacy Analytics*. https://doi.org/10.1145/3445970.3451154.
72. Maués, R. (2024). *Artificial Intelligence on Secure Software Development*. [online] Conviso AppSec. Available at: https://blog.convisoappsec.com/en/artificial-intelligence-on-secure-software-development/
73. MicrosoftGuyJFlo (2024). *What is Microsoft Entra ID Protection? - Microsoft Entra ID Protection*. [online] learn.microsoft.com. Available at: https://learn.microsoft.com/en-us/entra/id-protection/overview-identity-protection.
74. MixMode. (2022). *What is self-learning AI, and how is it applied to cybersecurity?* [online] Available at: https://mixmode.ai/what-is/self-learning-ai/.
75. Mounica S (2024). Big Data For Fraud Detection and Prevention. [online] hyperverge.co. Available at: https://hyperverge.co/blog/big-data-fraud-detection/.
76. Nayyar, A., Gadhavi, L., & Zaman, N. (2021). Machine learning in healthcare: review, opportunities and challenges. In *Elsevier eBooks* (pp. 23–45). https://doi.org/10.1016/b978-0-12-821229-5.00011-2
77. Olubudo, P. (2024). *Advanced Threat Detection Techniques Using Machine Learning: Exploring the Use of AI and ML in Identifying...* [online] ResearchGate. Available at: https://www.researchgate.net/publication/380743475_Advanced_Threat_Detection_Techniques_Using_Machine_Learning_Exploring_the_Use_of_AI_and_ML_in_Identifying_and_Mitigating_Advanced_Persistent_Threats_APTs.
78. Palo Alto Networks (n.d.). What are the Risks and Benefits of Artificial Intelligence (AI) in Cybersecurity? [online] Palo Alto Networks. Available at: https://www.paloaltonetworks.com/cyberpedia/ai-risks-and-benefits-in-cybersecurity.
79. Palo Alto Networks (n.d.). What is the Role of AI in Threat Detection? [online] Palo Alto Networks. Available at: https://www.paloaltonetworks.com/cyberpedia/ai-in-threat-detection.
80. Palo Alto Networks (n.d.). What is the Role of AI in Threat Detection? [online] Palo Alto Networks. Available at: https://www.paloaltonetworks.com/cyberpedia/ai-in-threat-detection.
81. Palo Alto Networks. (n.d.). What Is Precision AITM? [online] Available at: https://www.paloaltonetworks.com/cyberpedia/what-is-precision-ai.
82. PgCert, A.R. (2024). *1. Introduction In an increasingly digital world, the need for robust, reliable, and user-friendly authentication and security measures has never been more critical.* [online] Linkedin.com. Available at: https://www.linkedin.com/pulse/ai-powered-behavioral-biometrics-andre-ripla-pgcert-5x5te/.
83. Pluralsight.com. (2023). *The impact of AI: Cybersecurity challenges and opportunities*. [online] Available at: https://www.pluralsight.com/resources/blog/cybersecurity/ai-impact-cybersecurity.
84. R. Sommer and V. Paxson, "Outside the Closed World: On Using Machine Learning for Network Intrusion Detection," 2010 IEEE Symposium on Security and Privacy, Oakland, CA, USA, 2010, pp. 305-316, doi: 10.1109/SP.2010.25.
85. Raimundo, R. and Rosário, A. (2021). The Impact of Artificial Intelligence on Data System Security: A Literature Review. *Sensors*, [online] 21(21), p.7029. https://doi.org/10.3390/s21217029.
86. Ravichandran, N., Tewaraja, T., Rajasegaran, V., Kumar, S. S., Gunasekar, S. K. L., & Sindiramutty, S. R. (2024). Comprehensive Review Analysis and Countermeasures for Cybersecurity Threats: DDoS, Ransomware, and Trojan Horse Attacks. *Preprints.org*. https://doi.org/10.20944/preprints202409.1369.v1
87. Reasonlabs.com. (2023). What are Enhanced security measures? Comprehensive Cybersecurity Strategies. [online] Available at: https://cyberpedia.reasonlabs.com/EN/enhanced%20security%20measures.html.
88. Research Gate (2021). *130+ million publications organized by topic on ResearchGate*. [online] ResearchGate. Available at: https://www.researchgate.net/publication.

89. ResearchGate. (2019). *Vinayakumar RAVI | Assistant Research Professor | Assistant Research Professor | Prince Mohammad University, Khobar | Center for Artificial Intelligence | Research profile*. [online] Available at: https://www.researchgate.net/profile/Vinayakumar-Ravi
90. ResearchGate. (2020). Joao Paulo JAVIDI DA COSTA | Professor | Dr.-Ing. | Hochschule Hamm-Lippstadt, Hamm | Department 2 | Research profile. [online] Available at: https://www.researchgate.net/profile/Joao-Paulo-Javidi-Da-Costa
91. ResearchGate. (2020). *Toan NGUYEN | Ph.D. | Research profile*. [online] Available at: https://www.researchgate.net/profile/Toan-Nguyen-26
92. ResearchGate. (2024). *Bandar ALOTAIBI | Professor (Associate) | University of Tabuk, Tabuk | Department of Computer Science | Research profile*. [online] Available at: https://www.researchgate.net/profile/Bandar-Alotaibi-4
93. Richberg, J. (2024). *The Power and Limitations of AI in Cybersecurity | CISO Collective*. [online] Fortinet Blog. Available at: https://www.fortinet.com/blog/ciso-collective/power-and-limitations-of-ai-in-cybersecurity.
94. Roshanaei, M., Khan, M.R. and Sylvester, N.N. (2024). Enhancing Cybersecurity through AI and ML: Strategies, Challenges, and Future Directions. *Journal of Information Security*, [online] 15(3), pp.320–339. https://doi.org/10.4236/jis.2024.153019.
95. Salem, A.H. *et al.* (2024) 'Advancing cybersecurity: a comprehensive review of AI-driven detection techniques,' *Journal of Big Data*, 11(1). https://doi.org/10.1186/s40537-024-00957-y.
96. Salem, A.H., Azzam, S.M., Emam, O.E. and Abohany, A.A. (2024). Advancing cybersecurity: a comprehensive review of AI-driven detection techniques. *Journal Of Big Data*, [online] 11(1). https://doi.org/10.1186/s40537-024-00957-y.
97. Salem, A.H., Azzam, S.M., Emam, O.E. and Abohany, A.A. (2024). Advancing cybersecurity: a comprehensive review of AI-driven detection techniques. *Journal Of Big Data*, [online] 11(1). https://doi.org/10.1186/s40537-024-00957-y.
98. Sanchez, P. (2023). *Denial Of Service (DoS) Attacks: A Complete Guide*. [online] CovertSwarm. Available at: https://www.covertswarm.com/post/denial-of-service-attack.
99. Santos, O. (2023) Securing AI: navigating the complex landscape of models, Fine-Tuning, and RAG. https://blogs.cisco.com/security/securing-ai-navigating-the-complex-landscape-of-models-fine-tuning-and-rag.
100. Seng, Y. J., Cen, T. Y., Raslan, M. a. H. B. M., Subramaniam, M. R., Xin, L. Y., Kin, S. J., Long, M. S., & Sindiramutty, S. R. (2024). In-Depth Analysis and Countermeasures for Ransomware Attacks: Case Studies and Recommendations. *Preprints.org*. https://doi.org/10.20944/preprints202408.2261.v1
101. Serkan Salturk and Kahraman, N. (2024). Deep learning-powered multimodal biometric authentication: integrating dynamic signatures and facial data for enhanced online security. *Neural Computing and Applications*, 36(19), pp.11311–11322. https://doi.org/10.1007/s00521-024-09690-2.
102. Shah, I. A., Jhanjhi, N. Z., & Laraib, A. (2022). Cybersecurity and blockchain usage in contemporary business. In *Advances in information security, privacy, and ethics book series* (pp. 49–64). https://doi.org/10.4018/978-1-6684-5284-4.ch003
103. Sharma, R., Singh, A., Kavita, N., Jhanjhi, N. Z., Masud, M., Jaha, E. S., & Verma, S. (2021). Plant disease diagnosis and image classification using deep learning. *Computers, Materials & Continua/Computers, Materials & Continua (Print)*, *71*(2), 2125–2140. https://doi.org/10.32604/cmc.2022.020017
104. Sharma, V., Shah, D., Sharma, S. and Gautam, S. (2024). Artificial Intelligence based Intrusion Detection System – A Detailed Survey. *ITM Web of Conferences*, 65, p.04002. https://doi.org/10.1051/itmconf/20246504002.
105. Sindiramutty, S. R., Jhanjhi, N. Z., Tan, C. E., Khan, N. A., Shah, B., & Manchuri, A. R. (2024). Cybersecurity measures for logistics industry. In *Advances in information security, privacy, and ethics book series* (pp. 1–58). https://doi.org/10.4018/979-8-3693-3816-2.ch001
106. Sindiramutty, S. R., Jhanjhi, N. Z., Tan, C. E., Khan, N. A., Shah, B., Yun, K. J., Ray, S. K., Jazri, H., & Hussain, M. (2024). Future trends and emerging threats in drone cybersecurity. In *Advances in information security, privacy, and ethics book series* (pp. 148–195). https://doi.org/10.4018/979-8-3693-0774-8.ch007

107. Sindiramutty, S. R., Jhanjhi, N. Z., Tan, C. E., Yun, K. J., Manchuri, A. R., Ashraf, H., Murugesan, R. K., Tee, W. J., & Hussain, M. (2024). Data security and privacy concerns in drone operations. In *Advances in information security, privacy, and ethics book series* (pp. 236–290). https://doi.org/10.4018/979-8-3693-0774-8.ch010
108. Sindiramutty, S. R., Jhanjhi, N., Tan, C. E., Lau, S. P., Muniandy, L., Gharib, A. H., Ashraf, H., & Murugesan, R. K. (2024). Industry 4.0. In *Advances in logistics, operations, and management science book series* (pp. 342–405). https://doi.org/10.4018/979-8-3693-1363-3.ch013
109. Sindiramutty, S. R., Tan, C. E., & Wei, G. W. (2024). Eyes in the sky. In *Advances in information security, privacy, and ethics book series* (pp. 405–451). https://doi.org/10.4018/979-8-3693-0774-8.ch017
110. Sindiramutty, S. R., Tan, C. E., Shah, B., Khan, N. A., Gharib, A. H., Manchuri, A. R., Muniandy, L., Ray, S. K., & Jazri, H. (2024). Ethical considerations in drone cybersecurity. In *Advances in information security, privacy, and ethics book series* (pp. 42–87). https://doi.org/10.4018/979-8-3693-0774-8.ch003
111. Singhal, V., Jain, S. S., Anand, D., Singh, A., Verma, S., Kavita, N., Rodrigues, J. J. P. C., Jhanjhi, N. Z., Ghosh, U., Jo, O., & Iwendi, C. (2020). Artificial Intelligence Enabled Road Vehicle-Train Collision Risk Assessment Framework for Unmanned railway level crossings. *IEEE Access*, *8*, 113790–113806. https://doi.org/10.1109/access.2020.3002416
112. Splunk. (2023). *SIEM: Security Information & Event Management Explained | Splunk*. [online] Available at: https://www.splunk.com/en_us/blog/learn/siem-security-information-event-management.
113. Sreerangapuri, A. (2024). Post-Quantum Cryptography for AI-Driven Cloud Security Solutions. *IJFMR240529032*, [online] 6(5). Available at: https://www.ijfmr.com/papers/2024/5/29032.pdf.
114. Stanham, L. (2023). AI-Powered Behavioral Analysis in Cybersecurity | CrowdStrike. [online] Available at: https://www.crowdstrike.com/en-us/cybersecurity-101/artificial-intelligence/ai-powered-behavioral-analysis/.
115. T. M. Fernández-Caramés, "From Pre-Quantum to Post-Quantum IoT Security: A Survey on Quantum-Resistant Cryptosystems for the Internet of Things," in IEEE Internet of Things Journal, vol. 7, no. 7, pp. 6457-6480, July 2020, doi: 10.1109/JIOT.2019.2958788.
116. Technology and Operations Management. (2018). Darktrace: Can Artificial Intelligence lead the fight against Cyber Crime? - Technology and Operations Management. [online] Available at: https://d3.harvard.edu/platform-rctom/submission/darktrace-can-artificial-intelligence-lead-the-fight-against-cyber-crime/
117. Thesciencebrigade.com. (2024). *View of The Impact of AI on Cybersecurity: Emerging Threats and Solutions*. [online] Available at: https://thesciencebrigade.com/jst/article/view/232/226.
118. TLP:CLEAR TLP:CLEAR Joint Cybersecurity Information Deploying AI Systems Securely Best Practices for Deploying Secure and Resilient AI Systems. (n.d.). Available at: https://media.defense.gov/2024/Apr/15/2003439257/-1/-1/0/CSI-DEPLOYING-AI-SYSTEMS-SECURELY.PDF.
119. Traill, R. (2024). Get The Drop On Phishing Attacks Abusing Dropbox. [online] Darktrace.com. Available at: https://darktrace.com/blog/legitimate-services-malicious-intentions-getting-the-drop-on-phishing-attacks-abusing-dropbox.
120. Vinod, D. and Ashwin Adepu (2024). Leveraging Artificial Intelligence for Predictive Cyber Threat Intelligence. *International Journal of Creative Research In Computer Technology and Design*, [online] 6(6), pp.1–19. Available at: https://jrctd.in/index.php/IJRCTD/article/view/64.
121. Waheed, A., Seegolam, B., Jowaheer, M. F., Sze, C. L. X., Hua, E. T. F., & Sindiramutty, S. R. (2024). Zero-Day Exploits in Cybersecurity: Case Studies and Countermeasure. *preprints.org*. https://doi.org/10.20944/preprints202407.2338.v1
122. Wan, S., Meta Platforms, Inc., Saxe, J., Meta Platforms, Inc., Gomes, C., Meta Platforms, Inc., Chennabasappa, S., Meta Platforms, Inc., Rath, A., The University of Texas at Dallas, Sun, K., George Mason University, Wang, X., & The University of Texas at Dallas. (2024). Bridging the Gap: A Study of AI-based Vulnerability Management between Industry and Academia [Journal-article]. Bridging the Gap: A Study of AI-based Vulnerability Management Between Industry and Academia.

123. Wen, B. O. T., Syahriza, N., Xian, N. C. W., Wei, N. G., Shen, T. Z., Hin, Y. Z., Sindiramutty, S. R., & Nicole, T. Y. F. (2023). Detecting cyber threats with a Graph-Based NIDPS. In *Advances in logistics, operations, and management science book series* (pp. 36–74). https://doi.org/10.4018/978-1-6684-7625-3.ch002
124. What is Cyber Threat Intelligence? (2024). https://www.sentinelone.com/cybersecurity-101/threat-intelligence/cyber-threat-intelligence/.
125. What is Multi-layered security? - Comprehensive Cybersecurity (no date). https://cyberpedia.reasonlabs.com/EN/multi-layered%20security.html.
126. Wlosinski, L. (2022). Cybersecurity Incident Response Exercise Guidance. [online] ISACA. Available at: https://www.isaca.org/resources/isaca-journal/issues/2022/volume-1/cybersecurity-incident-response-exercise-guidance.
127. Wolf Eye Shell. (2023). Regular Software Updates and Patch Management In Cyber Security - Wolf Eye Shell. [online] Available at: https://wolfeyeshell.com/regular-software-updates-and-patch-management-in-cyber-security/.
128. World Economic Forum (2024). *Global Cybersecurity Outlook 2024*. [online] World Economic Forum. Available at: https://www3.weforum.org/docs/WEF_Global_Cybersecurity_Outlook_2024.pdf.
129. www.biia.com. (n.d.). Harnessing Predictive Analytics In Cybersecurity | BIIA.com | Business Information Industry Association. [online] Available at: https://www.biia.com/harnessing-predictive-analytics-in-cybersecurity/.
130. www.future-processing.com. (2023). *Artificial intelligence usage in multi-factor authentication | Blog - Future Processing*. [online] Available at: https://www.future-processing.com/blog/artificial-intelligence-usage-in-multi-factor-authentication/.
131. www.linkedin.com. (n.d.). *Protecting Your Castle: Understanding IDS, IPS, and EDR*. [online] Available at: https://www.linkedin.com/pulse/protecting-your-castle-understanding-ids-ips-edr-sohail-%D1%85%D0%B0%D0%BA%D0%B5%D1%80--ifhtf/.
132. www.screeb.app. (n.d.). *Understanding User Behavior Analytics (UBA) and Its Applications for Improved User Experience*. [online] Available at: https://screeb.app/blog/user-behavior-analytics-definition-methods-and-use-cases.