Scientific Research Publishing

Journal of Information Security, 2024, 15, 245-278
https://www.scirp.org/journal/jis
ISSN Online: 2153-1242
ISSN Print: 2153-1234

The Role of AI in Cyber Security: Safeguarding Digital Identity

Mohammad Binhammad[1], Shaikha Alqaydi[1], Azzam Othman[1], Laila Hatim Abuljadayel[2]

[1]Department of Computer Science, British University Dubai, Dubai, The United Arab Emirates
[2]Department of Cybersecurity, Dar Al Hekma University, Jeddah, Saudi Arabia
Email: m.hassan.y1994@gmail.com, azzamothman@gmail.com, shaikha.alqaydi@gmail.com, labuljadayel@yahoo.com

How to cite this paper: Binhammad, M., Alqaydi, S., Othman, A. and Abuljadayel, L.H. (2024) The Role of AI in Cyber Security: Safeguarding Digital Identity. *Journal of Information Security*, **15**, 245-278.
https://doi.org/10.4236/jis.2024.152015

Received: March 24, 2024
Accepted: April 27, 2024
Published: April 30, 2024

Abstract

This article signals the use of Artificial Intelligence (AI) in information security where its merits, downsides as well as unanticipated negative outcomes are noted. It considers AI based models that can strengthen or undermine infrastructural functions and organize the networks. In addition, the essay delves into AI's role in Cyber security software development and the need for AI-resilient strategies that could anticipate and thwart AI-created vulnerabilities. The document also touched on the socioeconomic ramifications of the emergence of AI in Cyber security as well. Looking into AI and security literature, the report outlines benefits including made threat detection precision, extended security ops efficiency, and preventive security tasks. At the same time, it emphasizes the positive side of AI, but it also shows potential limitations such as data bias, lack of interpretability, ethical concerns, and security flaws. The work similarly focuses on the characterized of misuse and sophisticated cyberattacks. The research suggests ways to diminish AI-generating maleficence which comprise ethical AI development, robust safety measures and constant audits and updates. With regard to the AI application in Cyber security, there are both pros and cons in terms of socio-economic issues, for example, job displacement, economic growth and the change in the required workforce skills.

Keywords

Artificial Intelligence, Cyber Attack, Cyber Security, Real-Time Mitigation, Social Media Security, AI-Driven Threat Intelligence

1. Introduction

Technology has evolved at an exceptional and unprecedented rate, and automation is increasing; the twenty-first century is witnessing AI adoption in different

fields more and more. Factors such as the immense leap in computation capability, the inseparability of IOT technologies in almost all industries, and business digitization through big data analytics stand for the utility of AI-based solutions. Nevertheless, on the road of artificial intelligence, technologies increasingly prevail, and questions about the security of society and the welfare of humanity are on a parallel rise [1].

AI applications appear to be very broad; therefore, this paper restricts its focus to AI vulnerability, opportunity, and future in the cyber security space. The objective and purpose of cyber security generally varies between different companies, and it is interpreted differently across diverse standards worldwide. According to the ISO/IEC 27032: In 2023 (en), Cyber security Internet safety will be paramount as people rely heavily on digital systems through various electronic devices. Cyber security guards internet-based networks with devices consisting of hardware, software, and data from malicious activities like hacking and unauthorized access. Cyber security, at its core, comprises a collection of technologies and processes to defend against potential threats to systems, networks, and data.

The policy could then be worded to protect against cyber risks such as hacking, data breaches, and privacy invasions. It could also help preserve defenseless assets such as critical infrastructures, sulfur dioxide levels, etc. Amid this, the related cyber security does not start anymore at the perimeter but is instead at the heart, that is, to identify the assets with value and specially localize the threat. The steps should be prioritized to ensure a strong defense perimeter that will safeguard the continuity of service [2].

In the last few years, cyberspace has been the target of Cyber-attacks that have critically affected water supply systems, petrochemical installations, nuclear power plants, and transportation infrastructure systems to shut down electricity supply and tamper with essential data. This fast-growing risk and the OT and IT merging have made the system much more complicated because of the integration of physical devices and computer networks with sensors and software. IEC encourages incorporating cyber resilience from such a complex and insecure situation through various concepts that affect everything ranging from processes to humans and technology to artificial intelligence.

AI-driven technologies can control cybercrime but also have both positive and negative effects. Some AI focuses on cyber resiliency, while others focus on infiltration. Academic studies on AI applications are diverse, but security should be the primary focus. A cyber defense survey suggested integrating AI to analyze large amounts of data and detect cyber threats. Research focused on applying neural network models to enhance cyber resilience in the military. Despite a decade, opportunities and threats of auto AI remain open for exploration and development [3].

2. The Impact of Artificial Intelligence Power Security on Cyber Security

AI-based cyber security technologies have evolved significantly, transforming

the way security professionals protect digital assets and overcome cyber threats. Initially used for automating routine tasks, AI now excels in sophisticated tasks like pattern recognition, anomaly detection, and predictive analysis, thanks to the efficiency of machine learning and deep learning.

Initially, AI integration into cyber security models focused on automating repetitive tasks to increase efficiency and reduce human errors. AI machines were used to detect network intrusions, malware, and security gaps. However, rule-based algorithms struggled to adapt to complex threats, leading to the need for advanced artificial intelligence practices. This led to the popularity of machine learning algorithms, which allow AI systems to learn from data and improve performance over time as shown in Figure 1. Machine learning algorithms allow AI systems to process vast amounts of data and identify potential security threats early, enabling safety experts to spot and mitigate risks more successfully. These algorithms also perform predictive analysis, allowing security experts to anticipate and avoid security breaches before they occur. This shift in AI practices has led to a surge in popularity in areas like natural language processing [1].

The 21st-century AI sector has seen significant advancements due to scientific research, with Turin's Turing test in the mid-19th century and McCarthy's work in the late 19th century demonstrating its potential. The 1990s saw increased computational power and data generation and processing systems, leading to the development of machine learning and the Neural Network (NN) architecture. The armed forces recognized the need for AI in National Intelligence Infrastructure (NII) deployment, particularly in border security. AI can analyze vast

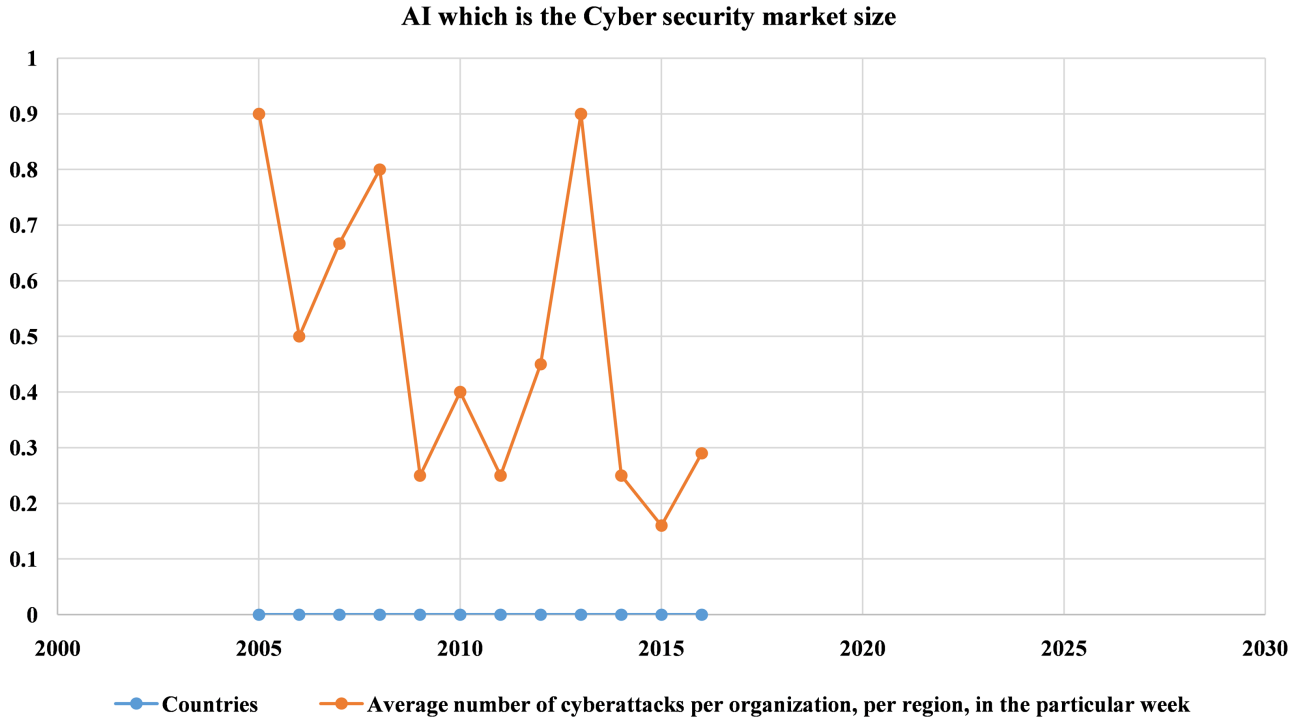

Figure 1. The artificial intelligence Artificial Intelligence (AI) is the Cyber security market size between the years, 2022 and 2032 (US billion) [3].

amounts of data, detect attack patterns and signatures, inform intrusion tools, and improve decision-making capabilities [4].

The government's need to take proactive measures to enhance effectiveness in the new environment is crucial for competitiveness. AI has been suggested for network intrusion detection and managing cyber-attack impacts, as it can analyze vast amounts of data, detect attack patterns, and improve decision-making capabilities. The concept of Big Data and its advanced AI models emerged in the early 20th century, transforming the way forward in the AI sector as shown in Figure 2.

Advancements in algorithms and hardware, such as the GPU, have enabled AI to handle data transfer and learn from millions of data at new speeds. Statistical approaches inspired ideological AI theories, and they have been applied to network monitoring tools for security and anomaly detection. HIDE and GAFT schemes are generalized anomaly and fault detection methods using preprocessing and artificial neural network classifications to create a responsive system for detecting threats and faults [5].

In cyber-terrorism, detecting threats involves determining patterns and the shape of data out of data. Probabilistic ontologies can help understand the likelihood of certain events and behaviors critical in detecting attacks or security breaches. Conventional graphs are limited in supporting mechanisms against physical attacks because they cannot address unknown hazards during real-world attacks [6].

The Adversary Courses of Action (ACoA) based on classical AI planning is proposed to address this issue. A Plan Domain Definition Auto language (PDDL)

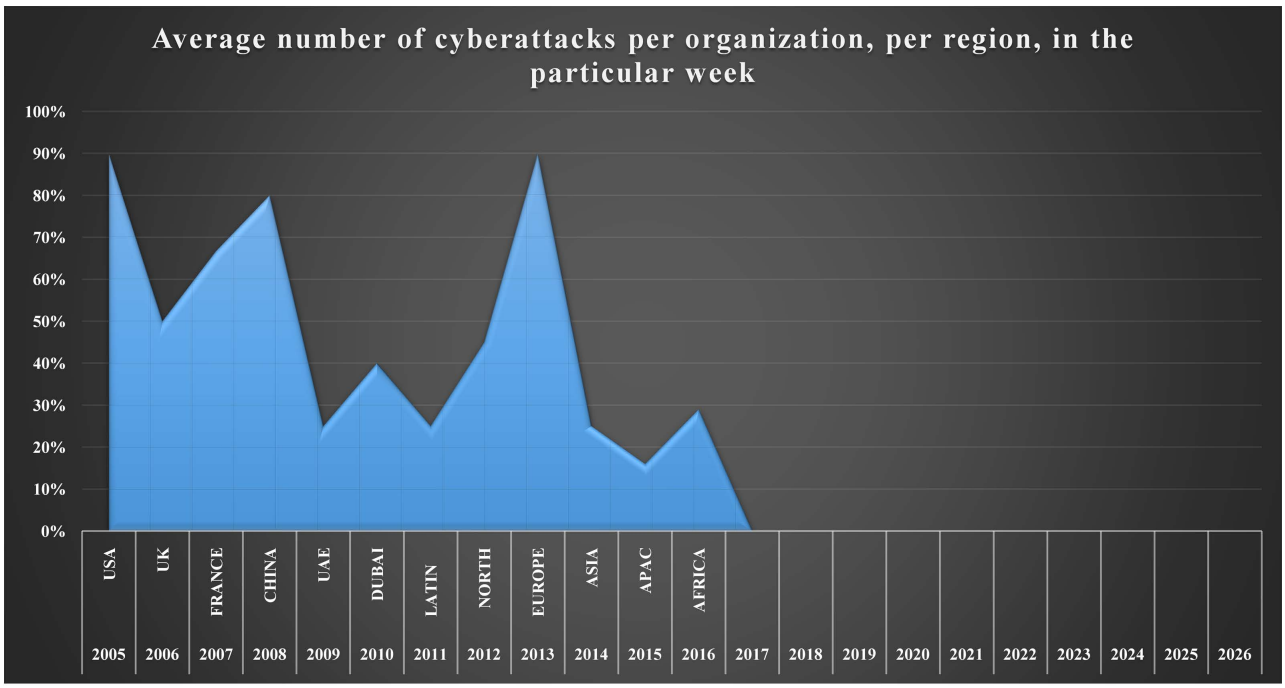

Figure 2. Average number of cyberattacks per organization, per region, in a particular week [3].

is used to create a simplified Document Management System (DMS) and a Course of Action (COA), allowing system operators to develop detailed threat plans that predict vulnerabilities and help prevent malicious acts involving insider threats [7].

The Intelligent Threat Prevention and Sensing Engine (ITPSE) was designed to support rule bases generated from expert knowledge systems and live traffic monitoring. An Intelligent Assistant System (IAS) is suggested, which incorporates a multidimensional architecture combining statistical models and machine learning methods for higher accuracy in security analysis. AI agents are created to challenge WWW browsing interfaces by allowing multivariate integration and increased efficiency among in-house computing environments.

As shown in **Figure 3**, AI has made significant strides in cyber security, with transformer models and multi-head self-attention mechanisms being particularly effective in detecting complex relationships in long patterns. Deep learning algorithms, modeled after the human brain's decision-making process, have significantly improved the precision and effectiveness of AI cyber security systems in identifying and addressing threats in real-time. AI is now a significant tool for protecting digital assets and preventing cyber-attacks, monitoring network traffic, detecting and dealing with malware, and solving potential security risks. AI-related technologies are also used to develop better security training and improve businesses' cyber security positions. As AI technologies advance, they are beginning to recognize systems that can address and deal with real-time threats, protecting organizations against cybercrimes and contributing to the stability and privacy of online interactions. This advancements will enable organizations to operate safely and securely in a digital world [9].

3. Application of AI in Cyber Security

3.1. Threats Detection and Prevention

About cyber security, prevention of threats and detection are essential, and artificial intelligence (AI) is utilized to robustify these capabilities. Such systems can effectively analyze enormous data volumes for threat detection correlating to malware, phishing, or insider threats. **Table 1**, shows how different organizations detect threats with assistance of artificial intelligence.

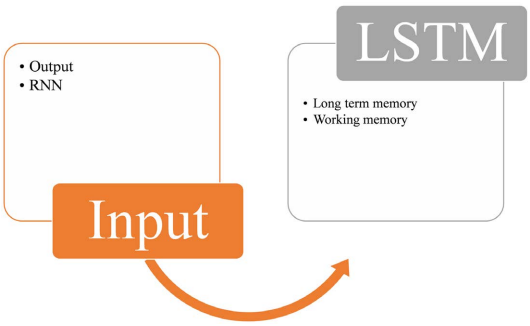

Figure 3. LSTM and RNN models [8].

Table 1. Threat detection and prevention in different organizations.

Organization	Threat Detection		
Financial Institution	It uses AI to analyze transaction data for anomalies that may indicate fraud.	Utilizes AI to implement multi-factor authentication, encryption, and access controls to prevent unauthorized access.	Financial Institution
Healthcare Provider	AI analyzes patient data to detect unusual patterns that may indicate fraudulent insurance claims or unauthorized access to medical records.	Implements AI-powered cyber security solutions to secure electronic health records (EHRs) and medical devices against cyber threats.	Healthcare Provider

AI is a powerful tool in threat detection due to its ability to peruse real-time data. Traditional methods, which rely on predefined rules, are time-consuming and inefficient. AI can adapt to new threats quickly by learning from new data, and identifying patterns and trends that analysts may not see. It can eliminate false positives by analyzing multiple data sources and identifying threats with greater accuracy than machines, reducing false alarms and focusing on genuine danger. AI can also be used for threat prevention, detecting and blocking threats that are hardly noticeable to humans, such as unapproved access issues or data leakage. Preliminary AI can add essentials to threat detection and prevention of Cyber security, determining real-time risks, reducing false positives, and proactively revealing hidden threats, improving cyber security [10].

3.2. Anomaly Detection

Anomaly detection is crucial in cyber security, identifying potential threats that break out of normal systems or networks. AI is advancing in this field, enabling improved detection and response to threats. AI-driven systems can handle vast amounts of data from various sources, such as network traffic, user behavior, and system logs. They can spot hidden anomalies that are beyond human intelligence. AI's unique advantage lies in its ability to learn and improve by studying these anomalies (as shown in Table 2), allowing it to apply itself to new and more dangerous threats. AI algorithms can accumulate new data and fine-tune their models to intercept emerging dangers, replacing static rule-based systems [11].

AI can improve anomaly detection efficiency by reducing false positives and identifying true outliers more accurately. It can gather data from multiple sources, enabling accurate identification of threats and reducing false alarms. AI also aids in detecting speed and accuracy, processing data as events develop, enabling companies to act promptly in identifying threats in the early stages. This has led to a cyber security anomaly phenomenon, enabling organizations to combat threats more efficiently. By incorporating AI technology, companies can significantly enhance their Cyber security stance and deal with cyber threats [12].

3.3. Incident Response

Incident response is an essential part of Cyber security, and it involves fast-track identification, management, and reduction of the effects of those incidents so that the organization can be protected. AI is now seen as a tool that enhances an organization's incident response capabilities, allowing it to respond to cyber threats more powerfully and quickly.

The AI-based incident response systems assist organizations in finding problems and coping with them in real-time (as shown in Table 3). These systems can analyze vast volumes of data from different sources like network traffic, logs, endpoint devices, etc.; they control security issues and prioritize them based on consequences [13].

AI plays a crucial role in incident response by performing repetitive tasks and enabling personnel to focus on complex security tasks. As attacks on information assets become more complex, AI is being used to improve incident response speed, agility, and accuracy. AI can detect incidents in real-time and proactively respond to prevent escalation, reducing false alarms and negatives. It also helps in reducing false alarms and negatives, ensuring security experts are alerted to natural threats. AI can also increase its capabilities in responding to incidents by discovering cause-and-effect relationships and providing solutions. AI algorithms can

Table 2. Anomaly detection in different organizations.

Organization	Anomaly Detection		
Financial Institution	It uses AI to detect unusual patterns in transaction data that may indicate fraudulent activity.	Financial Institution	It uses AI to detect unusual patterns in transaction data that may indicate fraudulent activity.
Healthcare Provider	Utilizes AI to identify anomalies in patient data that may indicate potential health issues or fraudulent activity.	Healthcare Provider	Utilizes AI to identify anomalies in patient data that may indicate potential health issues or fraudulent activity.

Table 3. Incident Response in different organizations.

Organization	Incident Response		
Financial Institution	Resorts to an AI played to automate initial steps like alert triage and preliminary investigation, leading human teams to focus on the more complex assignments.	Financial Institution	Resorts to an AI played to automate initial steps like alert triage and preliminary investigation, leading human teams to focus on the more complex assignments.
Healthcare Provider	AI is used by it to discover the anomaly in patient data that may lead to the access of medical records without authorization, enabling fast, easy solutions.	Healthcare Provider	AI is used by it to discover the anomaly in patient data that may lead to the access of medical records without authorization, enabling fast, easy solutions.

identify patterns of procedural weaknesses in past incidents, ensuring better-decentralized response processes and cyber security. Overall, AI enhances organizations' ability to respond to security incidents and creates better security for cyber threats [14].

3.4. Vulnerability Management

Vulnerability management is a crucial aspect of cyber security, where it identifies, evaluates, and mitigates vulnerabilities, namely, a weakness, flaw, or error in a system or application. Similarly, AI has extensively automated weakness detection techniques, allowing companies to eliminate and count the vulnerabilities in order of their significance.

Systemized AI vulnerability processing, as shown in **Table 4**, can examine continuous information flows and search for software and application weaknesses. Such systems can apply machine learning algorithms to analyze historical data, draw conclusions about the patterns, and predict the vulnerabilities before hackers bring them up [15].

AI plays a crucial role in vulnerability management by ranking threats based on their severity and potential damage risks. AI algorithms can perform risk scoring and coordinate vulnerabilities based on risk level and compromise potential. This tool automates the process and reduces human labor in finding and protecting vulnerabilities. AI-powered frameworks can automatically identify system and application vulnerabilities and provide remediation recommendations. AI also enhances the precision of vulnerability management by minimizing unneeded positives. By analyzing multiple data sources simultaneously, AI algorithms can identify actual vulnerabilities more precisely, resulting in fewer false alarms. AI has facilitated the transition to vulnerability management, making the detection and prioritization of cyber security weaknesses more precise. It is a powerful tool for organizations to sensitize vulnerability vigilance and

Table 4. Vulnerability Management in different organizations.

Organization	Vulnerability Management		
Financial Institution	Utilizes AI to automate vulnerability scanning and assessment processes, identifying and prioritizing vulnerabilities based on risk and potential impact.	Financial Institution	Utilizes AI to automate vulnerability scanning and assessment processes, identifying and prioritizing vulnerabilities based on risk and potential impact.
Healthcare Provider	AI is used for analyzing security vulnerabilities occurring in medical devices and systems; high patches and updates protect the risk of cyber-attacks and data breaches would be in a position.	Healthcare Provider	AI is used for analyzing security vulnerabilities occurring in medical devices and systems; high patches and updates protect the risk of cyber-attacks and data breaches would be in a position.

protect against cyber threats [16].

3.5. User Authentication

Access control is crucial in cyber security, ensuring only authorized users can access systems and data. AI has enhanced user authentication capabilities, making it easier for organizations to adopt secure and efficient methods. Biometric authentication is a critical application of AI, as weight sensors can replace physical contact methods by analyzing biometric data like fingerprints, voice patterns, or facial features. These systems provide security tools that surpass traditional password-based methods, as individual biometric data is unique and impossible to forge. AI also enhances the authenticity and speed of operations, impacting logistics, intelligence analysis, warfare operations, military platforms, and training as shown in **Table 5**. By studying user behavior patterns, AI can identify unique behavioral profiles, allowing systems to detect unusual patterns that could lead to unauthorized access to accounts.

On the other hand, AI is good for the security of biometric identification (i.e., multi-factor authentication like MFA) in authenticating mechanisms. MFA demands that users pass several authentication processes, for instance, a password or fingerprint verification, a one-time passcode, or many others, to get into systems and data. AI can get such an outlook from these many facets and will be able to determine if the user can gain any permission.

Recapitulating, AI has brought about a turning point in the cyber security industry by allowing organizations to implement secure and effective verification methods better. Organizations can deploy an AI-based system to boost their authentication process, a preventive measure for preventing unauthorized access, penetration, and cyber threats [17].

Table 5. User Authentication in different organizations

Organization	User Authentication		
Financial Institution	We do biometric authentication with AI, such as facial and voice recognition. Sounds biometric authentication has improved security and prevented unauthorized access to accounts.	Financial Institution	We do biometric authentication with AI, such as facial and voice recognition. Sounds biometric authentication has improved security and prevented unauthorized access to accounts.
Healthcare Provider	It provides MFA for AI-enabled devices, which implies a need for users to authenticate themselves using multiple-factor verification, such as passwords and biometrics required to access health records and systems.	Healthcare Provider	It provides MFA for AI-enabled devices, which implies a need for users to authenticate themselves using multiple-factor verification, such as passwords and biometrics required to access health records and systems.

3.6. Security Analysis

Security analysis is to determine what data was attacked, how it is used, how persistent the attack is, how much sensitive data is in the process if these are data breached, how far this data has been spread, and so on, and to choose the right tools to help deal with these attacks. Artificial Intelligence is the backbone of security analytics performance for companies to prevent threats and react to them fast enough.

The use of AI in security analysis as shown in Table 6, which is drawn from the study of extensive databases, is one of the primary applications of AI. On the other hand, artificial intelligence-supported tools can scan data from different inputs, including network traffic, logs, and end devices, and then extract the data with a specific pattern or anomaly that could indicate a security incident. This is possible inasmuch as AI does real-time data analysis, thus helping organizations identify, contain, and mitigate security vulnerabilities quickly.

Machine learning through AI can additionally offer precision in security analytics as it reduces the need for false positives. AI can analyze data from multiple sources and build relationships and patterns, and in this way, it can differentiate between real threats and false alarms more efficiently. Consequently, the cyber security operatives inspect and handle only actual emergencies.

Moreover, AI can assist in boosting the accuracy of security analytics by providing faster and more reliable information processing. AI-driven systems can process information in real time, enabling organizations to identify and tackle problems immediately. Moreover, AI can ultimately be used for automated security data analysis, regressing the focus on the manual effort needed [18].

Briefly, from AI, there is the revolution of security analytics in cyber security, the agents of which can now act in the best and quickest ways to prevent threats before they occur. Organizational security analysis can be further improved by

Table 6. Security Analytics in different organizations.

Organization	Security Analytics		
Financial Institution	AI should allow you to monitor transactions in real-time and user behavior and prevent fraudulent activities, assuring overall security and lowering financial risks.	Financial Institution	AI should allow you to monitor transactions in real-time and user behavior and prevent fraudulent activities, assuring overall security and lowering financial risks.
Healthcare Provider	Utilizes AI to analyze patient data and detect deviations that might be security threats like hacking or illegitimate access to medical records, apprising the professionals about such situations and ensuring privacy and security.	Healthcare Provider	Utilizes AI to analyze patient data and detect deviations that might be security threats like hacking or illegitimate access to medical records, apprising the professionals about such situations and ensuring privacy and security.

employing AI's power, and this approach will make it more difficult to detect cyber threats.

3.7. Threat Intelligence

Those procedures involving collecting, analyzing, and comprehending information on potential or concrete cyber threats refer to cyber threat intelligence, a crucial element of cyber security. It assists organizations' triage processes in identifying actors' methods, techniques, and plans (TTPs) and recognizing the most appropriate solutions to protect systems and data better as shown in Figure 4.

AI has revolutionized risk intelligence by turning the data flow from threat detection to countering the threats into an automated process in different organizations as shown in Table 7. AI algorithms can scan and proceed with every data source from network traffic, logs, social media, etc., to discover any patterns and constituents that may be recognized as a threat. These standards enable companies to find out where threats come from and how to render them incompetent much more quickly [19].

It notifies users about an insecure session and allows them to log out safely. AIs can monitor for threats and trends live, thus giving firms notifications of the current activities to help make security decisions.

Figure 4. Threat Intelligence in different organization.

Table 7. Threat Intelligence in different organizations.

Organization	Threat Intelligence		
Financial Institution	Utilizes AI for real-time monitoring of threats and vulnerabilities in the financial sector, providing timely insights to enhance security measures and protect against cyber-attacks.	Financial Institution	Utilizes AI for real-time monitoring of threats and vulnerabilities in the financial sector, providing timely insights to enhance security measures and protect against cyber-attacks.
Healthcare Provider	Uses AI to analyze threat intelligence data and identify potential security threats and vulnerabilities in medical systems and devices, ensuring patient data privacy and security.	Healthcare Provider	Uses AI to analyze threat intelligence data and identify potential security threats and vulnerabilities in medical systems and devices, ensuring patient data privacy and security.

In addition, AI can assist in upgrading the precision and applicability of threat intelligence information. Analyzing data from different sources combined with the connection of information can detect trends of the threats and rank them based on severity so that it can be an appraisal process. This makes it possible to prioritize attention on these high-priority threats before directing resources toward the more manageable ones.

In a nutshell, AI has altered threat intelligence in six ways-equipping organizations with technologies needed to detect, analyze, and respond to cyber threats more efficiently. Through AI, organizations are empowered to beef up their Cyber security measures to curb any likely threats [20].

3.8. Fraud Detection

Cyber security and financial security nurture the effectiveness of fraud detection, which is actually about detecting and preventing fraud, e.g., identity thefts, check fraud and account takeovers. Over the years, AI has been instrumental in boosting fraud detection in multiple organizations as well, as shown in **Table 8**. It provides the ability to analyze large volumes of data in real-time, determine characteristics likely to be associated with fraudulent behaviors and take the necessary measures to avoid crimes.

Deriving from the AI's powerful capabilities for fraud detection, as shown in **Figure 5**, is the quality to process and extract relevant information from multiple sources and, consequently, to understand intricate and unexpected patterns of fraudulent activity. Human computers can use AI systems' data analysis skills to easily detect anything suspicious, such as transactions, behavior, and historical patterns, to afford them more time for further investigation [21].

AI, in turn, can assist in mitigating falsely reported fraud cases. By using data more accurately and stopping the incineration of the actual instances of fraud, AI-powered intelligence systems would decrease the number of spurious alarms,

Table 8. Fraud detection in different organizations.

Organization	Fraud Detection		
Financial Institution	Utilizes AI for real-time monitoring of transactions and user behavior to detect and prevent fraudulent activities, enhancing overall security and reducing financial risks.	Financial Institution	Utilizes AI for real-time monitoring of transactions and user behavior to detect and prevent fraudulent activities, enhancing overall security and reducing financial risks.
Healthcare Provider	Uses AI to analyze billing data and patient records to detect anomalies indicating fraudulent insurance claims or billing practices, ensuring compliance and preventing financial losses.	Healthcare Provider	Uses AI to analyze billing data and patient records to detect anomalies indicating fraudulent insurance claims or billing practices, ensuring compliance and preventing financial losses.

Figure 5. Fraud detection through visualization in different organizations.

and organizations could concentrate on investigating fraud cases.

Moreover, AI can increase the speed of fraud detection. Thanks to robust AI-based systems capable of analyzing terabytes of data in milliseconds, businesses can now tackle fraud even when it happens at unprecedented speed. This real-time detection function of technology makes the fight against fraud in the internet or financial services virtually impossible. AI has been transforming the fraud detection process by giving organizations a platform on which it is possible to monitor the occurring anomalies, analyze the data, and take steps to avoid failure. Through AI utilization, institutions may improve their abilities to detect fraud while preventing a variety of similar kinds. Apart from possibly the most obvious application, AI in cyber security is vast and keeps growing as technology progresses. AI can transform online identities' detection, prevention, and response algorithms, thus, organizational Cyber security [22].

4. AI Proves to be an Indispensable Partner in a Security Journey with the Modern AI Technologies Which has been Riously Developed

AI, which again is a broad domain with a lot of cutting-edge technology, has different versions of its application that are shaped to the safety of the use of technology. Of the machine learning algorithms, Deep Learning (DL) and Neural Networks (NN), Expert Systems (ECS), and Natural Language Processing (NLP) are some of several models of AI. Applying these advanced intelligent algorithms can solve many of the complex problems concerning the new breed of previously unknown attacks in cyberspace, which is much more than a workforce effort. However, decoding how they work and seeing their meanings in practice requires that you dig deeper to reach their operating principles so that you get the whole picture of their use in the CS application [23].

4.1. Machine Learning Models

Machine learning (ML) is a method that uses mathematical data representation to make predictions and make decisions. It is a pre-processing step before transmitting data, essential for AI to learn and enhance cyber security. The mechanism is interpretable, generating insights based on feature saliency and ranking critical data attributes. Accuracy and correctness are crucial for detecting jerkiness. The next stage is feature engineering, which helps identify anomalies and outliers better at small changes in attack patterns as shown in **Figure 6**. However, DL models have limitations in analyzing complex patterns and long sequence data, making decision-explanation difficult. ML applications can be particularly useful in fields where adequacy is critical, as human populations often perceive and understand environmental issues differently than wildlife populations [23].

Intra-D-Tree is a futuristic intrusion detection system designed to protect IoT devices from Distributed Denial-of-Service (DDOS) attacks. It uses machine

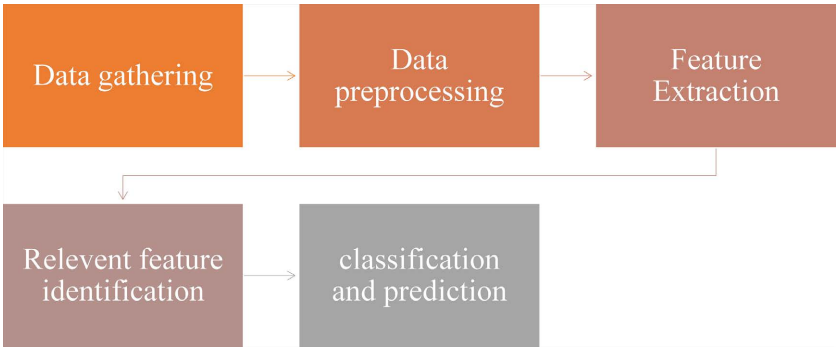

Figure 6. Different steps of machine learning model [23].

learning and tree models to generate security data from IoT devices. Intra-D-Tree incorporates IDS, streamlining decision-making processes by using less costly features and ranking features based on their significance. Tree-based models, which use classification and regression trees, are more effective than general-purpose intrusion detection models. ML models are essential for devices like mobile phones and Wi-Fi sensors, but they lack a dynamic detection mechanism, potentially leading to the adoption of dynamic detection models for network monitoring and cyber security in the future.

4.2. Deep Learning Technologies and Neural Network Based Concepts

Deep learning networks are constructed with nodes similar to input, hidden, and output layers, and their effectiveness depends on the arrangement and interconnection of computational layers. The central premise of deep learning models is to mimic the sophisticated neural network in our brains, which requires complex connections and cooperation for future-oriented decisions. Convolutional Neural Networks (CNNs) are the most commonly employed architecture of deep learning, leading to significant breakthroughs and impressive outcomes [23] [24].

The back-propagating gradient used in CNNs reached a point where it couldn't regain lost information due to gradient descent. However, loops built into RNN networks helped retrieve lost information. LSTM networks, Long Short-Term Memory (LSTM), and Auto-encoders are the most advanced technologies in RNN, with cells regulating the sequence of information and decoder producing reconstructed outputs from compressed input. Transformer models use self-attention mechanisms to extract long-range dependencies in sequential yet complicated network traffic patterns.

Deep learning models vary to develop anomalies and Information Discrimination Systems (IDS), which can spot complex patterns and intriguing behavior. These models can study large amounts of data to learn how to recognize the type of attack, pick the kind of threat, and even identify those that are known and hidden from our vision. The flow-based IDS in System Domain Networks (SDN) processes packet information using machine learning and deep learning ap-

proaches, enabling full access to the network from any part of the world and providing room for continuous routing policy change to support robust detection [25].

Conventional machine learning techniques have issues extracting little change from high-dimensional data, while DNNs can identify such little alterations from intricate features. Wide-level IoT devices often show a genetic rearrangement of the maneuvering manner, but the adoption of deep learning in systems enables the capture of covert malicious patterns via self-learning and compression technologies. RNNs can effectively classify and determine specific cyber-attack patterns, and LSTM-based IDS software can manage long-term data dependencies, allowing for balanced high detection rates with low false alarm rates.

A variant of RNN named LSTM has been invented, making it easier to process long data sequences with the gate mechanism built into it, preventing the issue of forgetting what happened earlier.

4.3. Expert Systems and the Reason Why They Rank

Artificial intelligence (AI) is a class of Expert Systems that combine human expertise and rule-based reasoning to make decisions and solve problems. This enables systems to respond to cyber security issues with high resilience and robustness. Rule-based IDS is an example of how predefined rules are coded in the detection process for abnormalities in a network or system. These rules can be classified as techniques for logging or evaluating live traffic, and if something seems amiss, an alert will be issued.

A new possible solution to this problem is the proposal of rules and decision-tree-based IDS-RDTIDS, which relies on an embedded rule-based expert system and decision-tree approach. This hybrid method based on fog computing has a three-tiered framework, focusing on DT modelling using REP Tree and rule-based classifier using JRIP. Forest PA refines the findings from the earlier two processes to obtain datasets that hint at actual data patterns.

An advanced rule-based IDS based on JESS is introduced, providing rules for pattern-based IDS PIDE, user behavior analysis. The suggested PIDE can track suspicious behavior like an ExS, using a predefined set of rules as a pattern recognition engine. By uniting the DL-classifier with the Rule-based selection method, a system relies more on the yield rate of correct determining and less on the inaccurately recognized ratio (FPR).

Innovative Exports can be a solution to cyber security challenges in industrial IoT applications based on extensive traffic surveillance with high precision. The hybrid architecture improves DL-based classifier performance by reducing the number of trained features and making more accurate discrimination [26].

4.4. Natural Language Processing

Organizations can extract valuable intelligence from social media and web pages

by using textual data analysis and interpretation tools to understand Cyber security trends and sentiment threats. Natural Language Processing (NLP) tools help strip malware code and vice versa from the textual material, enabling the identification of information needed from security logs. Recognizing phishing attacks is a common issue that people struggle with, especially with ISO and Deep Learning models. NLP can be effectively implemented with these models to support their classification and precise recognition of phishing attacks over many emails.

NLP techniques can be used to create domain ontologies using a two-fold approach: symmetrical/adjusted machine stage and symmetrical/adjusted machine stage. Ontologies provide an architecture plan with essential elements and an initial development infrastructure. The structure consists of four phases: cognitive analysis, data storage, visualization, and REST APIs.

Corporate security investigators often encounter personal information and communication online while dealing with a massive stock of web content and open-source materials. An automated tool called Doc2Vec is proposed to extract technology content from publicly available online information using the technique of Doc2Vec. This method uses natural language processing as a filter and triage to determine cyber security contents' feed data.

An analysis was conducted to identify different attack patterns in the Common Attack Pattern Enumeration and Classification (CAPEC) database, providing Cyber security experts with suggested attack events to implement mitigation measures. Topic modelling is used to group unstructured topics and extract concealed information from the attack description database [27].

Cyber-attack sensing and Information Extraction (CASIE) technology was proposed to add a more meaningful role to applying NLP in cyber threat identification from textual data. CASIE trains the event extractor on a set of news articles and recognizes the semantics of different aspect arguments for each news event, such as Data Breach, Phishing, Ransomware, Discovering Vulnerability, and Patch Vulnerability. However, there are few inaccuracy outcomes in cyber-threat analysis performances, which depend on the data quality and representativeness of trained models.

5. Opportunities

AI tools have the potential to address social, economic, and environmental problems. However, to fully realize their potential, research projects on technological infrastructure are crucial, especially in renewable energy. Adversarial machine learning is a prominent area of research that has received significant attention due to its economic benefits. The integration of AI into human endeavors is essential for the safety and security of both now and future.

AI provides cyber applications with tools to detect, respond, and prevent threats. Deep learning algorithms and machine learning models are used to process large data sets, enabling systems to distinguish patterns and detect

emerging threats. AI-based systems perform daily tasks in the automation industry, improving job security for cyber security operators. They can also prohibit access by unauthorized people through authentication techniques and biometrics.

AI-powered technologies improve incident response, enabling organizations to quickly benefit from cyber-attacks. It is a vital tool for intrusion detection and user behavioral analysis, identifying internal threats and unwanted behavior in businesses' networks. Advanced data analytics and practical insights have defined the period of growth in computer security, elevating responsiveness and cyber resilience. As AI advances, the role of this technology in protecting digital property and ensuring privacy could become more critical in the online system.

The two IDS planes, host-based and network-based, are the main applications for nearly all cyber security problems encountered in different fields of application. A taxonomy of contributions-in-it-of-itself of other AI solutions that improve cyber-security is provided, giving details on their performance and algorithm [28].

5.1. Anomaly Detection

Diversified modelling methods are used to identify behavior deviations from standards integrated with organizational or resilience norms. These models help security personnel make timely and intelligent decisions by gaining insight from past cyber-attack patterns. A hybrid Gaussian mixture model with multiple Kernel Principal Components was proposed for anomaly detection, along with intrusion detection integrating the decision tree model. The model considers discrepancies in the dataset or logging results, such as those that may cause the mechanism to detect occurrences to be misguided or have a high false rate.

The decision tree model is designed to work with conditional statement-driven rules to cut abnormal figures from normal ones. The attributes from the decision tree model are used to finalize the choice of hardening the rules using a regular expression and specific detection patterns. In the decision tree, each leaf node of the regular class is modelled GMM, which is based on the assumption that observations are created from various Gaussian distributions but the parameters are not fixed and are instead inferred from experimental data. Transaction volume weight varying by 10% of the typical distribution pattern helps speed up the detection of attacks.

In the hybrid approach, the category of the usual pattern is put into multiple classes to give the data a detailed and profound analysis and to identify the attack patterns. The involvement of IoT devices in innovative grid systems can expose Advanced Metering Infrastructure (AMI) security. GMM-based detection algorithms can improve this system by eliminating reliance on external expertise and scanning data at any time they found necessary.

To create an intelligent city that brings about smart ICT devices, IoT, and cloud storage systems for interconnectivity, anomaly detection and other network functions can be integrated into multiple network layers at an AI level. Hybrid architecture with a centralized to distributed approach is needed to de-

fend edge-to-cloud networks in smart cities and ensure the security of smart city data on the cloud. The joint effort of SDN, multiple controllers, and ML methods at the edge networks brings about more practical security measures from malicious or abnormal data and identifies corruption of system resources [29].

5.2. Signature Based Detection

Text signature plays a crucial role in identifying essential features in data sequences with time variation and urgency structure. Two types of detectors are split into rule-set or patterns: models that prevent unspecified actions in network activity zones or behavior patterns used to detect strange behavior from traffic background. Rule-based models, such as Snort, Suricata, and Zeek, use predefined rules to identify abnormal and abnormal patterns in data.

Fuzzy rule-based models have been used for hazard identification in cross-country product pipeline systems, applying GRT Fuzzy Rule Base (FRB) strategies to address inexact, uncertain, or subjective data sets. The rise of phishing, business email exploitation, and ransomware trends during pandemics has led to the development of systems combining Fuzzy Logic, Rule-based, and Data Mining to manage rising potential and threats.

Variations are a pattern-matching method used to evaluate uncommon data with no or minimal history. Identifying the exact type of pattern is still challenging, but models like absolute median deviation (AMD) provide better detection and prevention processes. Old attack lines, benign events, and hits from various logs are used to identify patterns and devise actions for unknown live data. Digital forensics face pattern recognition using PCA, NN, and GA-integrated models is developed and trained, with a central model of federated learning that leverages PCA for dimensionality reduction and GA for pattern optimization.

5.3. Cloud Security and Encryption

The EEIBDM approach is an efficient IoT framework for industrial-scale data management, focusing on secure authentication, resource access control, and privacy protection. It uses reinforcement techniques to strengthen cloud security and uses a digital-twin system with a virtual representation of industrial data management concepts. Federated learning is used for machine learning, and the approach uses biology-based Huffman encoding for data encryption. This method enhances the integrity and stability of the cloud system, especially through the Internet of Vehicles (IoV) and interconnection elements like sensors and traffic system management. The EAST algorithm is proposed as a steganography solution for data encryption, enabling multi-layering AI-made protection against cyber-attacks [30].

5.4. Incident Response and Mitigation

AI has significantly impacted detection facilities, attack-security methods, and

response efforts in modern technological network infrastructures. AI-based models, such as Statistical Decision Trees and Neural Network models, can handle security matters and improve endpoint security by monitoring and evaluating incidents from endpoint devices. These prevention tools enable quick recognition and systematic shutting down of devices' vulnerabilities, limiting the consequences of problems across the network. In wireless sensor networks, deep learning models are exploited to detect network-imminent threats, potentially prevented by CNN-based mechanisms.

AI-driven training and assessment can help reduce phishing cases and prevent network and system breaches. Cyberattackers are becoming more cunning by applying AI in their systems, and AI can be harnessed to develop real-time adaptive decision-making processes. An advanced Intelligent Decision Support System was proposed for risk management applications in critical infrastructures of industrial domains, relying on AI methods like multi-criteria assessment, causal network, inference, and data mining approach. The development of sensors and historical data contributes to the system's design, which is used to prevent and mitigate multi-layered cyber security threats [31].

5.5. Intelligence Information and Deceit Technology

Security systems technology has evolved rapidly, from active protection methods to more effective proactive mechanisms. One such mechanism is "Honey Pot", which collects essential data to understand attack signatures and tactics against a target. This real-time threat intelligence can help security analysts and forensic specialists isolate new attacks, injection techniques, and exploitation tactics to be brought down. Threat intelligence is crucial for cyber-resilience and staying ahead of attackers, and it needs to be updated to keep up with ever-changing security measures. The AttackKG system uses machine learning to analyze the relationship between threat patterns in unstructured Cyber Threat Intelligence (CTI) reports. Knowledge graphs are used to create new templates, and the old one is modified by writing a new stimulating algorithm for identification. AI in this complex environment will significantly increase the precision of threat detection, automate human intervention, and demonstrate the breakthroughs of the AI field in cyber security [31].

5.6. Difficulties Adhering AI Perfection

AI technology has the potential to revolutionize various fields, including applications, algorithms, and human society. However, addressing all cyber security problems facing organizations requires overcoming certain challenges and constraints. One significant issue is the quality of data and the availability of vast amounts of data for effective AI solutions. Without adequate data, AI may ignore crises or collapse the victim's system security. To address this issue, recent models and approaches have focused on using non-usual dimension and size data for threat detection and prediction. Algorithms like Transfer learning and

data Augmentation can train specific data patterns and assist in the detection process, but their disadvantages include insufficient training patterns that can influence performance in real-time test cases. Few-shot learning through active learning can function on real-time data with small datasets, but it still requires many data variations to comply with IT incidents with continually changing strategies to infiltrate and harm the organization's normal operation [32].

The proposed model categories regard high-level data of disparate classes as category-agnostic representations that can be well sampled. However, these models may be derailed from real-world novel data patterns or complex attack scenarios, which are some of the approaches for sophisticated attacks. Continuous examination of the base raw data and advanced detection and prevention software increases the computational cost of transactional processing and poses tough security and stability requirements.

According to statistics released in 2023, more than 80 million data breaches took place with botnets, and the other statistically assisted by that is estimated to be more than 80 million by now. AI systems have many problems manipulating and executing users' wishes in real time, and the measure 'detect and prevent' as soon as needed is a crucial factor that causes a small lag between detection and prevention. Traditional complex models like genetic algorithms and ant colony optimization were considered efficient in this task, but the trained dataset with insufficient diversity was not included in the analysis to respond more comprehensively to the speed versus accuracy problem.

Undetected attacks, such as Advanced Persistent Threats (APT), are the most general attacks directed at a system or infrastructure, and their signatures-based AI techniques are not easy to mitigate. These covert, advanced, and persistent types of attacks can cause major destruction to crucial infrastructures, making traditional AI-based methods difficult to defend against [33].

Applying game theory as a different approach to address multiple attackers and multiple defenders instead of the traditional approach of focusing on single attacker-single defender strategy is suggested, but it has a higher level of complexity and requires more computational power. The ethics of AI and biases in cyber security are more complicated than many tend to think, as ML, DL, and NN models are black box constructs, making them difficult to understand and use in safety-critical systems or critical infrastructures.

In conclusion, the rapid evolution of AI technology has led to its misuse, becoming a major danger to organizational networks and critical systems. To prevent AI from becoming a Bogeyman rather than a blessing in cyber security, a system of monitoring, regulation, and control is needed. Raising knowledge about threats related to vulnerabilities would also be greatly appreciated for maintaining adaptive and proactive defense against the pernicious section of AI applications.

6. Threats and Vulnerabilities

AI appliances offer new opportunities for reliable connections and secure net-

work monitoring, but they also raise concerns about cybercrime. Cybercriminals often exploit the modularity of AI to create fear in the automatic industry by using techniques designed to keep the system invincible [1]. The modularity of AI can be exploited, and the operationalization can be tailored to threats, destruction, and risk rather than security, safety, and thriving situations. AI methods have vulnerabilities that lead to a spectrum of security concerns. Attackers can manipulate algorithms, make them perform unnaturally, and launch adversarial attacks. AI has become a helpful tool and often serves a legitimate purpose, but cybercriminals often use it for personal gain, ignoring its purpose. AI-targeted attacks have gained a high profile due to the sensitivity of humans and the gentle memory of individuals. AI-incorporated attacks are often used to use an offense against a system since they stand in front of defenses that can identify, prevent, or mitigate the threat [34].

AI bias can be pronounced because it is trained on specific data with a limit in category distribution, elevating vulnerable areas and allowing hostile entities to exploit any considered vulnerability in the integrated system.

To make AI models more robust to various threats, including adversarial ones, AI security researchers must consider adversarial learning or deep learning models. These models operate under the premise established by labels and feature sets for diagnosing events' growing and abnormal patterns. The adversary can gain insights into relevance, specificity, and usability by studying the system model, monitoring learners, and detecting control requirements. Security systems cannot identify anomalies as they are programmed to make it through scanners during an attack (red team effort). The attacker can further antagonize this situation by changing the system model, leading to severe degradation or denial of service. The vulnerability of these systems allows the threat actor to slip through and damage critical infrastructure and services.

Defense model-building should involve adversarial training, exposure to possible attack scenarios during development, and testing to achieve efficiency and safety. Although the application of a DL-based IDS system in distinguishing tiny deviations made in normal and abnormal events is considered the most challenging thing, it still has security issues as an adversarial attack on these NIDS systems shows that the identification mechanism is not sound enough, causing numerous problems. In the IoT world, a highly successful attack has appeared to be a significant blow for the market of security structures that assess the effectiveness of the environment.

6.1. Privacy Concerns

Privacy issues are often raised when AI tools work with sensitive data, exposing individual data and potentially leading to privacy infringements. The development of AI instruments has exacerbated this anxiety by offering more inferences from aggregated data. Deep learning (DL) models are widely used for cyber security purposes due to their ability to extract essential features and provide ac-

curate results. However, the learning process considers many data sets, teaching the model to distinguish different designs using patterns. When a model knowingly or unknowingly identifies core features or facts, attackers can target the details of these models, including the security of personal and sensitive information [35]. This could lead to identity theft or financial fraud. Collaborative DL allows multiple stakeholders to share and learn from available DL model parameters, but there is a possibility that an adversarial network (GAN) can exploit the provided incomplete parameters during training. GAN is a combination of two neural network models: the generator and the discriminator-based network. The generator aims to create data naturally, while the discriminator identifies the difference between developed and original data.

The covert learning collaborative DL model demonstrates how GAN encrypts and regenerates data from shared and hidden parameters during attacks. Attackers don't have to touch the parameters of the shared model directly, but differential privacy is attempted to protect shared parameters. However, with GAN attacks, anonymity can be viewed as a relatively straightforward process, as the GAN model can mimic data using a separate parameter set without stealing any information [36].

To address this issue, stringent data protection measures, such as data anonymization, encryption, and data storage and access protocols, must be introduced within the system. AI algorithms must be built to manage security dilemmas while ensuring privacy simultaneously.

6.2. Potential Misuse of AI

Artificial intelligence technologies could be utilized to orchestrate cyber-attacks and extend their functional capability while protecting perpetrators' political and economic objectives via AI-based detection and deception strategies. AI can also be used as a framework to produce malware that evolves with the security mechanisms; it can be used to deploy phishing attacks of enormous scale that are so successful, and even fake news where audio and visual content can be manipulated. AI tools and techniques are increasingly accessible to cybercriminals, which means cyber-attacks are becoming more enhanced and targeted, fuelling the challenges of cybercrime attrition. It requires a hands-on Cyber security approach, where inspiring defensive strategies are being developed continuously to thwart AI sword threats [37].

6.3. Socio Economic Effect

In addition, we need to understand that the advancement and development caused by the use of technology revolutionized the standards of human life and resulted in the emergence of novel roles. Since AI will affect not only technological innovations but also various facets of the world's society, experiencing evolution, the technological revolution is not the only one that gets troubled. The upcoming paragraph will discuss the shifts in interactive practices and social

life transition in the digital atmosphere, as well as demonstrate in the future how a revolution in technology will change society. Besides, the range of influence could stimulate questions within workers about the reason for retaining job-related skills and the risk of total eradication of the workforce segment by the effect of revolution. These data on political changes and the workforce show us the direction of the role of AI in the modern digital era [38].

6.4. Social Impact

AI technologies are increasingly transforming the way people exchange personal data, offering numerous benefits such as affordability, security, and convenience [39]. However, not all efforts to protect private data, such as genetic traits and health records, are effective. Some security measures can be effective in protecting personal data, but the complex architecture of these devices and the open nature of voice channels make them vulnerable to security risks. The use of chatGPT in AI-enhanced technology may also lead to malicious attacks such as social engineering, automated hacking, and malicious software development. Additionally, cyber security technology development with AI cannot be equally distributed among organizations due to insufficient resources and financial constraints. Cyber thieves may focus on less secure networks, leading to cyber-attacks [40].

Ethical concerns have been raised regarding AI since its inception, and upholding the ethical use of AI worldwide requires an agreement between nations, workers, researchers, and individuals. Applying the same ethical principles for AI across different sectors will be challenging, as such incidents may lead to targeted cyber-attacks, further complicating the issue of security.

6.5. Cyber Security Workforce

AI is increasingly being used in Cyber security, transforming the workforce and requiring a data-literate workforce to assess and understand model conduct. This skill set ensures the safety and robustness of cyber security application designs. AI advancements have made it easier for employers to hire workers who can handle complex tasks, allowing cyber security experts to focus on critical strategic decisions and comprehensive cyber security framework management. However, AI will also perform moderate tasks with a smaller chance of essential thinking. This shift highlights the need for cyber security professionals to upgrade their skills and master them in a way that AI cannot replicate. This includes not only intelligibility but also knowledge and skills in incorporating moral, legal, and social matters in cyber security [41].

AI's role in the job process has both positive and negative consequences. It may decrease managerial roles and allow workers to participate in more professionally significant and influential work. Multidisciplinary competence is increasingly important, as experts must match technical information knowledge with insights about legal, ethical, and social issues [42].

6.6. Advantages of AI in Information Security

AI-driven systems can accelerate and precisely detect threats since they are capable of analyzing massive amounts of data that humans are unable to process. They are also capable of refining the incident response by incorporating them into the technologies, making the Organizations respond more quickly and effectively. AI can skyrocket the effective administration by means of automatizing a routine surveillance work, which releases professionals to emphasize on crucial matters. It may also be empowered to detect user behavior irregularities pointing out at unsupervised access or insider threats. Lastly, AI has the capability to further strengthen security by applying an advanced analysis and detection of continuously changing threats and weaknesses.

6.7. Drawbacks of AI in Information Security

The AI use in cyber security may pose challenges from the viewpoint of employees specializing in this sphere because AI systems do not perform without human programming and intelligence. One critical issue is a possible existence of false positives and negatives, which may be followed by unsuccessful attack prevention and malicious programs detection. Among the next important issues are data privacy concerns due to the need for large datasets for training; potential biases in training data and overran This exemplifies the challenges which can be met by firms with a tight budget, data privacy and legal issues can be emerged herein and there may be chance for in bias decision making. Over-trusting AI can induce ease of life and less human care, which change organizations power to wards new hazards.

6.8. Negative Consequences of AI in Information Security

The fact that AI can amplify the efforts of a cybercriminal, resulting in more complex and difficult to detect attacks, is concerning. The AI can said to have the potential of functioning likewise for the cyber-attacks, thus intensifying the number of the artificial intelligence cyber-weapons produced. Misuse or mismanagement of AI among cyber security measures might diminish trust and completely negate the (at all) possible benefits if such technology is used. Potential side-effects, including justification of disparities or creating instability in the seriously security processes, can emerge. However, ethical questions might be brought up in terms of how the AI would be used for surveillance or as a legion in security contexts.

6.9. AI in Network Intrusion Detection

AI machines are used in network intrusion detection systems (NIDS) to identify unauthorized access and malicious activities in computer networks. AI algorithms analyze network traffic to detect patterns indicative of suspicious behavior, such as port scanning, denial-of-service attacks, and data exfiltration. By continuously monitoring network traffic and analyzing behavior patterns,

AI-powered NIDS can detect and respond to threats in real-time, enhancing overall network security.

6.9.1. Advantages

AI machines that are applied in network intrusion detection systems (NIDS) aim at spotting unprivileged access and attacker behavior in hosting computers. AI which process network traffic to detect anomalies or usage signatures similar to those of attackers such as port scanning, denial-of-service attacks, and data exfiltration is an example of such a system. It is by proactively checking network traffic and examining the patterns of behavior that AI-based NIDSs can detect and react to security breaches in real time, thus enhancing overall network security.

6.9.2. Limitation

AI based NIDS might involve excessive financial and material resources as well as they are complex in nature due to their need for advanced computational resources. In addition to that, it is a complex phenomenon that comprises of the set of specialized skills; at the same time, these skills may become obstacles for an organization which has only a limited Cyber security expertise, and result in the possibility of false negatives.

6.10. AI in Malware Detection

The AI machines also serve to detect malware by their scan and they detect various illnesses such as viruses, worms, Trojans, and ransomware. The malware detection AI-systems perform analysis of file attributes, behavior patterns, and networking traffic in order to detect malicious activities. Through exploring machine learning algorithms, AI-enabled malware detection systems can autonomously stop threats of malware before they lead to damage in technological developments.

6.10.1. Advantages

Using AI for malware detection, you will examine the behavior of the files and applications that detect malware anomalies. They use heuristics to discover malware that does not have any predefined specific signatures and they are thus efficient against zero-day attacks. Consequently, these solutions greatly reduce security breach risks. Autonpixil operation would do this by isolating and deleting wrong and hacking files which reduces the user's manual activity. They represent a suitable approach for cyber protection at the enterprises of the highest level because of their ability to sort loads and files of large capacity.

6.10.2. Limitation

The use of AI throughout the malware detection system may further increase the workload, thus, thereby resulting in delays. AI algorithms may also be run against genuine files which could be deemed as malware, resulting in unnecessary alarms and errors. Also, the immoral actors can employ emerging adversarial tactics to bypass these programs.

6.11. AI in Security Gap Detection

AI machines play a role in identifying security vulnerabilities in the systems and networks like misconfigurations, ads-on and access control errors. AI-powered systems for security gaps detection scan configurations of systems, traffic, and behavior of the users to pinpoint vulnerabilities present. With the discovery of security flaws and their mitigation AI-driven solutions may increase overall Cyber security posture and reminisce threats of broken security.

6.11.1. Advantages

AI protocols are capable of automatic vulnerability assessment tasks, decrease time spent in the manual process and can continuously focus on monitoring systems for security leaks. This risk management practice prioritizes the risks according to their severity and the degree, of course, of impact providing organizations space to fight the most critical issues at the first place. AI-enabled systems can also naturally be able to shift security measures in line with emerging threats and the different vulnerability means, therefore the increase in protection level of the whole system. It is through such mechanism that the security measures equally become efficient and effective.

6.11.2. Limitation

Boolean or rule-based methods can produce false positives and therefore will cause unnecessary system security concerns in monitoring legal system configurations. Organizations with small cyber-security departments may be inadequate to adopt Artificial Intelligence software due to insufficient staff with expertise in cyber-security. The excessive reliance on AI may result in complacency and less supervision of humans, and it might give rise to new enterprises being successful only when they are open to changes.

7. Application of AI in Cyber Security

In this segment, we closely examine the role of AI in the arena of cyber security and we are taking into account AI's functionality in detecting and preventing threats, anomaly detection, incident response, vulnerability management, user authentication, security analytics, and threat intelligence. We go through the role AI plays in each of the important areas mentioned to increase information security and improve cyber security.

7.1. Logical Relation between Sections

The connection between Section 2 and Section 3 could be viewed as one flowing from the previous section to the next, or already expressed and now elaborated further. Part 2 offers such a baseline by explaining the advantages, disadvantages and ethical issues of AI in information security in a more comprehensive way. AI is a multifaceted concept that entails both advantages and disadvantages; this section provides an all-inclusive presentation of these.

Section 3 in this paper which is a follow-up of the principles covered in sec-

tion 2 will focus that how these positive, negative outcomes and problems come out in given different applications of AI in cyber security. This way, section 3, that is each application presented there, can be seen as a practical solution or an example of a broader concept that had been explained in section 2.

Similarly, AI's application on the stakeholder for threat detection and prevention in the Section 3 could be considered a demonstrated advantage related to Section 2, such as better threat detection and sooner incident response. Consequently, the phenomena of false alarms and misdiagnoses that were already mentioned under Section 2, namely, AI issues and limitations, should be considered when dealing with anomaly detection in incident response by Section 3.

7.2. Opportunities in AI

AI strengthens a decision-making process by studying a significant amount of data, determining patterns, and trends. It helps in achieving significance through the automatic accomplishment of routine duties and lessening of manual interventions. AI helps bring down costs by reducing expenses via automation of the processes and allocation of resources. AI-based chatbots and personal digital assistants lead to better user experience, because they provide a personal message. Besides that, AI enables predictive analytics, the ability predicting future behaviors using data from the past, in order to make business decisions on a solid foundations.

7.3. Logical Relation between Sections

Artificial intelligence fosters innovations in a range of sectors including healthcare, finance, manufacturing, retail and transportation. Healthcare uses AI mostly for personalized treatment plans, early diagnosis and treatment of diseases and medical image analysis. In finance, AI is being implemented for fraud detection, risk management and also, algorithmic trading. In production, AI is applied to predictive maintenance, quality control and, supply chain management. In retail, such as marketing, inventory control and customer service, AI is used for purposes of personalization. Transportation gets AI services for autonomous vehicles, prioritized lanes, and route planning. This encourages for a decrease in congestion and makes a significant contribution in many fields. Through all of that, AI is reconfiguring industries in many ways, like increasing efficiency, cutting costs down, and improving consumer experience.

8. Logical Relation between Sections

AI algorithms require high-quality data for training and the shortcoming may be just displaying the bias or inaccuracy during the period of operation. To consider this, data diversity and quality should be in place and algorithm audits should be conducted often for bias comparison. Cultivate explainable AI methods to enhance transparency and publicize-led accountability. Resolve ethical and legal problems such as security violations and biased outputs by using peaceful meas-

ures and legislative frameworks for AI creation. Increasing security strength and redundancy with rigorous testing, validating, and adversarial training is essential. Overcome scalability obstacles and resource constraints by generating advanced algorithms and architectures capable of effectively scaling and resourcing uptake. These measures are put in place to help mitigate ethics and beneficence-related deficiencies with AI systems.

Countermeasures for Difficulties in Adhering to AI Perfection

To build high quality AI data, utilize data preprocessing with techniques such as cleaning, normalization, and augmentation, along with bias detection testing. Improve model interpretability by the method of feature importance analysis and presenting the model visually. Bend the rules of ethical and legal compliance is unnecessary for the AI systems by implementing privacy protection mechanisms like data anonymization and encryption. Include robustness actions such as input validation and dissimilarity detection in order to shield against attacks. Keep security updated as per vulnerability identification to protect against threats. Apply scalable AI algorithms with resource use optimization that is done by load balancing and concurrent processing.

9. Future Directions

AI has been integrated into security tools and offers sophisticated defense systems, but it faces new complex challenges. The arms race in AI security measures and AI-based attacks requires progressive and innovative approaches to cyber security. Fast-paced R&D, ethical implications, robust privacy protection programs, and a resilient AI model are critical to managing this territory. One possible future of Cyber security lies in identifying, addressing, and addressing these challenges before they worsen [43].

One approach to conceptualizing AI-based technologies is an association with quantum systems, such as Quantum Computing. This can replace or modify existing cryptography methods, providing quantum-resistant algorithms and continued protection from future dangers. AI in cyber security targets should continue upgrading QKD and QKD encryption techniques to provide accurate security for modern moving attacks. Blockchain is another disruptive action that ensures integrity management systems based on security measures, transparency, and immutability of handled data to efficiently organize IoT networks. Although AI-based security models have struggled with scalability problems, the combination of Quantum computing and Blockchain can be a better prospect for improvement.

Explainable Artificial Intelligence (XAI) is significant in cyber systems AI, solving the problem of opaqueness of deep learning algorithms and making AI's decision process more transparent and understandable. This multi-faceted approach provides unprecedented capabilities, including explainable models, post-hoc explanations, visualization tools, feature attribution, model simplifica-

tion, natural language explanations, prototypes, and criticisms. Achieving explainable AI models in Cyber security builds confidence, fairness, and the power of AI systems in which the AI models are deployed [44].

10. Conclusion

Moreover, AI was recognized, and Saratov attained different applications in many domains. Apart from the cyber security field, the application of AI techniques is insecure because security is the only moving barrier that can block the way of personal information and services for the bad actors trying to get access to it. Thus, technologists of all experience levels should know AI capabilities and shortcomings to make the right choices. The foremost purpose of this work is to portray and comprehend the limitless possibilities of AI on the threshold of the Cyber security field and, on the other hand, to know the perils associated with this if the offered restrictions are ignored during the implementation phase. Furthermore, the evolution of AI will be provided within the research to assist researchers in envisioning how changes may have influenced the arrival of current AI technologies. Also, this information makes it much easier for researchers to connect the changes with the possibility of AI going forward in the world, as many researchers, professionals, and innovators did. Security of AI can be maintained by the development of cyber-attacks corresponding to the advancement of AI technology to the development level of the attacks if the development of AI can be equivalent to the attacks. If these applications incorporate other technologies, particularly cyber security capabilities, threat actors will be defended effectively, or the risks can be lessened significantly. To continue, the effect of AI on Cyber security in the other domains is seen to be transforming the workforce, in the sense that more intellectually demanding roles are filled and or are oriented towards greater diversity in skills, and also creating possibilities for development and a hiccup-free job transition in the changing tech field [45].

Acknowledgements

It is my pleasure to express my full sense of gratitude to all those who have helped me with this research. Finally, I would like to extend a very special thanks to my supervisor for taking their time to guide and support me through the whole research process. I am also grateful to the participants who willingly told their stories and assisted us in gathering information. In addition to that I am very thankful to my family and friends who helped me through this difficult time. They never went back on their help, and I couldn't have finished this research without them. Ultimately, I express my gratitude to all scientists and experts that contributed to my research and sparked inspiration.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Ness, S., Shepherd, N.J. and Xuan, T.R. (2023) Synergy between AI and Robotics: A Comprehensive Integration. *Asian Journal of Research in Computer Science*, **16**, 80-94. https://doi.org/10.9734/ajrcos/2023/v16i4372

[2] Khinvasara, T., Ness, S. and Tzenios, N. (2023) Risk Management in Medical Device Industry. *Journal of Engineering Research and Reports*, **25**, 130-140. https://doi.org/10.9734/jerr/2023/v25i8965

[3] Xuan, T. and Ness, S. (2023) Integration of Blockchain and AI: Exploring Application in the Digital Business. *Journal of Engineering Research and Reports*, **25**, 20-39. https://doi.org/10.9734/jerr/2023/v25i8955

[4] Nasnodkar, S., Cinar, B. and Ness, S. (2023) Artificial Intelligence in Toxicology and Pharmacology. *Journal of Engineering Research and Reports*, **25**, 192-206. https://doi.org/10.9734/jerr/2023/v25i7952

[5] Capuano, N., Fenza, G., Loia V. and Stanzione, C. (2022) Explainable Artificial Intelligence in CyberSecurity: A Survey. *IEEE Access*, **10**, 93575-93600. https://doi.org/10.1109/ACCESS.2022.3204171

[6] Edu, J.S., Such, J.M. and Suarez-Tangil, G. (2020) Smart Home Personal Assistants: A Security and Privacy Review. *ACM Computing Surveys*, **53**, Article No. 116. https://doi.org/10.1145/3412383

[7] Gupta, M., Akiri, C., Aryal, K., Parker, E. and Praharaj, L. (2023) From ChatGPT to ThreatGPT: Impact of Generative AI in Cyber Security and Privacy. *IEEE Access*, **11**, 80218-80245. https://doi.org/10.1109/ACCESS.2023.3300381

[8] Budzinski, O., Noskova, V. and Zhang, X. (2019) The Brave New World of Digital Personal Assistants: Benefits and Challenges from an Economic Perspective. *NETNOMICS*: *Economic Research and Electronic Networking*, **20**, 177-194. https://doi.org/10.1007/s11066-019-09133-4

[9] Hussain, S., Neekhara, P., Jere, M., Koushanfar, F. and McAuley, J. (2021) Adversarial Deepfake: Evaluating Vulnerability of Deepfake Detectors to Adversarial Examples. 2021 *IEEE Winter Conference on Applications of Computer Vision* (*WACV*), Waikoloa, 3-8 January 2021, 3347-3356. https://doi.org/10.1109/WACV48630.2021.00339

[10] Basit, A., Zafar, M., Liu, X., Javed, A.R., Jalil, Z. and Kifayat, K. (2021) A Comprehensive Survey of AI-Enabled Phishing Attacks Detection Techniques. *Telecommunication Systems*, **76**, 139-154. https://doi.org/10.1007/s11235-020-00733-2

[11] Hitaj, B., Ateniese, G. and Perez-Cruz, F. (2017) Deep Models under the GAN: Information Leakage from Collaborative Deep Learning. *Proceedings of the* 2017 *ACM SIGSAC Conference on Computer and Communications Security*, Dallas, 30 October-3 November 2017, 603-618. https://doi.org/10.1145/3133956.3134012

[12] Qiu, H., Dong, T., Zhang, T., Lu, J., Memmi, G. and Qiu, M. (2020) Adversarial Attacks against Network Intrusion Detection in IoT Systems. *IEEE Internet of Things Journal*, **8**, 10327-10335. https://doi.org/10.1109/JIOT.2020.3048038

[13] Rosenberg, I., Shabtai, A., Elovici, Y. and Rokach, L. (2021) Adversarial Machine Learning Attacks and Defense Methods in the Cyber Security Domain. *ACM Computing Surveys*, **54**, Article No. 108. https://doi.org/10.1145/3453158

[14] Abdelkhalek, M., Ravikumar, G. and Govindarasu, M. (2022) ML-Based Anomaly Detection System for DER Communication in Smart Grid. 2022 *IEEE Power & Energy Society Innovative Smart Grid Technologies Conference* (*ISGT*), New Orleans, 24-28 April 2022, 1-5. https://doi.org/10.1109/ISGT50606.2022.9817481

[15] Mehedi, S.T., Anwar, A., Rahman, Z., Ahmed, K. and Islam, R. (2022) Dependable Intrusion Detection System for IoT: A Deep Transfer Learning Based Approach. *IEEE Transactions on Industrial Informatics*, **19**, 1006-1017. https://doi.org/10.1109/TII.2022.3164770

[16] Li, Z., Zeng, J., Chen, Y. and Liang, Z. (2022) AttacKG: Constructing Technique Knowledge Graph from Cyber Threat Intelligence Reports. *Computer Security-ESORICS* 2022, Copenhagen, 26-30 September 2022, 589-609. https://doi.org/10.1007/978-3-031-17140-6_29

[17] Devi Priya, V.S. and Chakkaravarthy, S.S. (2023) Containerized Cloud Based Honeypot Deception for Tracking Attackers. *Scientific Reports*, **13**, Article No. 1437. https://doi.org/10.1038/s41598-023-28613-0

[18] Skulimowski, A.M.J. and Łydek, P. (2022) Adaptive Design of a Cyber-Physical System for Industrial Risk Management Decision Support. 2022 17*th International Conference on Control*, *Automation*, *Robotics and Vision* (*ICARCV*), Singapore, 11-13 December 2022, 90-97. https://doi.org/10.1109/ICARCV57592.2022.10004251

[19] Ansari, M.F., Sharma, P.K. and Dash, B. (2022) Prevention of Phishing Attacks Using AI-Based Cyber Security Awareness Training. *International Journal of Smart Sensor and Adhoc Network*, **3**, 61-72.https://doi.org/10.47893/IJSSAN.2022.1221

[20] Chandre, P.R., Mahalle, P. and Shinde, G. (2022) Intrusion Prevention System Using Convolutional Neural Network for Wireless Sensor Network. *International Journal of Artificial Intelligence*, **11**, 504-515. https://doi.org/10.11591/ijai.v11.i2.pp504-515

[21] Elhadad, A., Alanazi, F., Taloba, A.I., Abozeid, A., *et al.* (2022) Fog Computing Service in the Healthcare Monitoring System for Managing the Real-Time Notification. *Journal of Healthcare Engineering*, **2022**, Article ID: 5337733. https://doi.org/10.1155/2022/5337733

[22] Mukherjee, P., Pradhan, C., Tripathy, H.K. and Gaber, T. (2023) Kryptoschain—A Blockchain-Inspired, AI-Combined, DNA-Encrypted Secure Information Exchange Scheme. *Electronics*, **12**, Article 493. https://doi.org/10.3390/electronics12030493

[23] Stergiou, C.L. and Psannis, K.E. (2022) Digital Twin Intelligent System for Industrial Iot-Based Big Data Management and Analysis in Cloud. *Virtual Reality & Intelligent Hardware*, **4**, 279-291. https://doi.org/10.1016/j.vrih.2022.05.003

[24] Hassan, S., Wang, J., Kontovas, C. and Bashir, M. (2022) Modified FMEA Hazard Identification for Cross-Country Petroleum Pipeline Using Fuzzy Rule Base and Approximate Reasoning. *Journal of Loss Prevention in the Process Industries*, **74**, Article 104616. https://doi.org/10.1016/j.jlp.2021.104616

[25] Waleed, A., Jamali, A.F. and Masood, A. (2022) Which Open-Source IDS? Snort, Suricata or Zeek. *Computer Networks*, **213**, Article 109116. https://doi.org/10.1016/j.comnet.2022.109116

[26] Malek, Z.S., Trivedi, B. and Shah, A. (2020) User Behavior Pattern-Signature Based Intrusion Detection. 2020 *Fourth World Conference on Smart Trends in Systems*, *Security and Sustainability* (*WorldS4*), London, 27-28 July 2020, 549-552. https://doi.org/10.1109/WorldS450073.2020.9210368

[27] Ferrag, M.A., Maglaras, L., Ahmim, A., Derdour, M. and Janicke, H. (2020) RDTIDS: Rules and Decision Tree-Based Intrusion Detection System for Internet-of-Things Networks. *Future Internet*, **12**, Article 44. https://doi.org/10.3390/fi12030044

[28] Tripathi, K.P. (2011) A Review on Knowledge-Based Expert System: Concept and Architecture. *IJCA Special Issue on Artificial Intelligence Techniques Novel Ap-*

proaches & Practical Applications, **4**, 19-23.

[29] Sarker, P.S., Rafy, F., Srivastava, A.K. and Singh, R.K. (2023) Cyber Anomaly-Aware Distributed Voltage Control with Active Power Curtailment and DERs. *IEEE Transactions on Industry Applications*, **60**, 1622-1633. https://doi.org/10.1109/TIA.2023.3328850

[30] Kim, J., Kim, J. Le, H., Thu, T. and Kim, H. (2016) Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection. 2016 *International Conference on Platform Technology and Service* (*PlatCon*), Jeju, 15-17 February 2016, 1-5. https://doi.org/10.1109/PlatCon.2016.7456805

[31] Yin, C., Zhu, Y., Fei, J. and He, X. (2017) A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks. *IEEE Access*, **5**, 21954-21961. https://doi.org/10.1109/ACCESS.2017.2762418

[32] Diro, A.A. and Chilamkurti, N. (2018) Distributed Attack Detection Scheme Using Deep Learning Approach for Internet of Things. *Future Generation Computer Systems*, **82**, 761-768. https://doi.org/10.1016/j.future.2017.08.043

[33] Tang, T.A., Mhamdi, L., McLernon, D., Zaidi, S.A.R., Ghogho, M. and El Moussa, F. (2020) DeepIDS: Deep Learning Approach for Intrusion Detection in Software Defined Networking. *Electronics*, **9**, Article 1533. https://doi.org/10.3390/electronics9091533

[34] Xiao, K.H., Wu, E., Guo, J., Xu, C. and Wang, Y. (2021) Transformer in Transformer. *Advances in Neural Information Processing Systems*, **34**, 15908-15919.

[35] Du, S., Lee, J., Tian, Y., Singh, A. and Poczos, B. (2018) Gradient Descent Learns One-Hidden-Layer CNN: Don't be Afraid of Spurious Local Minima. *Proceedings of the* 35*th International Conference on Machine Learning*, Stockholm, 10-15 July 2018, 1339-1348.

[36] Kriegeskorte, N. and Golan, T. (2019) Neural Network Models and Deep Learning. *Current Biology*, **29**, R231-R236. https://doi.org/10.1016/j.cub.2019.02.034

[37] Bui, D.T., Tsangaratos, P., Nguyen, V.-T., Liem, N.V. and Trinh, P.T. (2020) Comparing the Prediction Performance of a Deep Learning Neural Network Model with Conventional Machine Learning Models in Landslide Susceptibility Assessment. *CATENA*, **188**, Article 104426. https://doi.org/10.1016/j.catena.2019.104426

[38] Sarker, I.H., Abushark, Y.B., Alsolami, F. and Khan, A.I. (2020) IntruDTree: A Machine Learning Based Cyber Security Intrusion Detection Model. *Symmetry*, **12**, Article 754. https://doi.org/10.3390/sym12050754

[39] Srinidhi, C.L., Ciga, O. and Martel, A.L. (2021) Deep Neural Network Models for Computational Histopathology: A Survey. *Medical Image Analysis*, **67**, Article 101813. https://doi.org/10.1016/j.media.2020.101813

[40] Wang, H., Lei, Z., Zhang, X., Zhou, B. and Peng, J. (2016) Machine Learning Basics. *Deep Learning*, **1**, 98-164.

[41] Wu, Z., Zhang, H., Wang, P. and Sun, Z. (2022) RTIDS: A Robust Transformer-Based Approach for Intrusion Detection System. *IEEE Access*, **10**, 64375-64387. https://doi.org/10.1109/ACCESS.2022.3182333

[42] Rieck, K., Trinius, P., Willems, C. and Holz, T. (2011) Automatic Analysis of Malware Behavior Using Machine Learning. *Journal of Computer Security*, **19**, 639-668. https://doi.org/10.3233/JCS-2010-0410

[43] Costa, P.C.G., Laskey, K.B. and Alghamdi, G. (2006) Bayesian Ontologies in AI Systems. https://www.academia.edu/66571238/Bayesian_ontologies_in_AI_systems?auto=do

wnload

[44] Manikopoulos, C. and Papavassiliou, S. (2002) Network Intrusion and Fault Detection: A Statistical Anomaly Approach. *IEEE Communications Magazine*, **40**, 76-82. https://doi.org/10.1109/MCOM.2002.1039860

[45] Stuart, J.A. and Owens, J.D. (2011) Multi-GPU MapReduce on GPU Clusters. 2011 *IEEE International Parallel & Distributed Processing Symposium*, Anchorage, 16-20 May 2011, 1068-1079. https://doi.org/10.1109/IPDPS.2011.102