SURVEY

Open Access

# Advancing cybersecurity: a comprehensive review of AI-driven detection techniques

Aya H. Salem[1*], Safaa M. Azzam[1], O. E. Emam[1] and Amr A. Abohany[2]

*Correspondence:
aya.salem@fci.helwan.edu.eg

[1] Faculty of Computer and Artificial Intelligence, Helwan University, Cairo, Egypt
[2] Faculty of Computer and Information, Kafr El-Sheikh University, Cairo, Egypt

**Abstract**

As the number and cleverness of cyber-attacks keep increasing rapidly, it's more important than ever to have good ways to detect and prevent them. Recognizing cyber threats quickly and accurately is crucial because they can cause severe damage to individuals and businesses. This paper takes a close look at how we can use artificial intelligence (AI), including machine learning (ML) and deep learning (DL), alongside metaheuristic algorithms to detect cyber-attacks better. We've thoroughly examined over sixty recent studies to measure how effective these AI tools are at identifying and fighting a wide range of cyber threats. Our research includes a diverse array of cyberattacks such as malware attacks, network intrusions, spam, and others, showing that ML and DL methods, together with metaheuristic algorithms, significantly improve how well we can find and respond to cyber threats. We compare these AI methods to find out what they're good at and where they could improve, especially as we face new and changing cyber-attacks. This paper presents a straightforward framework for assessing AI Methods in cyber threat detection. Given the increasing complexity of cyber threats, enhancing AI methods and regularly ensuring strong protection is critical. We evaluate the effectiveness and the limitations of current ML and DL proposed models, in addition to the metaheuristic algorithms. Recognizing these limitations is vital for guiding future enhancements. We're pushing for smart and flexible solutions that can adapt to new challenges. The findings from our research suggest that the future of protecting against cyber-attacks will rely on continuously updating AI methods to stay ahead of hackers' latest tricks.

**Keywords:** Cyber-attacks, Artificial intelligence, Machine learning, Deep learning, Cyber security, Intrusion detection

## Introduction

In the face of evolving digital advancements such as software-defined networking (SDN), big data, and fog computing, the growth of the internet has been remarkable. However, these advancements come with significant cybersecurity challenges with major implications for critical infrastructure. Traditional security methods, with their reliance on fixed security controls like firewalls and intrusion detection and prevention systems, have struggled to keep pace with the sophisticated nature of contemporary cyber threats [1]. Deep Learning (DL) has emerged as a transformative force, unlocking new possibilities for data access, enhanced performance, and potential maximization. It has revolutionized not just artificial intelligence (AI) applications across images, voice, and behavioral analysis but has also led to groundbreaking advancements in areas including robotics, speech, and facial recognition. In the field of cybersecurity, DL has evolved to play vital roles in detecting intrusions and monitoring for malware. This represents a notable advancement from the earlier uses of machine learning (ML) [2]. While ML has shown promise, its reliance on manual feature extraction has become a noticeable limitation, particularly cybersecurity. The manual compilation of malware features for ML-based recognition is a case in point, limiting the efficiency and accuracy of threat detection to predefined features and overlooking unidentified characteristics. Consequently, ML's adeptness largely depends on the precision of feature extraction and recognition [3]. DL offers a strategic edge in cyber defense through its ability to uncover complex, nonlinear correlations within data, thus enabling the recognition of new file types and previously unknown threats. Notably, DL has propelled advancements in preventing Advanced Persistent Threat (APT) attacks, even recognizing the subtle, high-level features used in the most evasive tactics.

With the ubiquity of the Internet of Things (IoT), the explosion in network connectivity, and many associated applications, cybersecurity has become a focal point of contemporary security concerns. The need to identify a variety of cyber threats effectively and develop robust intrusion detection systems has never been more pressing [4]. Advancements in cloud computing have prompted various organizations to outsource their data and computational requirements, emphasizing the need for a secure platform, particularly within cloud-based systems. Understanding malware behavior in the context of behavioral space is pivotal to enhancing the effectiveness of traditional security measures, especially given the vast and varied nature of cybersecurity data [5]. ML stands at the forefront of automating behavior analysis through informative feature extraction from network packets, paving the way for developing sophisticated intrusion detection systems. The essence of ML is to endow computers with the capability to learn and adapt autonomously without human intervention [6]. In an era where cybersecurity encompasses a range of techniques, policies, and procedures aimed at preserving data confidentiality and integrity, the goal is to reduce the risk of attacks and safeguard against unauthorized access. With the frequency and complexity of attacks escalating, there's an acute need for systems capable of identifying significant indicators of potential breaches. Despite its complexity, DL holds the promise of delivering precise outcomes when properly trained, representing a significant step forward in cybersecurity methodologies [6]. This paper aims to explore the application of DL, ML, and Metaheuristics in modern cybersecurity practices, evaluating their effectiveness in addressing cyber-attack threats and proposing future directions for research and implementation.

### Motivation

In today's rapidly advancing digital era, the cybersecurity landscape faces unprecedented challenges due to the sophistication and frequency of modern cyber threats. Traditional security measures, such as firewalls and intrusion detection systems, rely on static controls and manual processes, which are increasingly ineffective against the dynamic and complex nature of contemporary attacks. These methods often fail to detect advanced techniques like malware, phishing, APTs, botnets, and insider threats, leading to high false-positive rates, slower response times, and a significant demand for human intervention.

The motivation for this paper stems from the urgent need to bridge the gap left by these traditional methods. By exploring AI-driven techniques, specifically DL and ML, we aim to harness the power of AI to process vast amounts of data in real time, identify intricate patterns, and adapt to new threats quickly. AI's ability to learn from data and continuously evolve makes it an invaluable tool in developing more resilient and scalable cybersecurity solutions by addressing challenges such as insider threats that pose significant risks from within the organization, making them difficult to detect using conventional methods, and botnets, which are networks of compromised devices controlled by malicious actors capable of executing coordinated attacks that overwhelm traditional defenses.

Many surveys must comprehensively cover recent studies on the three main topics of ML, DL, and metaheuristic technologies. Additionally, they often fail to address the full spectrum of cyber-attacks or include dedicated sections for future work and limitations. In contrast, our survey provides thorough coverage of these areas, including specific sections for future works and limitations. Moreover, we have listed all recent datasets alongside their corresponding studies, ensuring our survey is grounded in the latest research. This research seeks to demonstrate how AI can transform cybersecurity by providing more accurate, efficient, and adaptive defenses. By addressing the limitations of traditional methods, such as manual feature extraction and static controls, and leveraging AI's strengths, we aim to enhance the detection and mitigation of sophisticated cyber threats, including insider threats and botnets, ultimately contributing to a more secure digital environment.

### Main contribution

The primary contribution of this paper is a carefully compiled collection of recent studies on different types of cyber-attacks and their corresponding detection technologies. It outlines current challenges and research gaps, providing scholars with a clear overview and a solid foundation for further research into the use of AI to detect and mitigate cyber threats. The contributions of the research are organized into the following points:

- An evaluation of ML, DL, and metaheuristic techniques in detecting cyber-attacks, including a focused analysis on anomaly detection, classification, and analysis methodologies.
- A review and critical examination of more than sixty scholarly articles from the last four years focused on cybersecurity and the implementation of foundational AI techniques.
- The article were assessed based on several criteria: cybersecurity focus, AI techniques used, data sets, methods for reducing data complexity, ways of sorting data, comparing methods, and measuring performance.
- Evaluation of open access cyber-security datasets and primary classification of cyber-attack types found within these datasets.
- The studies' key elements are organized into comparative tables, which serve as a streamlined reference for understanding the varied approaches and outcomes.

- A conclusion that addresses the challenges encountered with these AI methods in cybersecurity and proposes future solutions.

### Paper structure

The paper is organized into six clearly defined sections to provide a systematic exploration of AI methods in cyber-attack detection. In the ’Introduction’ section, we introduce the research contributions and motivations. In the ’Background’ section, we discuss the research background and describe the main topics of the research. In the ’Literature Review’ section, we review related literature. In the ’Experiments and Setup’ section, we briefly describe the research methodology used. In the ’Results and Discussion’ section, we analyze the results and discuss the reasons behind them. Finally, in the ’Conclusion’ section, we summarize the entire text.

## Background

The growth of computer networks has transformed how societies function, leading to an increase in cyberattacks’ frequency and complexity. Cyberattacks are disruptive activities that target computer systems, networks, or data. They are usually organized and well-planned and involve synchronized steps to achieve their goals [3]. Intending to cause damage, unauthorized access, or service interruptions that cause severe data loss or financial damage and often lead to long-lasting consequences [7], these are the insider threats that represent a significant and growing segment of these attacks, usually committed by disgruntled or rogue employees who exploit their authorized access to steal data or cause harm. These threats can also emerge from intrusive applications that users accidentally install on their devices, allowing these apps to access and misuse sensitive information. Advanced behavioral anomaly detection and auto-resiliency mechanisms are being developed to combat these threats by proactively identifying and mitigating malicious actions at both the employee and application levels [8].

There is a broad spectrum of cyber-attacks that represent a variety of threats in the digital world. Insights into several critical types of these attacks are provided, as highlighted in the literature and summarized in Table 1. This information emphasizes the complexity and wide range of cyber threats, illustrating the many attacks organizations and individuals may encounter in today’s interconnected environment [3]. Botnets, another critical cyber threat, are networks of infected computers controlled by an attacker to perform coordinated malicious activities, such as DDoS attacks, data theft, and spamming. These networks can be vast, comprising thousands or even millions of compromised devices, which makes them incredibly difficult to dismantle. Botnet operators use sophisticated methods to infect devices and maintain control, continuously evolving their techniques to avoid detection [9].

The range of cyber-attack types of points to the vital need for effective cybersecurity strategies. It’s crucial to guard sensitive data and keep digital services running smoothly. Figure 1 categorizes these cyber-attack types. As cyber threats are ever-changing, it’s essential to remain alert and continuously invest in advanced security solutions. Staying ahead of cyber threats means actively adapting to new risks,

**Table 1** Shows types of cyber-attacks [10]

| Attack type | Description |
|---|---|
| Phishing | This attack involves tricking individuals into divulging sensitive information, such as login credentials and financial data, by masquerading as trustworthy in electronic communication. It’s a prevalent method for attackers to gain unauthorized access to personal or corporate data |
| Malware | Short for "malicious software," malware includes viruses, worms, Trojans, and ransomware. It is designed to cause damage or unauthorized access to computer systems and data. Malware is a broad category that encompasses various forms of harmful software deployed during cyber-attacks |
| DDOS | Denial of Service (DoS)/Distributed Denial of Service (DDoS): These attacks aim to overwhelm a system’s resources, making it unable to respond to legitimate service requests. DDoS attacks use multiple compromised computer systems as sources of attack traffic, exacerbating the scale of the assault |
| MitM | Man-in-the-Middle (MitM) Attack: this type of attack intercepts the communication between two parties without their knowledge. Attackers can steal, alter, or fabricate messages between the communicating parties, leading to data breaches or eavesdropping |
| SQL injection | This attack technique exploits vulnerabilities in a database-driven website by injecting malicious SQL statements into a query. If successful, an attacker can read, modify, and delete database information, potentially accessing sensitive data |
| Zero-Day Exploit | Refers to an attack on or before the first or "zeroth" day of a vendor becoming aware of a vulnerability. These attacks exploit vulnerabilities before they can be patched, making them particularly dangerous and difficult to defend against |
| Ransomware | Ransomware is a form of malware that encrypts the victim’s files, making them inaccessible until the attacker pays a ransom for the decryption key. It represents a direct financial threat to individuals and organizations by holding data or systems hostage |
| XSS | Cross-site Scripting (XSS) attacks use third-party web resources to run scripts in the victim’s web browser or scriptable application |
| APT | Advanced Persistent Threats (APT): When an individual or group acquires unauthorized access to a network and goes unnoticed for a long time, attackers may exfiltrate important data, obviating the need for the organization’s security staff to investigate |
| BEC | Business Email Compromise (BEC) attacks target employees with financial authority, using detailed research to trick them into sending money to the attacker’s account |
| Crypto-jacking | Crypto-jacking confidentially uses a victim’s computing resources to mine cryptocurrency, posing a hidden threat by draining organizational network resources |
| Password attack | A password attack involves trying to crack a user’s password using methods like Brute-Force, Dictionary, Rainbow Table, Credential Stuffing, Password Spraying, and Keylogger attacks, including phishing for passwords |
| Insider threat | Insider Threats stem from individuals within an organization who misuse their authorized access to the company’s systems and data |
| Botnet attack | Botnet attack involves a network of compromised computers controlled remotely by an attacker to execute coordinated malicious activities |

employing best practices, and leveraging technology to safeguard against the diverse tactics used by the attackers [11].

The cybersecurity community has strongly focused on attack detection as a cornerstone strategy in response to these growing threats. This approach comprehensively monitors network activities, system status, and usage patterns to preemptively identify and neutralize unauthorized access or attacks. Within this landscape, AI and its subsets, including ML and DL, offer promising solutions to support cybersecurity. AI’s capacity to rapidly evolve and handle large datasets makes it well-suited for identifying and responding to sophisticated cyber threats. By analyzing patterns and learning from experience, AI-based systems can detect malware, insider threats, botnets, network intrusions, phishing attempts, and other malicious activities [12].

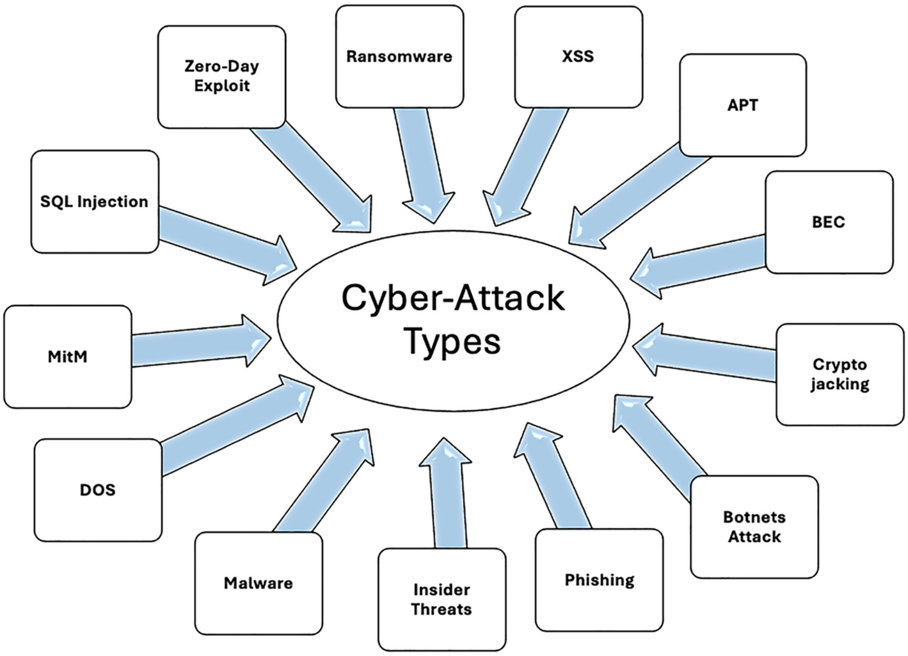

**Fig. 1** Cyber-attack types

### Artificial intelligence overview

AI delineated by pioneers such as John McCarthy in 1956 [13], refers to the science and engineering of making intelligent machines. Over the years, AI has evolved into a foundation of computer science, focusing on simulating human cognitive processes through complex mathematical algorithms. This interdisciplinary field combines elements from various domains to adopt machines that can learn, reason, and make decisions based on the data they process. Besides, it encompasses both the replication of human thought and behavior in machines, categorized respectively into thinking and acting both humanly and rationally [14]. AI applications range from simple tasks to complex problem-solving domains such as cybersecurity, where it addresses sophisticated cyber threats. This transformative technology continues to push the boundaries of what machines are capable of, aiming to enhance human capabilities and automate tasks through assisted, augmented, and autonomous intelligence [15].

The use of AI in cybersecurity is increasingly critical due to its capacity to analyze vast amounts of data rapidly, detect patterns, and identify potential threats with high efficiency. In a digital era characterized by ever-evolving cyber threats, traditional security measures often fall short in both the speed and sophistication needed to counteract modern cyberattacks, including zero-day threats [16]. AI's ability to learn from data enables the development of systems that can adapt to new, previously unknown attacks, enhancing the ability to secure information infrastructure from a broad spectrum of threats [15]. The benefits of integrating AI into cybersecurity include improved decision-making capabilities, enhanced detection of network intrusions, and the management of cyber-attack impacts. This progression in technology not only allows for real-time threat detection and response but also significantly reduces the rate of false positives, which are common in more traditional methods of cyber defense. Furthermore, AI's predictive analytics can foresee potential vulnerabilities before they are exploited, offering a proactive form of security rather than a reactive one. In essence, AI empowers cybersecurity with advanced analytical tools, making it an indispensable ally in the battle against cybercrime [17].

AI technologies encompass several approaches useful in cybersecurity, including:

- ML: Algorithms that enable computers to learn from data without explicit programming, allowing for improved threat detection and classification [18].
- DL: Advanced neural networks that can process large amounts of data and learn from experience, mimicking human brain functions to recognize complex patterns [19].

These AI techniques provide robust defenses against cyberattacks by enabling real-time monitoring, automated responses, and continuous learning to adapt to new threats. Furthermore, integrating metaheuristic algorithms with learning models offers significant advantages in the detection of cyberattacks [20]. Metaheuristic algorithms are vital in improving the efficiency and accuracy of various detection learning by enhancing the learning as they expand the search space explored during model training, potentially uncovering superior solutions that traditional methods might miss. This is particularly beneficial in cybersecurity, where the landscape and attack patterns can change rapidly. And by enabling models to adapt more dynamically to new or evolving types of cyber threats, thus enhancing the model Cybersecurity Deep: Approaches, Attacks Dataset, and Comparative Study's ability to generalize across different scenarios and datasets [21].

Advantages of Metaheuristic Algorithms in Cyber Attack Detection [20]:

- *Optimization*: Metaheuristic algorithms are better find optimal solutions to complex problems that are otherwise too challenging for conventional methods.
- *Automation*: By automating the tuning of detection parameters, these algorithms minimize the need for human intervention, making the detection process both faster and more reliable.
- *Speed*: They often achieve faster convergence to effective solutions, which is essential in time-sensitive cybersecurity environments where threats must be quickly identified and mitigated.

## Machine learning

ML is a domain that empowers computers to solve problems and interpret them without explicit programming. It forecasts outcomes by analyzing past data. This section aims to offer an overview of ML paradigms, classifications, and architectures. The learning technique consolidates various ML algorithms, which differ extensively, and categorizes them according to the nature of the tasks they perform or the complexity of their operations [22].

ML algorithms are divided into supervised, unsupervised, semi-supervised, and RL, as shown in Fig. 2. A few more categories have emerged in more detail: semi-supervised,

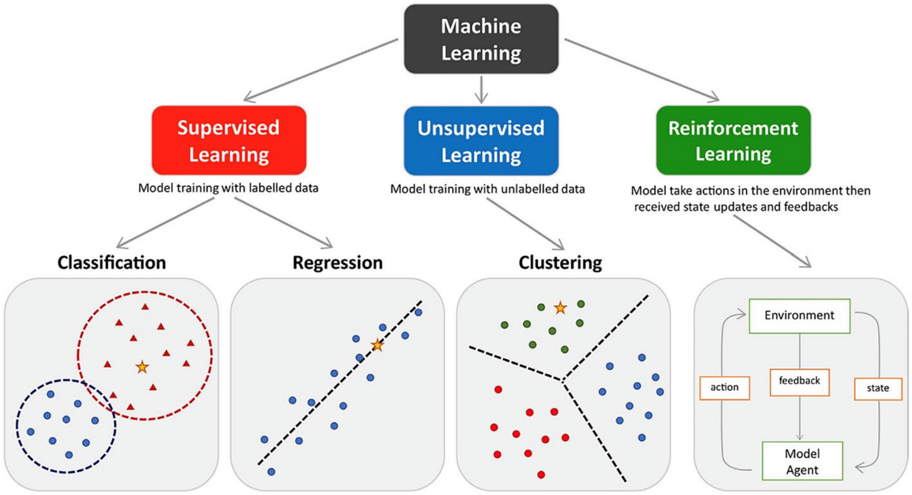

**Fig. 2** Main types of ML [23]

**Table 2** ML techniques [18]

*Supervised learning*: Supervised learning involves a set of labeled input–output pairs that guide the model during training. The process includes developing a mapping function from inputs (x) to outputs (y) by analyzing the data. Common applications include Classification and Regression tasks

*Unsupervised learning*: This approach does not utilize labeled data for training, aiming to uncover patterns or structures within the data based on its inherent characteristics. Without predetermined labels or outputs, it focuses on tasks like Dimensionality Reduction, Clustering, and Association Rule Learning. Language models in unsupervised learning are often targeted by attacks

*Reinforcement learning*: Reinforcement learning (RL) operates by interacting with an external environment and learning through trial and error. It develops predictions about future outcomes based on accumulated experiences, notably without reported privacy attacks in this learning paradigm

*Semi-supervised learning*: Combining elements of both supervised and unsupervised learning, this method uses a mix of labeled and unlabeled data to train models, enhancing interpretation with the unlabeled data before refining tasks with the labeled portion. It is commonly applied in Classification and Regression tasks

*Active learning*: Active learning strategies select training data purposefully to minimize the need for extensive labeled datasets, thereby optimizing the cost and time required to gather labeled training data

*Ensemble learning*: Ensemble learning involves merging multiple weak classifiers to create a robust classifier that makes decisions based on the aggregate predictions of individual models. Techniques such as boosting and bagging exemplify ensemble learning strategies

active, and ensemble learning, each suited for different types of data and problems as discussed in Table 2 [18].

A variety of supervised and unsupervised learning techniques have been applied to develop advanced and effective models capable of identifying and categorizing attacks. Some of the ML Algorithms are briefly described in Table 3.

## Deep learning

DL is a specialized area within ML focused on representation learning through multilayer transformations, leading to enhanced accuracy in detection and prediction tasks. In cybersecurity, DL-enhanced defense mechanisms are increasingly deployed to automate the identification of cyber threats, with these systems continuously evolving and enhancing their effectiveness over time [28].

**Table 3** ML Algorithms

| | |
|---|---|
| LR | Logistic Regression: This technique is applied to classification challenges, predicting the outcome for a categorical dependent variable. It is specifically designed for binary classification tasks, where the outcome needs to be a categorical or distinct value [24] |
| GNB | Gaussian Naive Bayes: A probabilistic classification method based on the Gaussian distribution. It operates under the assumption that each feature independently contributes to the probability of the outcome, aggregating these probabilities to determine the most likely class [25] |
| SVM | Support Vector Machine works by transforming data into a higher-dimensional space to find a separating hyperplane between different classes. This method is effective even when the data are not linearly separable in the original space [26] |
| DT | Decision Tree: this algorithm builds a model in the form of a tree structure. It splits data based on attributes, using branches to represent decisions and leaf nodes to represent outcomes, aiming to predict a target variable through simple decision rules [27] |
| ETC | Extra Trees Classifier: generates multiple DTs using random samples of the dataset and features, choosing split values randomly rather than calculating them. This randomness leads to a diverse and uncorrelated ensemble of trees [25] |
| VC | Voting Classifier: Combines the predictions of several classifiers by taking a majority vote to determine the final class. This approach leverages the strengths of various models to improve overall performance [25] |
| RF | Random Forest: An ensemble method that improves prediction accuracy and controls overfitting by averaging the predictions of numerous DT classifiers, each trained on different data samples [27] |
| KNN | K Nearest Neighbor: A method that classifies each data point based on the majority class among its KNNs. It is a form of semi-supervised learning that relies on the proximity of data points to determine their classification [24] |
| BC | Bagging Classifier: An ensemble technique that trains base classifiers on random subsets of the original dataset and aggregates their predictions to form a final prediction, either by voting or averaging [25] |
| GB | Gradient Boosting enhances predictive accuracy by combining multiple weak prediction models, typically DTs, into a stronger model. This method sequentially corrects the errors of the weak learners, leading to improved performance [25] |
| AC | AdaBoost Classifier: An ensemble boosting method that combines multiple weak learners to form a more accurate prediction model. It focuses on correcting the errors of individual learners by adjusting their contributions to the final model [25] |
| XB | XGBoost: An optimized gradient boosting library that is efficient and scalable for training. It employs ensemble learning to combine the predictions of weak models, enhancing its capacity to handle large datasets and achieve superior performance [25] |

DL's basic structure consists of the input layer, hidden layer/s, and output layer, depending on the computational layers' there are several models, as shown in Table 4, which encompass a range of predictive models based on Artificial Neural Networks (ANNs), which are networks of interconnected neurons transmitting information among themselves. The distinction between Deep Neural Networks (DNNs) and simpler single-hidden-layer neural networks lies in the DNNs' substantial depth, marked by many hidden layers facilitating intricate pattern recognition. A DNN typically comprises an input layer, several hidden layers, and an output layer, with each layer containing neurons that output nonlinear responses, as shown in Fig. 3. Data progresses from the input layer to hidden layers, where neurons compute weighted sums and apply activation functions like ReLU or tanh, before reaching the output layer for final result presentation [19].

These architectures have broad applications in cybersecurity, from detecting false data injection and network anomalies to developing advanced defense strategies and intrusion detection systems.

## Metaheuristic

Metaheuristic algorithms are optimization methods that aim to find optimal or near-optimal solutions to complex problems by exploring and exploiting the search space.

**Table 4** DL Learning models [29]

| | |
|---|---|
| CNN | Convolutional Neural Networks (CNNs) are tailored for processing multi-array data structures, such as images or sequences, using local connections and shared weights for efficiency. CNNs often include convolutional and pooling layers, culminating in fully connected layers, and are used in cybersecurity for tasks like user authentication and malware detection, as shown in Fig. 4 |
| RNN | Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks excel in learning sequential data patterns, incorporating memory elements to handle temporal dependencies. LSTMs address RNNs' vanishing gradient problem by using cell memory units with gate mechanisms, making them suited for analyzing time-dependent data |
| FNN | Feedforward Neural Networks (FNN) process data from input to output layers through hidden layers in a single direction. They are simple and effective for tasks like image and speech recognition but not suitable for sequential data |
| GRU | Gated Recurrent Units (GRU) are RNNs with update and reset gates to manage long-term dependencies and mitigate the vanishing gradient problem. They excel in language modeling, speech recognition, and time series prediction |
| VAE | Variational Autoencoders (VAE) use a probabilistic approach to encode data into a latent space for generating new, similar data samples. They are used in image generation, anomaly detection, and data augmentation |
| GNN | Graph Neural Networks (GNN) process graph-structured data by aggregating and updating node features through message-passing mechanisms. They are effective for node classification, link prediction, and graph classification in various domains |
| AE | Autoencoders (AEs) aim to reconstruct their input at the output, utilizing encoder and decoder components for dimensionality reduction and feature learning. AEs are employed in unsupervised learning and are useful for tasks such as intrusion and spam detection |
| DBN | Deep Belief Networks (DBNs) generate models from stacked Restricted Boltzmann Machines (RBMs), learning data reconstruction in an unsupervised manner and enabling classification tasks through additional training |
| GAN | Generative Adversarial Networks (GANs) consist of generative and discriminative networks that learn to produce data indistinguishable from real data, addressing issues like data imbalance in cybersecurity by generating synthetic data samples |
| DRL | Deep Reinforcement Learning (DRL) combines RL with DNNs to develop agents that optimize long-term rewards through actions, effectively addressing dynamic and complex security challenges |
| LSTM | Long Short-Term Memory (LSTM) is a specialized RNN used for sequence prediction tasks, such as language processing or time series analysis. It includes three gates—input, forget, and output—to control the flow of information, solving the vanishing gradient problem of traditional RNNs. While effective for time series prediction and classification, LSTMs require extensive training and are unsuitable for non-sequential data |

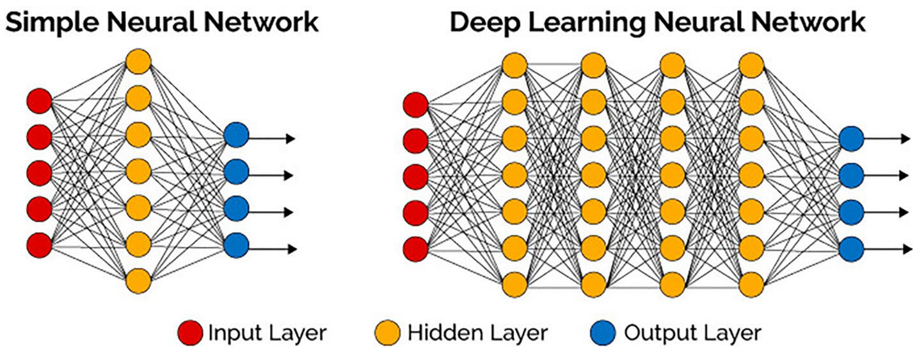

**Fig. 3** NN VS DNN [30]

They are derivative-free, flexible, and effective in avoiding local optima. These algorithms initiate their optimization process with one or multiple randomly generated solutions and do not require derivative calculations like gradient-based methods.

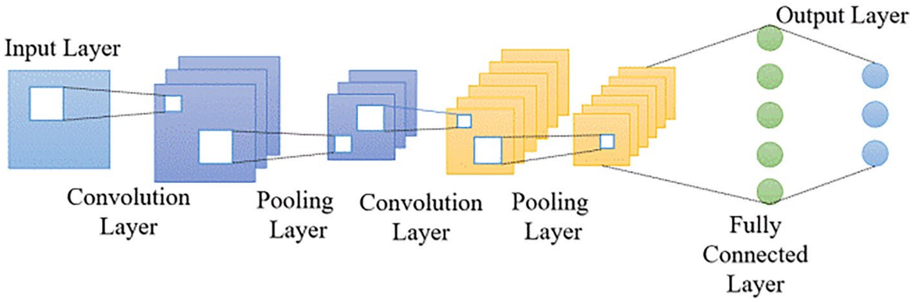

**Fig. 4** CNN structure [31]

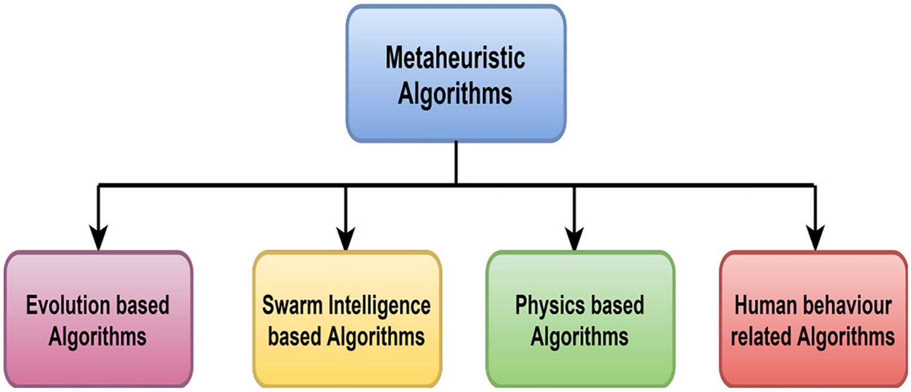
Metaheuristic Algorithms
Evolution based Algorithms
Swarm Intelligence based Algorithms
Physics based Algorithms
Human behaviour related Algorithms

**Fig. 5** Metaheuristic algorithms classification

Metaheuristics make a balance between exploration (investigating promising search space) and exploitation (local search of promising areas) [32].

Metaheuristic algorithms are sophisticated global optimization strategies derived from simulations and nature-inspired methodologies. These strategies, inspired by the social and swarm behaviors observed in various species, such as fish, birds, ants, and other animals, have been recognized for their effectiveness over several decades. The collective intelligence demonstrated by these creatures in solving complex problems efficiently has paved the way for the development of optimization algorithms. These algorithms have demonstrated considerable success across a diverse range of real-world optimization challenges, leveraging the principles of collective behavior to derive optimal solutions [33].

The metaheuristic algorithms are classified into four main categories as shown in Fig. 5: evolution-based, swarm intelligence-based, physics-based, and human-related algorithms. This classification is based on their behavior and inspiration sources, ranging from natural processes and animal behavior to physics principles and human activities as shown in Table 5.

To conclude the background section, we've dug into the complex world of cybersecurity, highlighting the cyberattacks and the critical need for effective detection mechanisms. Integrating cutting-edge technologies such as AI, ML, and DL with metaheuristic algorithms has been showcased as an efficient approach. This powerful

**Table 5** Metaheuristic classifications [21]

| |
|---|
| *Evolution-based method* |
| • Genetic Algorithm (GA): Based on biological evolution principles, involve reproduction, crossover, and mutation to evolve solutions<br>• Differential Evolution (DE): Similar to GA, focusing on population-based mutation, crossover, and selection to generate new solution vectors<br>• Genetic Programming (GP): Evolves computer programs to perform a specific task by mimicking the process of natural evolution<br>• Evolution Strategies (ES): Uses techniques inspired by biological evolution to optimize real-valued functions |
| *Swarm-based method* |
| • Particle Swarm Optimization (PSO): Inspired by social behaviors of bird flocking and fish schooling, optimizing solutions based on the swarm's collective intelligence<br>• Butterfly Optimization Algorithm (BOA): Mimics butterfly behavior in seeking mates or food through fragrance emission, utilized for optimization<br>• Ant Colony Optimization (ACO): Mimics the foraging behavior of ants to find optimal paths through graphs<br>• Artificial Bee Colony (ABC): Simulates the foraging behavior of honeybees to optimize numerical problems<br>• Firefly Algorithm (FA): Uses the flashing behavior of fireflies to find global optima by attracting each other based on their brightness |
| *Human-based method* |
| • Teaching–Learning-Based Optimization (TLBO): Emulates the teaching and learning process in a classroom, where the best solutions are iteratively improved through teacher and learner phases<br>• Harmony Search (HS): Inspired by the musical process of finding harmonious states, focusing on memory consideration, pitch adjustment, and randomization<br>• Cultural Algorithm (CA): Uses the concept of cultural evolution, where knowledge is shared and improved over generations<br>• Group Search Optimization (GSO): Based on the group behavior of animals searching for food, such as pack hunting strategies |
| *Physics-based method* |
| • Gravitational Search Algorithm (GSA): Based on Newton's law of gravitation, where the search agents are objects with mass attracting each other to find optimal solutions<br>• Simulated Annealing (SA): Mimics the annealing process in metallurgy, where a material is heated and then slowly cooled to remove defects<br>• Electromagnetic Metaheuristic (EM): Uses principles of electromagnetism, considering solutions as charged particles that attract or repel each other<br>• Harmony Search (HS): Listed under human-based due to its dual inspiration, it simulates the physical process of finding a state of harmony |

combination significantly improves detecting and responding to cyber threats. We've provided an introductory overview of each technological element.

## Literature review

Recent advancements in computing technology, particularly AI, have significantly impacted everyday life and work by introducing systems capable of performing tasks that traditionally required human intelligence. AI systems excel in real-time analysis and decision-making, leveraging vast data volumes to solve complex problems across various scientific and technological domains. This capability is increasingly critical in cybersecurity, where the sheer volume of data makes manual analysis impractical, and the sophistication of threats, including AI-based threats, continuously evolves. Employing AI can dramatically reduce the costs and time associated with developing threat recognition algorithms despite the high expenses linked to specialist employment [15]. AI's role in cybersecurity is multifaceted. It includes the efficient and accurate analysis of large data sets, utilizing historical threat data to anticipate and mitigate future attacks, even as attack methodologies evolve. AI's adaptability makes it an invaluable tool in cyber defense. It is capable of identifying significant changes in attack patterns, managing large-scale data, and enhancing continuous learning within AI security systems to improve threat response [34].

However, the deployment of AI in cybersecurity is challenging. AI systems require extensive data to function effectively, and processing such volumes can be resource-intensive. Moreover, the risk of false alarms can undermine user trust in AI systems, and delayed responses to threats may compromise system effectiveness [34]. Furthermore, Cyber-attacks are significant risks to AI-based security systems. Despite these challenges, ongoing research enhances AI's robustness against cyberattacks. In our survey, we provide a comprehensive scope, covering a broad range of AI techniques, including ML, DL, and metaheuristic algorithms, to address various cyber threats such as malware, network intrusions, insider threats, botnets, and spam, over sixty recent studies, and the comparison of multiple AI methodologies. Unlike many surveys that do not comprehensively cover all three main topics of ML, DL, and metaheuristic technologies, we ensure that our survey addresses these areas thoroughly. Furthermore, we encompass a wide array of cyber-attacks and provide specific sections for future works and limitations, which many surveys lack. It also includes a detailed list of recent datasets and their corresponding studies, ensuring our findings are grounded in the latest research. The use of diverse benchmark datasets ensures comprehensive validation. The paper emphasizes the practical integration of AI and ML models across various environments, such as IoT, cloud computing, and traditional networks, making its findings highly applicable. Additionally, it highlights practical advantages like automation and real-time threat response, showcasing the operational benefits of AI in cybersecurity. The future recommendations are detailed and actionable, focusing on continuous improvement, the development of new datasets, transparency, explainability, and the early integration of AI in the cybersecurity lifecycle for proactive measures. In [35] it covers AI, ML, DL, and RL applications in cybersecurity, including malware detection, intrusion detection, and vulnerability assessment. However, it could benefit from exploring the integration of AI with blockchain for enhanced data integrity and real-world case studies for practical insights. A detailed comparison with traditional methods and discussions on legal, ethical considerations, and privacy issues are needed. Highlighting human-AI collaboration, robustness against adversarial attacks, and cross-disciplinary approaches would provide a broader perspective. Future AI evolution to tackle emerging threats, scalability, deployment challenges, and organizational readiness, including necessary training for cybersecurity professionals, are also crucial areas to address. These additions would make the study more comprehensive and practical for implementing AI and ML in cybersecurity. In [36] a comprehensive survey on AI system vulnerabilities to cyberattacks categorizes these threats into manipulation and extraction attacks. The technologies discussed include ML, DL, and various defense mechanisms such as adversarial training, feature squeezing, and robust aggregation methods. The paper identifies attack types like adversarial attacks, poisoning attacks, model inversion, and extraction attacks. Diverse benchmark datasets are utilized for thorough validation; only a specific domain has been introduced. A comparative analysis of several prominent studies has been conducted to provide a comprehensive understanding of the current landscape in AI and cybersecurity research. The comparison encompasses a range of criteria, including the objective and scope, methodology, data sources, and used environments. This analysis highlights the distinct approaches and findings of each study, thereby positioning the present research within the broader context of existing literature. In [37] it provides an extensive review of how AI and ML are utilized in cybersecurity. Key technologies include supervised learning for intrusion detection, malware detection, and network security. Unsupervised learning techniques to identify new threats. The paper covers applications like security automation, threat intelligence, vulnerability management, and security education, including malware, intrusion attempts, ransomware, crypto-jacking, and IoT attacks, but still, a variety of techniques for cyber-attack detection needs to be mentioned. In [38] a systematic review of AI applications in cybersecurity categorizes 236 studies within the NIST framework. It highlights AI's role in automating tasks, enhancing threat detection, and improving response accuracy using ML, DL, natural language, and RL technologies. Key areas include asset management, threat hunting, vulnerability assessment, incident response, and addressing malware, phishing, APTs, and insider threats. Despite its comprehensive scope, the paper needs more analysis of practical deployment challenges, ethical implications, and potential biases in AI models. It must also include discussions on standardized benchmarks and evaluation metrics for AI effectiveness. Future research should focus on these aspects to ensure robust and ethical AI applications in cybersecurity.

Some research has focused on identifying software vulnerabilities and malware, which are key areas where AI can significantly impact. Techniques such as data mining, ML classifiers like KNNs and SVMs, and DL architectures and metaheuristic algorithms have been extensively applied to improve malware detection and software security [39]. Asiri et al. introduced the Hybrid Metaheuristics Feature Selection with Stacked DL-Enabled Cyber-Attack Detection (HMFS-SDLCAD) model to address cyber security in the IoT environment. Their novel approach employs the Salp Swarm Optimization based on PSO (SSOPSO) for feature selection alongside a Stacked Bidirectional Gated Recurrent Unit (SBiGRU) for detecting and classifying cyber-attacks. The model also uses the Whale Optimization Algorithm (WOA) for optimizing the hyperparameters. This comprehensive system, validated against benchmark datasets, showed substantial improvements over existing models, demonstrating its efficacy in real-time cyber-attack detection [40]. Caviglione et al. present an in-depth analysis of current malware threats, illustrating how cyberattacks are increasingly sophisticated due to a combination of technological advancements and innovative exploitation methods. The paper highlights the rising incidence of malware attacks driven by cybercriminals seeking profits with relatively low risk compared to traditional crimes. The study also discusses the challenges of modern malware detection due to the complexity and diversity of attacks, emphasizing the need for constant evolution in detection techniques. The authors review the state of malware and its detection, focusing on ML techniques, which are gaining traction as a means to combat the rapid evolution of malware. Their work underscores the importance of staying ahead in the ongoing arms race between attackers and defenders in the cyber field [41]. An et al. advanced the field of cyber security by developing a CNN-based model (V-CNN) for the automated detection of vulnerabilities, utilizing DL to outperform traditional static analysis. Their approach leveraged a comprehensive dataset from MITRE's CVE/CWE and redefined vulnerabilities for enhanced detection. The V-CNN model demonstrated a remarkable 98% accuracy in identifying security vulnerabilities, indicating a significant improvement over the 95% accuracy of the RF model used in their comparisons. This work signifies a pivotal step towards integrating AI algorithms with vulnerability detection to create more resilient cybersecurity systems [42].

In network security, AI technologies have demonstrated fundamental effectiveness in developing Intrusion Detection Systems (IDS) and DoS/DDOS attacks with the support of metaheuristic algorithms. These systems benefit from AI's ability to rapidly process and analyze data, thereby reducing false alarms and optimizing feature selection to improve the reliability of network intrusion detections [43, 44]. ElDahshan, AlHabshy, and Hameed proposed a hierarchical intrusion detection system based on meta-heuristic optimization algorithms to enhance network security. Their system focuses on using extreme learning machines (ELMs) optimized through novel meta-heuristic algorithms, including the Grey Wolf Optimizer and Archimedes Optimization Algorithm. These algorithms are employed to select optimal hyperparameters and feature sets, aiming to maximize detection rates while minimizing false alarm rates. Their approach leverages the complexity of ELMs to efficiently handle multi-class classification problems in network security, providing robust solutions against various attack vectors on network systems [45]. Soliman, Oudah, and Aljuhani addressed the escalating concerns of cyber security within the Industrial Internet of Things (IIoT) by proposing a DL-based intrusion detection system. Their approach involved reducing the feature dimensionality using Singular Value Decomposition (SVD) and employing Synthetic Minority Over-sampling Technique (SMOTE) to handle imbalanced datasets. The effectiveness of their intelligent detection system was validated on the ToN_IoT dataset, yielding an exceptional accuracy rate of 99.99% for binary classification and 99.98% for multi-class classification. This study contributes to the cybersecurity landscape by offering a robust model for detecting various cyberattacks in IIoT networks [46]. Psychogyios et al. introduced a novel DL architecture for enhancing Intrusion Detection Systems (IDSs) in time series data. Their approach combined CNNs, long short-term memory networks (LSTMs), and attention mechanisms to forecast malicious network activity proactively. The researchers utilized the UNSW-NB15 dataset, converting it into a time series format, to evaluate their model. Their findings demonstrated that the model achieved comparable F1 scores and area under the curve (AUC) values within 1% and 3% of conventional real-time detection systems, respectively. Additionally, the architecture offered an ~8% improvement in F1 score over a standalone LSTM model, underlining the efficacy of integrating multiple DL techniques for threat detection in cybersecurity [47]. For the DoS/DDoS attacks, Reddy SaiSindhuTheja and Shyam developed an innovative DoS attack detection system for cloud environments, utilizing a metaheuristic algorithm known as the Oppositional Crow Search Algorithm (OCSA). This algorithm enhances feature selection in tandem with an RNN classifier to effectively identify and classify attack patterns. The integration of OCSA improves the precision and recall of the system, effectively reducing the dimensionality of data and focusing on the most significant features. Their approach, validated against benchmark datasets, showed notable superiority over traditional methods by achieving high-performance metrics, including a precision of 98.18% and an accuracy of 94.12%, thereby setting a new benchmark in DoS attack detection in cloud computing [48]. Sanjeetha et al. addressed the critical challenge of securing Software Defined Networks (SDNs) from DDoS attacks. They developed a real-time detection and mitigation model that calculates application-specific threshold limits using a ML model, distinguishing between benign and DDoS traffic. Unlike static thresholds commonly used in prior research, their dynamic threshold adapts to current network traffic conditions for more accurate DDoS detection. Upon identifying DDoS traffic, their system blocks only the offending traffic stream, allowing legitimate traffic to proceed unhindered. This approach not only improves the efficiency of DDoS attack mitigation in SDNs but also minimizes disruptions to non-malicious network applications [49]. Gaur and Gujral have made an enhanced cybersecurity measure for IoT devices by evaluating ML classifiers aimed at the early detection of DDoS attacks. Utilizing the CICDDoS2019 dataset, their proposed system demonstrated superior performance through a hybrid feature selection methodology. By implementing chi-square, Extra Tree, and ANOVA on classifiers like RF, DT, KNNs, and XGBoost, they achieved an impressive 98.34% accuracy with ANOVA for XGBoost while significantly reducing feature dimensions by 82.5%. This methodology not only enables the early detection of DDoS attacks on IoT devices but also contributes to the efficiency and robustness of cyber defense mechanisms [50].

Additionally, AI has proven effective in phishing and spam detection, where AI algorithms help identify and filter malicious content. Innovative approaches using neural networks, RL, and combinations of ML techniques like SVM, NB, NN, and DL have been particularly effective in distinguishing between legitimate communications and potential security threats, beside metaheuristic algorithms. Asiri et al. have developed an innovative NN model for cyberattack detection that utilizes an Enhanced Whale Optimization Algorithm (EWOA). This model specifically addresses the growing threat of credential stuffing attacks, which exploit the common practice of reusing credentials across multiple platforms. By optimizing the training of the NN, the EWOA significantly enhances the system's ability to detect such cyber threats. Their empirical analysis shows that this model outperforms traditional methods in detecting credential stuffing, demonstrating its potential to significantly improve digital security measures [51]. Atawneh and Aljehani investigated the application of DL to enhance email phishing detection mechanisms. Utilizing a dataset comprised of both phishing and benign emails, their research exploited advanced DL models, including CNNs, LSTMs, RNNs, and BERT, for more accurate identification of fraudulent emails. Their work, notable for leveraging natural language processing techniques, achieved a breakthrough accuracy of 99.61% by employing BERT and LSTM models. This study is a testament to the potential of sophisticated DL techniques in fortifying cybersecurity measures against the evolving threats of phishing [52]. Asiri et al. made significant advancements in real-time phishing detection systems using DL. They tackled the increasingly sophisticated phishing attacks, which often exploit legitimate web development techniques to deceive victims. Their system was designed to detect Tiny Uniform Resource Locators (TinyURLs), Browsers in the Browser (BiTB), and regular phishing attacks. By splitting their detection system into a DL model, browser extension, and docker container, they were able to achieve precision, recall, and F1 score of 99%. The model employed a Bidirectional Long Short-Term Memory (BiLSTM) with an attention mechanism to classify URLs efficiently. Furthermore, they introduced three decision strategies—Single Phishing

Strategy (SPhS), Mean Sum Strategy (MSS), and Weighted Average Strategy (WeAS)—to enhance the decision-making process on whether a webpage is benign or malicious. Their system demonstrated the best results with the WeAS, proving the importance of strategic decision-making processes in cybersecurity [53]. Alohali et al. have proposed a phishing detection model using metaheuristics and DL to enhance cybersecurity in sustainable environments. Their method combines an Improved Simulated Annealing-based Feature Selection (ISA-FS) with Long Short-Term Memory (LSTM) networks, optimized further using the Bald Eagle Search (BES) algorithm for hyperparameter tuning. This multifaceted approach not only improves the accuracy of detecting phishing websites but also optimizes the process to reduce computational overhead, achieving an impressive accuracy rate of 95.78% in their evaluations [54]. For email spam detection, Sharma and Sahni have explored the use of multi-layer perceptrons (MLPs) for efficient classification of network traffic, which is vital for identifying malicious activities in real-time. Their research proposed a hybrid model integrating MLP with DTs to enhance the predictive accuracy and speed of network intrusion detection systems (NIDS). The authors demonstrated that their model could achieve high accuracy rates, significantly outperforming traditional methods in terms of speed and detection rates. Their findings suggest that the hybrid approach not only reduces the false positive rates in detecting network anomalies but also efficiently handles large-scale data, making it a promising solution for future cybersecurity applications [55]. Butt et al. investigated cloud-based strategies for combating email phishing attacks using ML and DL techniques. Their study emphasized the effectiveness of using a combination of SVM, Naive Bayes (NB), and long short-term memory (LSTM) algorithms to classify and detect phishing emails with high accuracy. These algorithms allowed for a significant advancement in identifying malicious activities, with the SVM algorithm achieving the highest accuracy of 99.62%. This work underlines the potential of integrating various ML approaches to enhance cybersecurity measures against sophisticated email threats [56].

Collectively, these AI-based technologies address the challenges in cybersecurity more effectively than traditional methods. They do so by automating the detection and response processes, enhancing the speed and accuracy of threat detection. Understanding the mechanisms and operations behind these AI models offers more profound insights into their application, highlighting their critical role in developing more resilient cybersecurity systems. Overall, metaheuristic algorithms support the robustness and adaptability of cyber-attack detection systems, making them more effective against a wide range of cyber threats. They ensure that detection systems are not only accurate but also remain relevant over time as attack methods evolve.

## Experiments and setup

Various methods for detecting cyber-attacks have been proposed. To systematically explore these, we developed a research protocol following the systematic literature review (SLR) methodology, illustrated in Fig. 6. This protocol includes identifying the research topic, preparing research questions, selecting studies, and extracting data. By using a mixed-methods approach, combining qualitative and quantitative techniques, we can provide more straightforward and comprehensive data analysis [57].

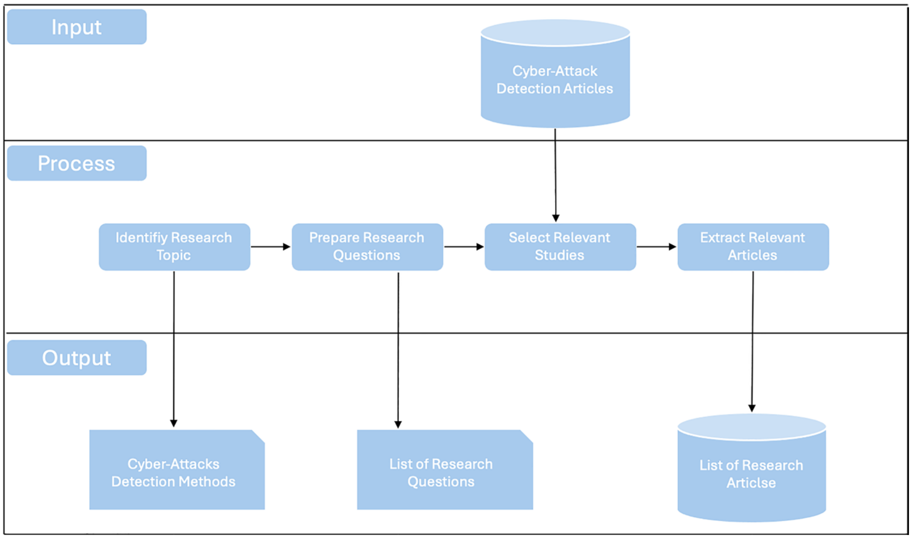
Input
Cyber-Attack Detection Articles
Process
Identifiy Research Topic
Prepare Research Questions
Select Relevant Studies
Extract Relevant Articles
Output
Cyber-Attacks Detection Methods
List of Research Questions
List of Research Articlse

**Fig. 6** Systematic literature review process

### Identify research topic

In our research, we carefully organize a selection of studies to provide a comprehensive yet focused examination of AI-driven detection techniques in cybersecurity. Our primary objective was to highlight this vast domain's most impactful and innovative contributions. The selection criteria included relevance to current trends, methodological accuracy, diversity of approaches (ML, DL, and metaheuristic algorithms), and recent datasets. By doing so, we aimed to ensure that our review remained manageable and provided meaningful insights without being overwhelming. This approach allowed us to delve deeper into these methods' effectiveness, limitations, and potential improvements, ultimately presenting a clearer picture of the current state and future directions of AI in cybersecurity. While thousands of contributions exist, our selection strategy focused on those studies that present significant advancements and practical applications in detecting various types of cyber-attacks.

To gather a wide range of studies, we used a neutral search strategy focusing on key review articles about detecting cyber-attacks. Our search was thorough and well-planned, with "cyber-attack detection" chosen as our main search term to capture relevant articles. Before starting our review, we evaluated three databases: Scopus, Google Scholar, and Web of Science. We chose Scopus because it includes major publishers like ACM, Springer, and IEEE, and provides a more selective coverage compared to Google Scholar, which includes some non-peer-reviewed works like technical reports.

Our main focus is on recent methods of cyber-attack detection published from 2020 to 2024. We found that a total of 12,931 articles were published in Scopus, with 9084 of these from 2020 to 2024. Google Scholar listed 112,000 total articles, with 21,100 from this period. Web of Science had a total of 664 articles, with 419 published between 2020 and 2024 as shown in Fig. 7.

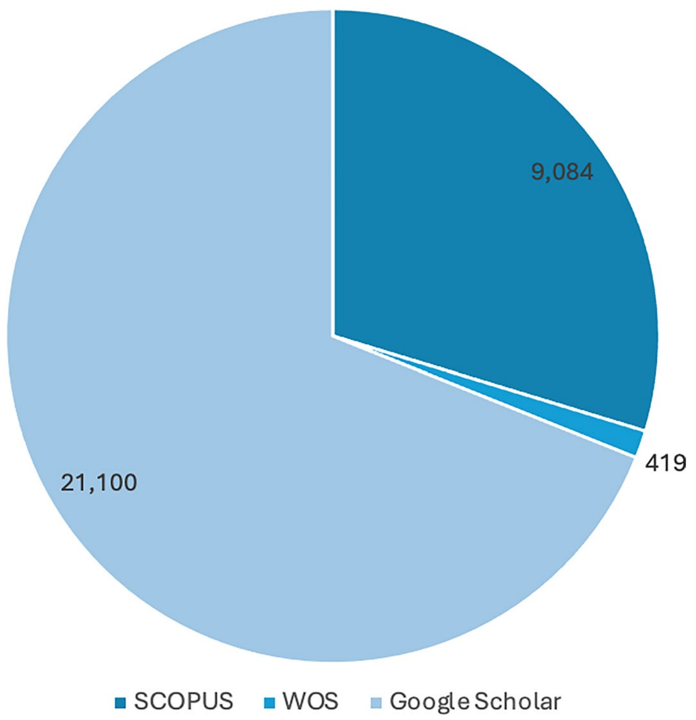

**Fig. 7** Published articles from 2020 to 2024 in different DBs

### Prepare research questions

We formulated ideas for paper analysis and established specific research questions (RQs) to guide our study. Initially, we explored the types of techniques used (RQ1, RQ2, and RQ3), which helped us understand how these methods are developed and applied. Subsequently, we also examined how these methods are evaluated, introduced their challenges (RQ5, RQ6, and RQ7), the datasets that are used (RQ4), and finally, the proposed future work (RQ8, RQ9, and RQ10). The research questions posed in this study are as follows:

RQ1: What ML models are utilized to detect cyber-attacks?

RQ2: What DL models are utilized to detect cyber-attacks?

RQ3: What metaheuristic algorithms are employed in cyber-attack detection?

RQ4: What are the most commonly used datasets for detecting cyber-attacks?

RQ5: What are the limitations of ML models used in cyber-attacks detection?

RQ6: What are the limitations of DL models used in cyber-attacks detection?

RQ7: What are the limitations of metaheuristic algorithms used in cyber-attacks detection?

RQ8: What future work is suggested for ML models in cyber-attack detection?

RQ9: What future work is suggested for DL models in cyber-attack detection?

RQ10: What future work is suggested for metaheuristic algorithms in cyber-attack detection?

**Table 6** Inclusion and exclusion criteria

| | |
|---|---|
| Inclusion criteria | Papers related to software attack detection |
| | Studies utilizing ML, DL, or Metaheuristic (MH) algorithms |
| Exclusion criteria | Works published in journals or conferences, and peer reviewed |
| | Studies focusing on attack detection methods outside of ML, DL, and MH |
| | Research lacking empirical analysis, surveys, or results |
| | Studies where the full text is unavailable |

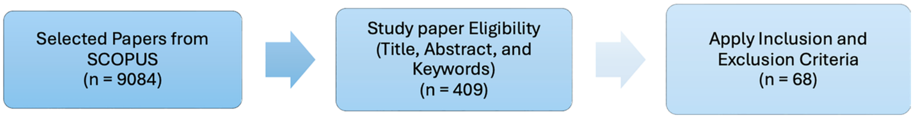
Selected Papers from SCOPUS (n = 9084)
Study paper Eligibility (Title, Abstract, and Keywords) (n = 409)
Apply Inclusion and Exclusion Criteria (n = 68)

**Fig. 8** Paper selection criteria

### Select relevant studies

After selecting the digital repositories, it was necessary to address a search string to perform an exhaustive search to select the main studies. The process of defining the search string involved four steps:

1- Using the predefined research questions to identify the relevant outcomes.
2- Identifying synonyms and alternative spellings for each key term.
3- Verifying the presence of search terms in titles, abstracts, and keywords of the articles.
4- Applying Boolean operators such as "AND" and "OR" to formulate the search string effectively.

The resultant search string, addressed using the above steps, was: ("Software Cyber-attacks" OR "SW Cyber-attacks") AND ("Detection") AND ("Methods") AND ("Cyber-attack" OR "Cyber attack" OR "Cybersecurity" OR "Cyber threats") AND ("Machine Learning" OR “ML” OR "Deep Learning" OR “DL” OR "Metaheuristic Algorithms" OR "Optimization Algorithms") OR ("Artificial Intelligence" OR "AI"). This string was employed to aggregate all published papers, aiming to maximize the scope of relevant literature within a set timeframe from 2020 up to now.

To refine the selection of primary studies, we established inclusion and exclusion criteria as shown in Table 6.

Initially, 9084 studies were identified through the search string. After screening these studies based on titles, abstracts, and keywords, we narrowed them down to 409 primary studies. Further application of the inclusion and exclusion criteria reduced the selection to 68 relevant studies. These studies were thoroughly reviewed to confirm their quality and relevance to our research goals, focusing on their scientific rigor and contributions to the field of cyber-attack detection.

### Extract relevant articles

The papers retrieved from the database were filtered based on the criteria we established, as presented in Fig. 8. Initially, we gathered 9084 studies using our specified search string. We then excluded primary studies based on their titles, abstracts, and keywords, resulting in 409 primary studies. By strictly applying the inclusion and exclusion criteria, we further narrowed this down to 68 studies.

We categorized and analyzed each paper by extracting key information. This included details from the paper, like the publication year, authors, citation count, field of study, model, evaluation metrics, and dataset. We also focused on each paper based on the type of learning method used. So, to gather this information, we primarily reviewed the title, abstract, and introduction of each paper, which typically contained all the necessary details. We only consulted the full text for additional information when needed.

We screened papers for quality, length, and type to obtain a collection that could be used effectively for research and analysis.

## Results and discussions

In this section, we explore and analyze our research findings in detail, specifically focusing on the effectiveness of various cyber security attack detection methods. Our analysis thoroughly addresses the research questions outlined in "Prepare research questions" section, providing detailed answers that not only enhance understanding but also contribute effectively to the field. These insights are presented in a way that other researchers can easily build upon, serving as a solid foundation for further exploration and development in cyber security measures. Based on our comprehensive review and assessment, we put forward clear recommendations designed to improve the detection capabilities of systems against cyber threats. These suggestions are practical and geared towards enhancing the ability of systems to detect and respond to evolving cyber threats more effectively.

### RQ1: What ML models are utilized to detect cyber-attacks?

In addressing the research question concerning the application of ML in the detection of various cyber-attacks, we conducted a comprehensive literature review. This involved a systematic collection of recent research papers focused on ML strategies within the cyber security domain. We specifically selected those studies that proposed solutions to detect an array of cyber threats. These selected papers have been summarized, with key details and findings extracted and presented in Table 7. This table serves as a valuable structure of current research trends and the effectiveness of different ML models in the field, providing a consolidated reference for further investigation and development of cyber-attack detection systems.

### RQ2: What DL models are utilized to detect cyber-attacks?

For the DL-focused research question, our systematic review encompassed a vast array of recent studies that utilize DL techniques to improve cyber-attack detection. The breadth of this investigation includes papers that use DL's capabilities to analyze

**Table 7** Recent studies in ML models in cyber-attacks detection

| Ref. | Year | Attack problem | Used method | Result | Dataset | Pro | Cons | Environment | Future work |
|---|---|---|---|---|---|---|---|---|---|
| [58] | 2020 | Cyber-attack detection | DBN, DT, and SVM | Models accuracy: DBN: 97.50, DT: 99.96, SVM: 95.11 | Multiple including NSL-KDD, DARPA | Enhance cybersecurity capabilities | Requires large datasets, computational resources | Network based environment | Develop new and hybrid models |
| [59] | 2024 | Botnet Attacks | Network Traffic Analysis and Machine Learning | 99.8% botnet traffic filtered, 100% accuracy | Live botnet attack dataset | High accuracy, efficient traffic filtering | Challenging deployment on resource-constrained devices | Network infrastructure, IoT devices | Improve traffic analyzer, enhance honeypot, train ML on recent datasets |
| [60] | 2024 | Insider Threat Monitoring & Detection | Hybrid detection: Machine Learning + Statistical Criteria | Accuracy: 98.48% | CERT r4.2 | Handles bias and data imbalance; improves prediction accuracy | High computational cost; needs real-time implementation | Organizational security | Develop employee-specific trained models; deploy client–server architecture for threat detection API |
| [61] | 2024 | Intrusive applications in Android | Continuous Threat Monitoring Framework (CTMF) | Identified 378,480 intrusive apps with 77.919% accuracy | Google Play Store data | High accuracy in detecting intrusive apps | Resource-intensive, time-consuming new dataset preparation | Android | Enhance tools for analyzing AndroidManifest.xml, develop dedicated Google Play API |
| [62] | 2022 | DOS/DDOS attacks in IoT | "Looking-Back" concept for detection with RF classifier | 99.81% accuracy | Bot-IoT dataset | Precise handling of DoS/DDoS attacks | Computational complexity | IoT | Test against smarter attacks |
| [63] | 2022 | Detect Cyber-attacks in IOT | MFO-RELM model | Accuracy: 99.79%, Precision: 98.84%, Recall: 98.84%, F-score: 98.84% | N-BaIoT | Effective cybersecurity threat identification | Model's implementation complexity | IOT | Incorporate novel clustering and outlier removal |
| [64] | 2023 | DDoS for Software-defined Network | Improved binary grey wolf optimization with ML algorithms | 99.13% accuracy | CSE-CIC-IDS2018 | New solution for SDN security | Extensive parameter tuning | SDN | Classify different types of malware |
| [65] | 2023 | Intrusion Detection | FFO for detecting intrusions and PNN for categorization | 98.99% accuracy | KDD-CUP 99 | High performance across various cyber-attacks | Handling zero-day attacks, computational requirements | Network | Develop more adaptable IDS technologies |
| [66] | 2024 | Network Intrusion Detection | GSAFS-OQNN model | Accuracy: 99.79%, Specificity: 99.88%, MCC: 98.72% | UNSW-NB15 | Optimal feature selection, adaptability | Computational complexity | Network | Improve the model's adaptability and scalability |
| [67] | 2024 | DDOS Attacks | Evolutionary optimization algorithms with ML techniques | 99.99% accuracy with XGB-GA | KDD Cup 99, CIC-IDS 2017 | Advances DDoS attack detection, enhances cybersecurity | Complexity of implementation and tuning | Network | Further, improve detection capabilities |

**Table 7** (continued)

| Ref. | Year | Attack problem | Used method | Result | Dataset | Pro | Cons | Environment | Future work |
|---|---|---|---|---|---|---|---|---|---|
| [68] | 2023 | Cloud computing attacks | Supervised learning algorithms, including SVM, LR, RF, DT, NB, XGBoost, and KNN | Detection rate over 99% | Private cloud environment dataset | High detection rate and efficiency in cloud computing security | Scalability and adaption to evolving attacks | Cloud | Detect specific cloud attacks, explore more classifiers |
| [69] | 2024 | Enhance IDS | LR, SVM, DT, and RF | RF achieves an F1 score of 97.80% | UNSW-NB15 | Data-driven cybersecurity with early threat detection | Computational complexity, scalability issues | Network | Test with actual network traffic data |
| [70] | 2023 | Various cyber-attacks in IoT | ML algorithms including ADA Boost, LSVM, Auto Encoder Classifier, QSVM, MLP | ADA Boost has 98.3% accuracy | UNSW-NB15 | Improves detection accuracy of IDS | Security challenges in IoT-based smart cities | IoT | Enhance security in IoT networks with AI and GA |
| [71] | 2024 | Phishing attacks | LR and RF | 92% accuracy in detecting URLs | Phishing Dataset from Kaggle | Real-time monitoring, high accuracy | Potential false positives/negatives in classification | Web and Internet Services | Expand data, enhance detection range |
| [72] | 2023 | DDOS and MiTC Attack in the Cloud | DT, SVM, NB, KNN | 99.96% accuracy with DT | Simulated DDOS and MiTC datasets | Adaptive identification of evolving attack techniques | Need for extensive training data | Cloud | Detect other types of attacks in cloud environments |
| [73] | 2022 | Cyber-attacks detection | RF, J48, NB, Multi-Layer Perceptron algorithms | 99.76% accuracy | NSL-KDD | Efficient cyber-attack detection accuracy | Requires extensive datasets for training | Network | Enhance technique performance |
| [74] | 2024 | Malware Detection | ML algorithms with feature selection based on TFIDF | RF has 97.68% accuracy | UNSWNB15 | Detects multiple types of attacks, enhances network security | Computational demands, large datasets needed | Cloud | Investigate Hidden Markov Models and DL |
| [75] | 2023 | Network Intrusion Recovery in SDN | MLBNIR approach including LR, DT, SVM, RF | Reduced intrusion recovery time by up to 90% | InSDN dataset | Effective detection and reduced recovery time in SDNs | Computational complexity | SDN | Enhance IDS performance with new algorithms |
| [76] | 2023 | DDOS Attacks in SDN | MTD approach using probabilistic linear models, NN, and trees | Efficient detection within 3 s | Kaggle | Effective solution for sophisticated cyber threats in SDNs | Significant computational resources are needed | SDN | Improve IDS performance with new algorithms |

**Table 8** Recent studies of DL models in cyber-attacks detection

| Refs. | Year | Attack problem | Used method | Result | Dataset | Pro | Cons | Environment | Future work |
|---|---|---|---|---|---|---|---|---|---|
| [53] | 2024 | Phishing Attack | Hybrid methods with URL extraction and DL model | Precision, Recall, F1 Score: 99% | Phishing URLs and benign URLs dataset | Real-time detection, high accuracy | Complexity, potential for delayed performance | Web and Internet Services | Expand dataset, refine models, focus on URL paths |
| [77] | 2024 | Adversarial attacks on DL-based IDS | DLL-IDS with LID method | LID method Improved detection accuracy from 17.9% to 71.7% under CW attack | NSL-KDD, CIC-IDS2018 | High accuracy, adaptable, maintains accuracy on clean samples | Computational overhead, optimization needed for real-time use | Network | Enhance IDS security across domains, explore robust methods against attacks |
| [78] | 2024 | Intrusion detection using hybrid models | Dugat-LSTM model with chaotic Honey Badger optimization | Achieves 98.76% accuracy on TON-IOT, 99.65% on NSL-KDD | TON-IOT, NSL-KDD | High accuracy, effective feature selection, low memory usage | Potential to fall into local optimum, complexity of hybrid model | IoT | Hybridize HBO with other techniques to reduce computational cost and improve selection |
| [47] | 2024 | Intrusion in network traffic | Proactive IDS with CNN, LSTM, and attention models | F1 score: 91% for T = 20 packets, AUC within 3% of real-time detection | UNSW-NB15 | Advanced model leveraging DL for proactive prediction | Extensive computation, complex model setup | Network | Explore additional DL architectures for improved prediction accuracy |
| [79] | 2024 | DDoS Detection | Hybrid Deep Learning (CNN, LSTM) | 99.995% (CICIoT2023), 98.75% (TON_IOT) | CICIoT2023, TON_IOT | High accuracy | High computational cost, imbalanced datasets | IoT | Optimize parameters, reduce training times, develop cost-efficient IDS |
| [80] | 2023 | Intrusion detection in networks | Deep Convolutional NN | Achieves 99.79% to 100% accuracy | ISCX-IDS 2012, DDoS (Kaggle), CICIDS2017, CICIDS2018 | High accuracy, real-time detection, low false positives | High computational requirements, issues with imbalanced datasets | Networks | Enhance model generalization and develop real-time DDoS response |
| [81] | 2020 | Web Spam Detection in IoT | Cognitive spammer framework using LSTM and NN | High accuracy (96.96%) | WEBSPAM-UK 2007 | An active solution to web spam reduces computational costs, improves reliability in IoT environments | Need for large datasets, computational intensity | IoT | Explore additional DL algorithms and manage large-scale data with data-streaming models |
| [82] | 2021 | Malware Detection | Deep-learning-based architecture integrates ResNet-50 and AlexNet | High accuracy (97.78%) | Malimg, Microsoft BIG 2015, Malevis | Efficiently identifies malware variants, reduces feature space, enables high-accuracy classification | Significant computational resources and large datasets are needed for training | Cloud | Detection system for malware using obfuscation techniques |

**Table 8** (continued)

| Refs. | Year | Attack problem | Used method | Result | Dataset | Pro | Cons | Environment | Future work |
|---|---|---|---|---|---|---|---|---|---|
| [83] | 2022 | DDOS attack detection | Intrusion detection system using DNN, CNN, and LSTM | 99.99% accuracy for binary classification, 99.30% for multiclass | CIC-DDoS2019 | High detection accuracy, real-time performance, adaptive recognition of various DDoS patterns | High computational resources and large datasets for effective training | Network | Test architecture against different IoT datasets, enhance resilience to adversarial learning attacks |
| [84] | 2023 | Malware Detection | DL methodology with multiple layers and model combinations | Over 99% detection rate, 99.80% accuracy | Generated new Malware dataset | Substantial improvements over existing methods, effective against known and zero-day malware | Challenge of identifying emerging and sophisticated malware | Cloud | Expand application to classifying specific malware types, evaluate effectiveness on additional datasets |
| [85] | 2024 | Ransomware Detection | Ebola optimization search algorithm for enhanced DL-based detection | Accuracy: 99.88%, Sensitivity: 99.88%, Specificity: 99.88% | Dataset with 840 samples including goodware and ransomware | High accuracy and robustness in detection | Computational complexity, substantial training data needed | IOT | Novel approach combining Ebola optimization with DL for improved detection in IoT devices |
| [86] | 2023 | DDOS, DoS, Brute-force, and botnet attacks in the Cloud | Hybrid DL-based approach for intrusion detection using PCA, SMO-FCM, and AE | Achieved 95% accuracy | CSE-CIC-IDS-2018 | Significant effectiveness in detecting and classifying malicious activities | Need for continuous model updates, potential computational demands | Cloud | Enhance adaptability to new cyber threats, integrate more advanced AI techniques |
| [87] | 2022 | IDS | DL Model based on LSTM | Up to 99% detection accuracy | CSE-CIC-IDS-2018 | High accuracy in feature extraction, adept at analyzing large datasets | Significant computational resources and extensive training data are required | Network | Enhance model accuracy, reduce error rates, and speed up training by identifying the most relevant features |
| [88] | 2024 | IoT cyber-attack detection | DNN, RNN, and CNN models | High accuracy of 96.56% | CICDIoT2023 | High accuracy, versatility in application, leverages DL for enhanced detection | Substantial computational resources, potential overfitting, large datasets needed | IoT | Incorporate incremental training, explore real-time cyberattack detection system integration |
| [89] | 2023 | Malicious attacks detection | Bi-directional LSTM model | 99% precision and recall rates | UGR'16 | High precision and recall rates, effective identification and classification | Computational demands, extensive datasets for training | Network | Explore other feature engineering techniques with Bi-LSTM for improved performance |

**Table 8** (continued)

| Refs. | Year | Attack problem | Used method | Result | Dataset | Pro | Cons | Environment | Future work |
|---|---|---|---|---|---|---|---|---|---|
| [90] | 2023 | Web-based attacks detection | CNNs and RNNs | High accuracy 94–96% | KDD Cup 1999 and CICIDS2017 | Effective at identifying sophisticated cyber threats, adaptable to cybersecurity challenges | High computational demand, complex model training | Web and Internet Services | Integrate DL with other AI methods like RL for robust, adaptive systems |
| [91] | 2023 | Cyber-attacks in IoT networks | Distributed DL framework using FFNN and LSTM | Up to 99.95% accuracy | NSL-KDD and BoT-IoT | Effective identification of sophisticated threats, adaptable to cybersecurity dataset challenges | Significant computational resources, extensive training datasets required | IoT | Explore integration of DL with other AI methods like adversarial learning for more robust systems |
| [92] | 2022 | Network Cyber-attacks | Combination of DL and ML algorithms with statistical techniques | Best: LSTM + CNN (Acc: 99.5%, Prec: 100%, Rec: 99%, F1: 99.5%) | HTTP DATASET CSIC 2010 | Effective use of DL in cybersecurity, detects multiple attacks | Requires large computational resources | Networks | Explore additional deep learning configurations and datasets |
| [93] | 2022 | DOS, U2R, probe, and R2L attacks | Spider Monkey Optimization, Stacked Deep Polynomial Network | Acc: 99.02%, Prec: 99.38%, Rec: 98.91%, F1: 99.14% | NSL-KDD | Significant effectiveness in identifying and classifying malicious activities | Computational demands necessity for extensive datasets | IoT | Further analysis with respect to payload-based detection, explore different classifiers and datasets |
| [19] | 2022 | Cybersecurity attack detection using deep learning | CNN, RNN, DNN | CNN: Acc: 98.64% with binary class, RNN: Acc: 97.75%, DNN: Acc: 96.81% | USTC-TRC2016, NSL-KDD | High accuracy and precision in attack detection | Intensive computational and data processing requirements | Network | Improve DL methodologies for better accuracy and efficiency in real-time applications |
| [94] | 2023 | Intrusion Detection in SDN | Transfer Learning with Reptile-TL | Anomaly Detection: 0.71, Attack Type Identification: 0.91 | InSDN | Effective for unknown attacks and small sample scenarios | Requires long training times (over 4 h) | SDN | Combine sampling techniques with few-shot learning |
| [95] | 2023 | Detection and Defense against DDoS Attacks in SDNs | Adversarial DBN-LSTM | Accuracy: 91.23%, Sensitivity to adversarial DDoS | CICDDoS 2019 | Effective against adversarial attacks, efficient training | High computational costs, complex model architecture | SDN | Refine models to enhance resistance to adversarial inputs |

complex and large-scale data, which are characteristic of today's digital ecosystems. These scholarly works have been critically summarized, covering various DL architectures like CNNs, RNNs, and more advanced models. The essence of these papers, along with their novel contributions and findings, has been efficiently encapsulated in Table 8. This collection serves as a robust analytical tool for researchers to identify gaps and opportunities for innovation in the landscape of DL applications in cyber security.

### RQ3: What metaheuristic algorithms are employed in cyber-attack detection?

Concerning the research question on metaheuristic algorithms, we have a list of research articles that explore the application of metaheuristic algorithms for cyber-attack detection. Metaheuristic algorithms, known for their ability to find optimal or near-optimal solutions for complex optimization problems, are increasingly applied to enhance the detection rates in cyber security systems. The research papers chosen for review present various metaheuristic approaches, including Genetic Algorithms, Particle Swarm Optimization, and Ant Colony Optimization, among others. We have refined the core of these papers and summarized their core methodologies and outcomes in Table 9. This table provides a synthesized viewpoint on how metaheuristic algorithms are being leveraged to advance the state-of-the-art in detecting and mitigating cyber threats, thereby offering a strategic starting point for future research endeavors in this domain.

### RQ4: What are the most commonly used datasets for detecting cyber-attacks?

Common datasets for detecting cyber-attacks vary in characteristics, with both old and modern datasets offering unique benefits and limitations [111]. These datasets are summarized at the below Table 10.

In addition to the commonly used benchmark datasets for detecting cyber-attacks, we have recently introduced several modern datasets that come with their own unique set of advantages and limitations:

1 *PhiUSIIL phishing URL dataset*: This dataset, generated between October 2022 and May 2023, includes 134,850 legitimate URLs and 100,945 phishing URLs. It features attributes like top-level domains, URL length, subdomains, and obfuscated characters. Its balanced nature and recent phishing techniques enhance model accuracy, although it requires significant computational resources and may over-specialize due to continuous training [115].

2 *CICEV2023 dataset*: Created in 2023, this dataset focuses on DDoS attacks on EV authentication within smart grid infrastructure. It includes metrics such as "Time delta," "Instruction overhead," and "CPU cycle overhead," with 5,284 normal and 58,000 attacks EV authentication attempts. It provides a robust foundation for developing detection models but focuses more on infrastructure status than actual EV charging records [116].

3 *Edge-IIoTset dataset*: This dataset, generated from November 2021 to January 2022, includes 61 features and covers various attacks, such as DoS/DDoS, information gathering, and malware. It supports centralized and federated learning modes and

**Table 9** Recent studies of metaheuristic models in cyber-attacks detection

| Refs. | Year | Attack problem | Method | Result | Dataset | Pro | Cons | Environment | Future work |
|---|---|---|---|---|---|---|---|---|---|
| [96] | 2023 | Cyber-attack detection in IoT networks | Hybrid feature selection scheme with NSGA-II | 99.48% accuracy | ToN-IoT | Efficient feature minimization | Computational complexity | IoT | Efficient and fast IDS for IoT |
| [97] | 2023 | IoT Cyber threat detection | Framework using BGSA and BGWO for optimized feature selection | High accuracy (99.41%) | UNSW-NB15 | High accuracy and low FPR | Requires complex feature selection | IoT | Optimizing the feature selection and classification framework |
| [98] | 2023 | Email Spam Detection | PSO and GA with ML techniques | Accuracy(97.66%) Precision (94.21%) Recall (97.23%) | Spam Email and Enron Spam Dataset | Adapts to new spam strategies | Requires complex implementation and tuning, computation intensive | Email System | Explore integration with other meta-heuristic techniques |
| [99] | 2024 | Intrusion cyber-attacks | Deep Stacked Ensemble, GWO | 99% accuracy | MSU-ORNL | Adaptability and learning capability | Computational complexity | Network | Integration of more sophisticated AI and ML techniques |
| [100] | 2023 | Android malware detection | RHSODL-AMD Model with RHSO for feature and ARAE model optimized with Adamax | 99.05% accuracy | Andro-AutoPsy | Effective discrimination between malware and legitimate apps | Computational efficiency | Android | Improve RHSODL-AMD technique's performance |
| [101] | 2023 | Cyber-attacks on SCADA systems | RBM with artificial root foraging optimization | Up to 97.8% accuracy | Oak Ridge National Lab Dataset, MSU | High accuracy and adaptability | Requires intensive computation | Smart grid power systems, SCADA systems | Enhance performance and extend the algorithm |
| [102] | 2024 | Traffic diversion attacks in SDN | GA in SDN framework | Over 70% accuracy | Public dataset | High adaptability to SDN environments | Computational demands | SDN | Further improvements in the algorithm's performance |
| [103] | 2023 | Network Intrusion Detection | GWDTO hybrid metaheuristic optimization | 98.1% accuracy and high stability | IoT-IDS | Enhanced performance in feature selection | Computational complexity | Network | Enhancements to GWDTO algorithm and broader application |
| [104] | 2023 | Network Intrusion Detection | FCM and NN classifier with GA, PSO | 99.97% accuracy | CICIDS2017 | Improved accuracy in detecting cyber-attacks | High computational resources required | Network | Further evaluate against recent datasets |
| [105] | 2024 | DOS Attack | Hybrid optimization-based DL with DBN, AO, DHOA | 92.8% accuracy | NSL-KDD, BOT-IoT | Effective DoS attack detection | Computational complexity | Network | Test the model across more datasets |
| [106] | 2024 | Network intrusion Detection | Bio-inspired optimization with DL for attack detection | Over 98.8% accuracy | CSE-CIC-IDS2018 | High accuracy with reduced feature sets | Requires meticulous feature selection to achieve high accuracy | Network | Enhance detection capabilities |

**Table 9** (continued)

| Refs. | Year | Attack problem | Method | Result | Dataset | Pro | Cons | Environment | Future work |
|---|---|---|---|---|---|---|---|---|---|
| [107] | 2021 | Network intrusion Detection | TS-RF using Tabu Search and RF | 83.12% accuracy. FR > 60% | UNSW-NB15 | Improved classification accuracy, Reduced time complexity and feature space | Requires intensive computation for optimization | Network | Explore new algorithms and models |
| [48] | 2021 | DOS Attacks in Cloud Computing | OCSA for feature selection and RNN for detection | 94.12% accuracy | KDD Cup 99 Dataset | High accuracy in classifying malicious activities | Requires intensive computation | Cloud | Enhance detection of various attacks |
| [108] | 2023 | Network Intrusion Detection | MQBHOA with HOA and quantum computing | 99.8% accuracy | NSL-KDD, CSE-CIC-IDS2018 | Effective solution for sophisticated cyber threat detection | Computational resources need | Network | Enhance IDS performance in cloud computing |
| [45] | 2022 | Hierarchical Intrusion Detection | Meta-heuristic optimization with ELM | Accuracy: 98.93%, DR: 99.63%, FAR: 0.01% | UNSW-NB15, CIC-IDS2017 | High accuracy, low FAR, effective for multiple attack types | Complex setup and computational intensity | Network | Refinement of FS techniques, explore additional datasets |
| [109] | 2024 | Clustering-Based DDoS Attack Detection | Cuckoo Search (CSA), Flower Pollination (FPA), Firefly (FSA) | CSA: FPR 0.02 (best for lower approximations), FPA: FPR 0.015 (best for upper approximations), FSA: FPR 0.03 | CICIDS2017 | Effective for specific ranges of approximation in feature selection | Limited by switching approximations, less effective for middle ranges | Network | Refine switching probabilities, broader dataset testing |
| [110] | 2023 | Malicious attacks in cloud computing | Hybrid Metaheuristics (PSO, FFA, SFLA), CNN | Improved classification accuracy; FFA achieved the best results (99.84%) | Microsoft Malware prediction database and GitHub database | Optimized feature selection for cloud data classification | Computational complexity and parameter tuning needs | Cloud | Further comparative studies and exploration of additional metaheuristic algorithms |
| [40] | 2022 | Cyberattacks in IoT | Hybrid Metaheuristics, Deep Learning, SBiGRU, SSOPSO | High accuracy 99.77%, outperformed older models | Benchmark dataset | High accuracy and robustness against attacks | Computational costs | IoT | Incorporate feature reduction and outlier removal |

**Table 10** Datasets benchmark overview [112–114]

| Dataset | Year | Records | Benign % | Malicious % | Attacks |
|---|---|---|---|---|---|
| CICIDS2017 | 2017 | ~2.8 M | 83.1 | 16.9 | DDoS, DoS, BruteForce, and others |
| MQTTset | 2020 | ~1 M | 70 | 30 | DoS, Publish Flood, Malformed data and more |
| CSE-CIC-IDS-2018 | 2018 | ~16 M | 85 | 15 | DDoS, DoS, Bot, SQL Injection, and more |
| CIC-DDoS2019 | 2019 | 50 M | 0.11 | 99.89 | Volumetric DDoS |
| UNSW-NB15 | 2015 | 2.5 M | 90 | 10 | Various, including 9 attack types |
| ADFA-WD | 2013 | ~50 K | 71.4 | 28.6 | Brute force, Windows local exploits, Meterpreter, Hydra |
| ADFA-LD | 2013 | ~50 K | 79.2 | 20.8 | Brute force, Linux local exploits, Meterpreter, Hydra |
| UNSW-BotIoT | 2019 | 72 M | 60 | 40 | DDoS, DoS, OS, and Service Scan |
| DoHBrw2020 | 2020 | 1.4 M | 67.6 | 32.4 | Malicious DNS traffic |
| ISCX-URL-2016 | 2016 | 80 K | 30.9 | 69.1 | Benign, spam, phishing, malware URLs |
| DARPA1998 | 1998 | ~4,9 M | 97 | 3 | Evaluation of intrusion detection systems |
| KDD Cup 99 | 1999 | ~4 M | 98 | 2 | DOS, Probe, R2L, U2R |
| NSL-KDD | 2009 | 148 K | 70 | 30 | DOS, Probe, R2L, U2R |
| CAIDA | 2007 | ~20 M Flows | 78 | 22 | Various typrs of DdoS attacks |
| Kyoto 2006+ | 2006 | ~90 M | 20 | 80 | DoS, Probing, U2R, R2L |

comprises 421,417 normal and 399,417 malicious records. While comprehensive, its short generation period may limit long-term trend analysis [113].

4 *CIC-Malmem-2022 dataset*: Released in 2022 by the Canadian Institute of Cybersecurity, this dataset includes 58,596 samples with 56 features, focusing on memory-based obfuscated malware across Trojan, Spyware, and Ransomware. It is suitable for contemporary threat detection but requires advanced models and extended training times [117].

5 *X-IIoTID dataset*: Collected over a week, this dataset captures periodic effects and includes 820,834 instances with 67 features. It covers diverse IIoT protocols and attack types, ensuring interoperability and comprehensive labeling. However, it lacks predefined training and testing splits and has limited coverage of cyber-physical attacks on PLCs [118].

These modern datasets offer valuable resources for cybersecurity research, enabling the development and validation of more effective detection models and security frameworks. However, each comes with its own set of challenges that must be considered when integrating them into research and practical applications.

### RQ5: What are the limitations of ML models used in cyber-attacks detection?

ML models are a key asset in cyber-attack detection, but they come with a set of limitations that impact their practical application in cybersecurity:

- *Large datasets requirement*: require huge training, accurately labeled data, which is often hard to source in the cybersecurity area [52].

- *Computational demand*: Training and implementing these models require significant computational power, presenting a challenge for resource-limited configurations [53, 54].
- *Vulnerability*: ML models are subject to malicious attacks like adversarial attacks trick ML models with fake inputs, causing errors, evasion attacks alter data to bypass detection, data poisoning adds harmful data to training sets, weakening models, and model inversion extracts sensitive information by reverse-engineering models. These issues highlight the need for robust defenses like adversarial training, regular updates, and thorough validation [35].
- *Complexity and interpretability*: The complex architecture of ML models, leads to difficulties in understanding their decision-making process, which is critical in cybersecurity for establishing trust [55, 58].
- *Adaptability*: Models often need retraining to keep up with new or changing attack methods, risking the oversight of zero-day attacks [56, 62].
- *Scalability challenges*: It is challenging to scale ML models to handle large data volumes and provide real-time analysis [59, 60].

These constraints highlight the need for continuous research into more advanced, adaptable, and efficient ML methods tailored to the dynamic field of cybersecurity.

### RQ6: What are the limitations of DL models used in cyber-attacks detection?

DL models face notable challenges in cyber-attack detection:

- *Dataset requirement*: DL models require large training datasets, leading to a high computational load [55].
- *Resource constraints*: Effective training and operation need effective computational resources, which may not be feasible in all environments [68–70]
- *Regular updates needed*: To track evolving threats, continuous updates of DL models are necessary to maintain their effectiveness [72].
- *Complex algorithms*: The advanced algorithms used in DL models add to their computational complexity [71].
- *Labeled data shortage*: There's often a lack of readily available, well-labeled training data [74].
- *Delayed real-time detection*: The heavy computational demands can slow down real-time detection capabilities [73].
- *Vulnerability attacks*: DL models can be sensitive to sophisticated malicious attacks, indicating a need for stronger defenses [81].

Despite their high accuracy in detecting various attacks, these challenges highlight the need for continued enhancement of DL models for cybersecurity.

### RQ7: What are the limitations of metaheuristic algorithms used in cyber-attacks detection?

Metaheuristic algorithms, while advantageous for cyber-attack detection due to their flexibility and efficiency, face several limitations:

- *Computational complexity*: They require significant processing power and can be time-consuming, especially with large or complex datasets [82, 83].
- *Feature selection*: Their effectiveness is highly dependent on careful feature selection, with wrong selection leading to poor performance [89].
- *Resource demands*: Powerful computational resources are needed for training and running these algorithms, posing challenges in resource-limited settings [92].
- *Data preprocessing ne*eds: The importance of data preprocessing adds to the complexity and time of deployment [85].
- *New attacks*: Adapting to new and changing cyber threats may require retraining or significant adjustment, which could reduce accuracy over time [85, 87].

Improving these algorithms affects addressing their computational needs, feature selection, managing resources efficiently, streamlining data preprocessing, and enhancing their ability to adapt to new attacks.

### RQ8: What future work is suggested for ML models in cyber-attack detection?

- *Integration of emerging technologies*: ML in cybersecurity will integrate with virtualization, blockchain, big data, and cloud computing to enhance threat detection, data protection, and scalability, addressing complex and large-scale cyber threats efficiently [58].
- *Development of a new detection system*: A new ML-based detection system will incorporate modern technologies to improve identification of attacks, manage high-dimensional data, and handle anomalies, creating a robust and efficient cybersecurity solution [59, 61].
- *Improving the robustness of AI models*: ML models will be enhanced through adversarial testing, defining robustness metrics, using ensemble models, preprocessing, regularization, transfer learning, and continuous updates to ensure resilience against evolving threats [66].
- *Combining RL with other techniques*: Future work will combine RL with ML techniques to develop adaptive and efficient cybersecurity systems capable of real-time threat response [118].
- *Addressing the dual challenge of AI tools*: Efforts will focus on preventing the misuse of ML tools for malicious content, addressing privacy and transparency concerns, and mitigating misleading information to ensure ML tools positively contribute to cybersecurity [36].

### RQ9: What future work is suggested for DL models in cyber-attack detection?

- *Hybrid models*: Future research should integrate symbolic AI with DL to create models that combine reasoning and learning, enhancing robustness and interpretability through neurosymbolic systems [121].

- *Generalization*: Develop DL models resilient to adversarial attacks and capable of generalizing across different domains, using techniques like transfer learning and domain adaptation [36].
- *Efficient learning techniques*: Advance few-shot, zero-shot, self-supervised, and unsupervised learning methods to enable effective model training with limited or unlabeled data [85].
- *Scalable and distributed learning*: Enhance federated learning for privacy-preserving model training across distributed devices and improve algorithms for efficient distributed training [79].
- *Real-time and online learning*: Create DL models that can adapt and learn in real time from evolving data streams, using online and incremental learning approaches to avoid catastrophic forgetting [35].
- *Advanced neural architectures*: Utilize neural architecture search to design optimal network architectures and expand the application of graph NN to model complex data relationships [78].
- *Multi-modal learning*: Develop models that integrate and learn from multiple data types simultaneously, enhancing performance through cross-modal and multi-task learning frameworks [33].
- *Ethical and responsible AI* : Research techniques to identify and mitigate biases in DL models, ensure fairness, and establish frameworks for AI governance and compliance with ethical standards [122]

### RQ10: What future work is suggested for metaheuristic algorithms in cyber-attack detection?

- *Integration of metaheuristic algorithms*: Combining various metaheuristic algorithms (e.g., genetic algorithms, simulated annealing, ant colony optimization) to enhance feature selection methods [99].
- *Adaptation to high-dimensional data*: Developing metaheuristic-based feature selection methods specifically for high-dimensional datasets to effectively handle the curse of dimensionality and identify relevant features [105].
- *Dynamic feature selection*: Implementing methods that can adapt to changing data environments and evolving feature sets, allowing real-time updates to the selected feature subset as new data is available [109].
- *Improving computational efficiency*: Focus on optimizing the computational efficiency of metaheuristic-based feature selection methods, potentially using optimized search strategies or parallel and distributed computing techniques [98].
- *Hybrid approaches*: Exploring hybrid methods that combine metaheuristic algorithms with other techniques (filter-based, wrapper-based) to leverage the strengths of multiple approaches for better accuracy and efficiency [119].
- *Application to different domains*: Applying metaheuristic-based feature selection methods to various domains like cybersecurity, bioinformatics, and IoT to assess their effectiveness in different contexts, each with unique challenges and opportunities [106].

- *Automating the selection process*: Developing automated frameworks for feature selection that minimize human intervention, including creating user-friendly tools that can automatically select and tune metaheuristic algorithms based on dataset characteristics and problem requirements [65].

## Conclusion

The exploration undertaken in this research provides a comprehensive review of AI methodologies utilized in the area of cyber-attack detection. Our analysis underscores the pivotal role of ML, DL, and metaheuristic algorithms in refining the responsiveness and precision of cybersecurity systems. Key findings indicate that while AI technologies significantly enhance detection rates, they are also challenged by high computational demands and the necessity for vast, accurate new datasets.

In this comprehensive study, we have accurately analyzed various methods employed for the detection of cyber-attacks. Our evaluation has been gathered from an initial collection of 9084 papers, a methodical review based on title, abstract, and keywords, followed by a stringent application of inclusion and exclusion criteria and generated 68 high-quality studies. Key details such as publication details, citation counts, fields of study, methodologies, and datasets were extracted primarily from the introductory sections of each paper, ensuring a concentrated source of research for further analysis. These papers focus primarily on advanced techniques such as DL, ML, and metaheuristic approaches. A detailed comparison of these techniques against traditional methods, which has been systematically documented in comparative tables within the study. These tables serve as a valuable resource, showing the effectiveness of new and traditional approaches for various types and datasets of cyber-attacks.

A significant insight derived from this research is the key role of dimension reduction and feature selection in enhancing the efficacy of intrusion detection systems. We explored how various techniques affect performance metrics, underscoring the necessity of optimizing AI algorithm hyperparameters through heuristic methods to significantly improve effectiveness.

### Author contributions

Aya Hesham Salem wrote the manuscript and designed all figures and tables. Prof. Amr A. Abohanya reviewed and supervised the manuscript content. Dr. Safaa Azzam supervised the manuscript content. Prof. Osama Emam supervised the manuscript content.

### Funding

Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

### Data availability

No datasets were generated or analysed during the current study.

## Declarations

### Competing interests

The authors declare no competing interests.

Received: 4 May 2024 Accepted: 13 July 2024
Published online: 04 August 2024

## References

1. Parkar P, Bilimoria A. A survey on cyber security IDS using ML methods. Proceedings—5th International Conference on Intelligent Computing and Control Systems, ICICCS 2021, no. ICICCS, pp. 352–360, 2021, https://doi.org/10.1109/ICICCS51141.2021.9432210.
2. Musa NS, Mirza NM, Rafique SH, Abdallah AM, Murugan T. Machine learning and deep learning techniques for distributed denial of service anomaly detection in software defined networks—current research solutions. IEEE Access. 2024;12(January):17982–8011. https://doi.org/10.1109/ACCESS.2024.3360868.
3. Eswaran M, et al. Survey of cyber security approaches for attack detection and prevention. IEEE Access. 2023;12(1):1–6. https://doi.org/10.17762/turcomat.v12i2.2406.
4. Alsamiri J, Alsubhi K. Internet of things cyber attacks detection using machine learning. Int J Adv Comput Sci Appl. 2019;10(12):627–34. https://doi.org/10.14569/ijacsa.2019.0101280.
5. Ahmad W, Rasool A, Javed AR, Baker T, Jalil Z. Cyber security in IoT-based cloud computing: a comprehensive survey. Electronics (Switzerland). 2022;11(1):1–34. https://doi.org/10.3390/electronics11010016.
6. Morovat K, Panda B. A survey of artificial intelligence in cybersecurity. Proceedings—2020 International conference on computational science and computational intelligence, CSCI 2020, pp. 109–115, 2020, https://doi.org/10.1109/CSCI51800.2020.00026.
7. Uma M, Padmavathi G. A survey on various cyber attacks and their classification. Int J Netw Secur. 2013;15(5):390–6. https://doi.org/10.6633/IJNS.201309.
8. Rauf U, Mohsen F, Wei Z. A taxonomic classification of insider threats: existing techniques, future directions and recommendations. J Cyber Secur Mobil. 2023;12(2):221–52. https://doi.org/10.13052/jcsm2245-1439.1225.
9. Thanh SN, Stege M, El-Habr PI, Bang J, Dragoni N. Survey on botnets: incentives, evolution, detection and current trends. Future Internet. 2021. https://doi.org/10.3390/fi13080198.
10. Perwej Y, Qamar Abbas S, Pratap Dixit J, Akhtar N, Kumar Jaiswal A. A systematic literature review on the cyber security. Int J Sci Res Manag. 2021;9(12):669–710. https://doi.org/10.18535/ijsrm/v9i12.ec04.
11. AbuBakar A, Zolkipli MF. Cyber security threats and predictions: a survey. Int J Adv Eng Manag (IJAEM). 2023;5(2):733. https://doi.org/10.35629/5252-0502733741.
12. Parizad A, Hatziadoniu CJ. Cyber-attack detection using principal component analysis and noisy clustering algorithms: a collaborative machine learning-based framework. IEEE Trans Smart Grid. 2022;13(6):4848–61. https://doi.org/10.1109/TSG.2022.3176311.
13. Philosophical logic and artificial intelligence. Springer Netherlands, 1989. https://doi.org/10.1007/978-94-009-2448-2.
14. Pomerol J-C. Artificial intelligence and human decision making. Eur J Oper Res. 1997;99(1):3–25. https://doi.org/10.1016/S0377-2217(96)00378-5.
15. Dokur NB. Artificial Intelligence (AI) applications in cyber security. https://www.researchgate.net/publication/367253331.
16. Hua Li J. Cyber security meets artificial intelligence: a survey. Front Inf Technol Electron Eng. 2018;19(12):1462–74. https://doi.org/10.1631/FITEE.1800573.
17. Welukar JN, Bajoria GP. Artificial intelligence in cyber security—a review. Int J Sci Res Sci Technol. 2021. https://doi.org/10.32628/IJSRST218675.
18. Thomas T, Vijayaraghavan AP, Emmanuel S. Machine learning approaches in cyber security analytics. 2019. https://doi.org/10.1007/978-981-15-1706-8.
19. Barik K, Misra S, Konar K, Fernandez-Sanz L, Koyuncu M. Cybersecurity deep: approaches, attacks dataset, and comparative study. Appl Artif Intell. 2022. https://doi.org/10.1080/08839514.2022.2055399.
20. Nordin NS, et al. A comparative analysis of metaheuristic algorithms in fuzzy modelling for phishing attack detection. Indonesian J Electr Eng Comput Sci. 2021;23(2):1146–58. https://doi.org/10.11591/ijeecs.v23.i2.pp1146-1158.
21. Agrawal P, Abutarboush HF, Ganesh T, Mohamed AW. Metaheuristic algorithms on feature selection: a survey of one decade of research (2009–2019). IEEE Access. 2021;9:26766–91. https://doi.org/10.1109/ACCESS.2021.3056407.
22. Kuntla GS, Tian X, Li Z. Security and privacy in machine learning: a survey. Issues Inf Syst. 2021;22(3):224–40. https://doi.org/10.48009/3_iis_2021_242-258.
23. Peng J, Jury EC, Dönnes P, Ciurtin C. Machine learning techniques for personalised medicine approaches in immune-mediated chronic inflammatory diseases: applications and challenges. Front Pharmacol. 2021;12(September):1–18. https://doi.org/10.3389/fphar.2021.720694.
24. Alduailij M, Khan QW, Tahir M, Sardaraz M, Alduailij M, Malik F. Machine-learning-based DDoS attack detection using mutual information and random forest feature importance method. Symmetry (Basel). 2022;14(6):1–15. https://doi.org/10.3390/sym14061095.
25. Gawand MKSP. A comparative study of cyber attack detection & prediction using machine learning algorithms. Researchgate. 2013. https://doi.org/10.21203/rs.3.rs-3238552/v1
26. Sarker IH. CyberLearning: effectiveness analysis of machine learning security modeling to detect cyber-anomalies and multi-attacks. Internet Things. 2021;14:100393. https://doi.org/10.1016/j.iot.2021.100393.
27. Hasan M, Islam MM, Zarif MII, Hashem MMA. Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet Things. 2019;7:100059. https://doi.org/10.1016/j.iot.2019.100059.
28. Sarker IH, Kayes ASM, Badsha S, Alqahtani H, Watters P, Ng A. Cybersecurity data science: an overview from machine learning perspective. J Big Data. 2020. https://doi.org/10.1186/s40537-020-00318-5.
29. Rodriguez E, Otero B, Gutierrez N, Canal R. A survey of deep learning techniques for cybersecurity in mobile networks. IEEE Commun Surv Tutor. 2021;23(3):1920–55. https://doi.org/10.1109/COMST.2021.3086296.
30. Pourafshin F. Big data mining in internet of things using fusion of deep features. Int J Sci Res Eng Trends. 2021;7(2):1089–93.
31. Gu H, Wang Y, Hong S, Gui G. Blind channel identification aided generalized automatic modulation recognition based on deep learning. IEEE Access. 2019;7:110722–9. https://doi.org/10.1109/ACCESS.2019.2934354.

32. Hassan IH, Mohammed A, Masama MA. Metaheuristic algorithms in network intrusion detection. In: Comprehensive metaheuristics. Elsevier; 2023. p. 95–129. https://doi.org/10.1016/B978-0-323-91781-0.00006-5.
33. Rajwar K, Deep K, Das S. An exhaustive review of the metaheuristic algorithms for search and optimization: taxonomy, applications, and open challenges. Artif Intell Rev. 2023. https://doi.org/10.1007/s10462-023-10470-y.
34. Role of AI in cyber security through Anomaly detection and Predictive analysis. J Inf Educ Res. 2023;3:2. https://doi.org/10.52783/jier.v3i2.314.
35. Ozkan-Okay M, et al. A comprehensive survey: evaluating the efficiency of artificial intelligence and machine learning techniques on cyber security solutions. IEEE Access. 2024;12:12229–56. https://doi.org/10.1109/ACCESS.2024.3355547.
36. Sangwan RS, Badr Y, Srinivasan SM. Cybersecurity for AI systems: a survey. J Cybersecur Privacy. 2023;3(2):166–90. https://doi.org/10.3390/jcp3020010.
37. Mohamed N. Current trends in AI and ML for cybersecurity: a state-of-the-art survey. Cogent Eng. 2023. https://doi.org/10.1080/23311916.2023.2272358.
38. Kaur R, Gabrijelčič D, Klobučar T. Artificial intelligence for cybersecurity: literature review and future research directions. Inf Fusion. 2023. https://doi.org/10.1016/j.inffus.2023.101804.
39. Bin Hulayyil S, Li S, Xu L. Machine-learning-based vulnerability detection and classification in internet of things device security. Electronics (Switzerland). 2023. https://doi.org/10.3390/electronics12183927.
40. Asiri MM, et al. Hybrid metaheuristics feature selection with stacked deep learning-enabled cyber-attack detection model. Comput Syst Sci Eng. 2023;45(2):1679–94. https://doi.org/10.32604/csse.2023.031063.
41. Caviglione L, et al. Tight arms race: overview of current malware threats and trends in their detection. IEEE Access. 2021;9:5371–96. https://doi.org/10.1109/ACCESS.2020.3048319.
42. An JH, Wang Z, Joe I. A CNN-based automatic vulnerability detection. EURASIP J Wirel Commun Netw. 2023. https://doi.org/10.1186/s13638-023-02255-2.
43. Lucky G, Jjunju F, Marshall A. A lightweight decision-tree algorithm for detecting DDoS flooding attacks. In Proceedings—companion of the 2020 IEEE 20th international conference on software quality, reliability, and security, QRS-C 2020, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 382–389. https://doi.org/10.1109/QRS-C51114.2020.00072.
44. . Mynuddin M, Hossain MI, Uddin Khan S, Islam MA, Mohammed Abdul Ahad D, Tanvir MS. Cyber security system using fuzzy logic. In International Conference on Electrical, Computer, Communications and Mechatronics Engineering, ICECCME 2023, Institute of Electrical and Electronics Engineers Inc., 2023. https://doi.org/10.1109/ICECCME57830.2023.10252778.
45. ElDahshan KA, AlHabshy AAA, Hameed BI. Meta-heuristic optimization algorithm-based hierarchical intrusion detection system. Computers. 2022. https://doi.org/10.3390/computers11120170.
46. Soliman S, Oudah W, Aljuhani A. Deep learning-based intrusion detection approach for securing industrial Internet of Things. Alex Eng J. 2023;81:371–83. https://doi.org/10.1016/j.aej.2023.09.023.
47. Psychogyios K, Papadakis A, Bourou S, Nikolaou N, Maniatis A, Zahariadis T. Deep learning for intrusion detection systems (IDSs) in time series data. Future Internet. 2024;16(3):73. https://doi.org/10.3390/fi16030073.
48. SaiSindhuTheja R, Shyam GK. An efficient metaheuristic algorithm based feature selection and recurrent neural network for DoS attack detection in cloud computing environment. Appl Soft Comput. 2021;100: 106997. https://doi.org/10.1016/j.asoc.2020.106997.
49. Sanjeetha R, Kanavalli A, Gupta A, Pattanaik A, Agarwal S. Real-time DDoS detection and mitigation in software defined networks using machine learning techniques. Int J Comput. 2022;21(3):353–9. https://doi.org/10.47839/ijc.21.3.2691.
50. Gaur V, Kumar R. Analysis of machine learning classifiers for early detection of DDoS attacks on IoT devices. Arab J Sci Eng. 2022;47(2):1353–74. https://doi.org/10.1007/s13369-021-05947-3.
51. Jyothi KK, et al. A novel optimized neural network model for cyber attack detection using enhanced whale optimization algorithm. Sci Rep. 2024. https://doi.org/10.1038/s41598-024-55098-2.
52. Atawneh S, Aljehani H. Phishing email detection model using deep learning. Electronics (Switzerland). 2023. https://doi.org/10.3390/electronics12204261.
53. Asiri S, Xiao Y, Alzahrani S, Li T. PhishingRTDS: a real-time detection system for phishing attacks using a Deep Learning model. Comput Secur. 2024;141: 103843. https://doi.org/10.1016/j.cose.2024.103843.
54. AbdullahAlohali M, et al. Metaheuristics with deep learning driven phishing detection for sustainable and secure environment. Sustain Energy Technol Assess. 2023. https://doi.org/10.1016/j.seta.2023.103114.
55. Zavrak S, Yilmaz S. Email spam detection using hierarchical attention hybrid deep learning method. Expert Syst Appl. 2023. https://doi.org/10.1016/j.eswa.2023.120977.
56. Butt UA, Amin R, Aldabbas H, Mohan S, Alouffi B, Ahmadian A. Cloud-based email phishing attack using machine and deep learning algorithm. Complex Intell Syst. 2023;9(3):3043–70. https://doi.org/10.1007/s40747-022-00760-3.
57. Kitchenham S, Charters B. Guidelines for performing systematic literature reviews in software engineering. Technical report, Ver. 2.3 EBSE, vol. 1, no. January 2007, pp. 1–54, 2007. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.117.471&rep=rep1&type=pdf.
58. Shaukat K, Luo S, Chen S, Liu D. Cyber threat detection using machine learning techniques: a performance evaluation perspective. 1st Annual international conference on cyber warfare and security, ICCWS 2020—Proceedings, 2020, https://doi.org/10.1109/ICCWS48432.2020.9292388.
59. Prasad A, Chandra S. BotDefender: a collaborative defense framework against botnet attacks using network traffic analysis and machine learning. Arab J Sci Eng. 2024;49(3):3313–29. https://doi.org/10.1007/s13369-023-08016-z.
60. Wei Z, Rauf U, Mohsen F. E-Watcher: insider threat monitoring and detection for enhanced security. Ann Telecommun. 2024. https://doi.org/10.1007/s12243-024-01023-7.
61. Mohsen F, Rauf U, Lavric V, Kokushkin A, Wei Z, Martinez A. On identification of intrusive applications: a step toward heuristics-based adaptive security policy. IEEE Access. 2024;12:37586–99. https://doi.org/10.1109/ACCESS.2024.3373202.

62. Mihoub A, Ben Fredj O, Cheikhrouhou O, Derhab A, Krichen M. Denial of service attack detection and mitigation for internet of things using looking-back-enabled machine learning techniques. Comput Electr Eng. 2022;98(2021): 107716. https://doi.org/10.1016/j.compeleceng.2022.107716.
63. Alrowais F, Althahabi S, Alotaibi SS, Mohamed A, Hamza MA, Marzouk R. Automated machine learning enabled cybersecurity threat detection in internet of things environment. Comput Syst Sci Eng. 2023;45(1):687–700. https://doi.org/10.32604/csse.2023.030188.
64. Liu Z, Wang Y, Feng F, Liu Y, Li Z, Shan Y. A DDoS detection method based on feature engineering and machine learning in software-defined networks. Sensors. 2023. https://doi.org/10.3390/s23136176.
65. Omer N, Samak AH, Taloba AI, Abd El-Aziz RM. A novel optimized probabilistic neural network approach for intrusion detection and categorization. Alex Eng J. 2023;72:351–61. https://doi.org/10.1016/j.aej.2023.03.093.
66. Aljehane NO, Mengash HA, Hassine SBH, Alotaibi FA, Salama AS, Abdelbagi S. Optimizing intrusion detection using intelligent feature selection with machine learning model. Alex Eng J. 2024;91(January):39–49. https://doi.org/10.1016/j.aej.2024.01.073.
67. Talpur F, Korejo IA, Chandio AA, Ghulam A. ML-based detection of DDoS attacks using evolutionary algorithms optimization. 2024;24(5):1672. https://doi.org/10.3390/s24051672
68. Kumar A, Dutta S, Pranav P. Supervised learning for attack detection in cloud. Int J Exp Res Rev. 2023;31:74–84. https://doi.org/10.52756/10.52756/ijerr.2023.v31spl.008.
69. More S, Idrissi M, Mahmoud H, Asyhari AT. Enhanced intrusion detection systems performance with UNSW-NB15 data analysis. Algorithms. 2024;17(2):64. https://doi.org/10.3390/a17020064.
70. Chohan MN, Haider U, Ayub MY, Shoukat H, Bhatia TK, Ul Hassan MF. Detection of cyber attacks using machine learning based intrusion detection system for IoT based smart cities. EAI Endorsed Trans Smart Cities. 2023;7(1):1–7. https://doi.org/10.4108/eetsc.3222.
71. Singh A, Shibargatti A, Jena MA, Manvi S. Machine learning based detection of phishing websites in chrome. 1st Int Conf Emma-2021. 2024;2742: 020072. https://doi.org/10.1063/5.0184539.
72. Rexha B, Thaqi R, Mazrekaj A, Vishi K. Guarding the Cloud: an effective detection of cloud-based cyber attacks using machine learning algorithm. Int J Online Biomed Eng. 2023. https://doi.org/10.3991/ijoe.v19i18.45483.
73. Özalp AN, Albayrak Z. Detecting cyber attacks with high-frequency features using machine learning algorithms. Acta Polytech Hungarica. 2022;19(7):213–33. https://doi.org/10.12700/APH.19.7.2022.7.12.
74. Azeem M, Khan D, Iftikhar S, Bawazeer S, Alzahrani M. Analyzing and comparing the effectiveness of malware detection: a study of machine learning approaches. Heliyon. 2024;10(1): e23574. https://doi.org/10.1016/j.heliyon.2023.e23574.
75. Hammad M, Hewahi N, Elmedany W. Enhancing network intrusion recovery in SDN with machine learning: an innovative approach. Arab J Basic Appl Sci. 2023;30(1):561–72. https://doi.org/10.1080/25765299.2023.2261219.
76. Ribeiro MA, Pereira Fonseca MS, de Santi J. Detecting and mitigating DDoS attacks with moving target defense approach based on automated flow classification in SDN networks. Comput Secur. 2023;134(August): 103462. https://doi.org/10.1016/j.cose.2023.103462.
77. Yuan X, Han S, Huang W, Ye H, Kong X, Zhang F. A simple framework to enhance the adversarial robustness of deep learning-based intrusion detection system. Comput Secur. 2024. https://doi.org/10.1016/j.cose.2023.103644.
78. Devendiran R, Turukmane AV. Dugat-LSTM: deep learning based network intrusion detection system using chaotic optimization strategy. Expert Syst Appl. 2024. https://doi.org/10.1016/j.eswa.2023.123027.
79. Yaras S, Dener M. IoT-Based Intrusion Detection System Using New Hybrid Deep Learning Algorithm. Electronics. 2024;13(6):1053. https://doi.org/10.3390/electronics.
80. Hnamte V, Hussain J. Dependable intrusion detection system using deep convolutional neural network: a novel framework and performance evaluation approach. Telemat Inform Rep. 2023. https://doi.org/10.1016/j.teler.2023.100077.
81. Makkar A, Kumar N. An efficient deep learning-based scheme for web spam detection in IoT environment. Futur Gener Comput Syst. 2020;108:467–87. https://doi.org/10.1016/j.future.2020.03.004.
82. Aslan O, Yilmaz AA. A new malware classification framework based on deep learning algorithms. IEEE Access. 2021;9:87936–51. https://doi.org/10.1109/ACCESS.2021.3089586.
83. Akgun D, Hizal S, Cavusoglu U. A new DDoS attacks intrusion detection model based on deep learning for cybersecurity. Comput Secur. 2022;118: 102748. https://doi.org/10.1016/j.cose.2022.102748.
84. Aslan Ö. Separating malicious from benign software using deep learning algorithm. Electronics (Switzerland). 2023. https://doi.org/10.3390/electronics12081861.
85. Alzahrani IR, Allafi R. Integrating Ebola optimization search algorithm for enhanced deep learning-based ransomware detection in Internet of Things security. AIMS Math. 2024;9(3):6784–802. https://doi.org/10.3934/math.2024331.
86. Balajee RM, Jayanthi Kannan MK. Intrusion detection on AWS cloud through hybrid deep learning algorithm. Electronics. 2023;12(6):1423. https://doi.org/10.3390/electronics12061423.
87. Farhan BI, Jasim AD. Performance analysis of intrusion detection for deep learning model based on CSE-CIC-IDS2018 dataset. Indonesian J Electric Eng Comput Sci. 2022;26(2):1165–72. https://doi.org/10.11591/ijeecs.v26.i2.pp1165-1172.
88. Abbas S, et al. Evaluating deep learning variants for cyber-attacks detection and multi-class classification in IoT networks. PeerJ Comput Sci. 2024;10: e1793. https://doi.org/10.7717/peerj-cs.1793.
89. Alamyar AM, License RF. Detecting malicious attacks using cyber-security models using deep learning approach. pp. 0–26, 2023.
90. Salam A, Ullah F, Amin F, Mohammad A. Deep learning techniques for web-based attack detection in, MDPI, pp. 1–18, 2023.
91. Jullian O, Otero B, Rodriguez E, Gutierrez N, Antona H, Canal R. Deep-learning based detection for cyber-attacks in IoT networks: a distributed attack detection framework. J Netw Syst Manage. 2023;31(2):1–24. https://doi.org/10.1007/s10922-023-09722-7.

92. Ghazal SF, Mjlae SA. Cybersecurity in deep learning techniques: detecting network attacks. Int J Adv Comput Sci Appl. 2022;13(11):221–30. https://doi.org/10.14569/IJACSA.2022.0131125.
93. Otoum Y, Liu D, Nayak A. DL-IDS: a deep learning-based intrusion detection framework for securing IoT. Trans Emerg Telecommun Technol. 2022;33(3):1–14. https://doi.org/10.1002/ett.3803.
94. Chuang HM, Ye LJ. Applying transfer learning approaches for intrusion detection in software-defined networking. Sustainability (Switzerland). 2023;15(12):1–24. https://doi.org/10.3390/su15129395.
95. Chen L, Wang Z, Huo R, Huang T. An adversarial DBN-LSTM method for detecting and defending against DDoS attacks in SDN environments. Algorithms. 2023. https://doi.org/10.3390/a16040197.
96. Dey AK, Gupta GP, Sahu SP. Hybrid meta-heuristic based feature selection mechanism for cyber-attack detection in IoT-enabled networks. Procedia Comput Sci. 2022;218:318–27. https://doi.org/10.1016/j.procs.2023.01.014.
97. Dey AK, Gupta GP, Sahu SP. A metaheuristic-based ensemble feature selection framework for cyber threat detection in IoT-enabled networks. Decis Anal J. 2023;7(January): 100206. https://doi.org/10.1016/j.dajour.2023.100206.
98. Mange P, Lule A, Savant R. Advanced spam email detection using machine learning and bio-inspired meta-heuristics algorithms. Int J Intell Syst Appl Eng IJISAE. 2023;2024(4s):122–35.
99. Naeem H, Ullah F, Srivastava G. Classification of intrusion cyber-attacks in smart power grids using deep ensemble learning with metaheuristic-based optimization. Expert Syst. 2024. https://doi.org/10.1111/exsy.13556.
100. Albakri A, Alhayan F, Alturki N, Ahamed S, Shamsudheen S. Metaheuristics with deep learning model for cybersecurity and android malware detection and classification. Appl Sci (Switzerland). 2023. https://doi.org/10.3390/app13042172.
101. Diaba SY, Shafie-Khah M, Elmusrati M. Cyber security in power systems using meta-heuristic and deep learning algorithms. IEEE Access. 2023;11(February):18660–72. https://doi.org/10.1109/ACCESS.2023.3247193.
102. Alshammari MA, Abd El-Aziz AA, Hamdi H. Detecting traffic diversion using metaheuristic algorithm in SDN. Int J Intell Syst Appl Eng. 2024;12(9):369–79.
103. Alkanhel R, et al. Network intrusion detection based on feature selection and hybrid metaheuristic optimization. Comput Mater Continua. 2023;74(2):2677–93. https://doi.org/10.32604/cmc.2023.033273.
104. Mjahed O, El Hadaj S, El Guarmah EM, Mjahed S. Improved supervised and unsupervised metaheuristic-based approaches to detect intrusion in various datasets. CMES Comput Model Eng Sci. 2023;137(1):265–98. https://doi.org/10.32604/cmes.2023.027581.
105. Thomas M, Meshram BB. DoS attack detection using Aquila deer hunting optimization enabled deep belief network. Int J Web Inf Syst. 2024. https://doi.org/10.1108/IJWIS-06-2023-0089.
106. Mohsenabad HN, Tut MA. Optimizing cybersecurity attack detection in computer networks: a comparative analysis of bio-inspired optimization algorithms using the CSE-CIC-IDS 2018 dataset. Appl Sci. 2024;14(3):1044. https://doi.org/10.3390/app14031044.
107. Nazir A, Khan RA. A novel combinatorial optimization based feature selection method for network intrusion detection. Comput Secur. 2021;102: 102164. https://doi.org/10.1016/j.cose.2020.102164.
108. Ghanbarzadeh R, Hosseinalipour A, Ghaffari A. A novel network intrusion detection method based on metaheuristic optimisation algorithms. J Ambient Intell Human Comput. 2023;14(6):7575–92. https://doi.org/10.1007/s12652-023-04571-3.
109. Zeinalpour A, McElroy CP. Comparing metaheuristic search techniques in addressing the effectiveness of clustering-based DDoS attack detection methods. Electronics (Switzerland). 2024. https://doi.org/10.3390/electronics13050899.
110. Goyal N, Trivedi MC. Metaheuristic algorithms for optimization and feature selection in cloud data classification using convolutional neural network. J Inf Technol Manag. 2023;15(3):99–112. https://doi.org/10.22059/JITM.2023.93627.
111. Prasad A, Chandra S. Machine learning to combat cyberattack: a survey of datasets and challenges. J Model Simul. 2023;20(4):577–88. https://doi.org/10.1177/15485129221094881.
112. Yang Z, et al. A systematic literature review of methods and datasets for anomaly-based network intrusion detection. Comput Secur. 2022. https://doi.org/10.1016/j.cose.2022.102675.
113. Ferrag MA, Friha O, Hamouda D, Maglaras L, Janicke H. Edge-IIoTset: a new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning. IEEE Access. 2022;10:40281–306. https://doi.org/10.1109/ACCESS.2022.3165809.
114. Khraisat A, Gondal I, Vamplew P, Kamruzzaman J. Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecurity. 2019. https://doi.org/10.1186/s42400-019-0038-7.
115. Prasad A, Chandra S. PhiUSIIL: a diverse security profile empowered phishing URL detection framework based on similarity index and incremental learning. Comput Secur. 2024. https://doi.org/10.1016/j.cose.2023.103545.
116. Kim Y, Hakak S, Ghorbani A. DDoS Attack Dataset (CICEV2023) against EV Authentication in Charging Infrastructure. In 2023 20th Annual International Conference on Privacy, Security and Trust, PST 2023, Institute of Electrical and Electronics Engineers Inc., 2023. https://doi.org/10.1109/PST58708.2023.10320202.
117. Shafin SS, Karmakar G, Mareels I. Obfuscated memory malware detection in resource-constrained IoT devices for smart city applications. Sensors. 2023. https://doi.org/10.3390/s23115348.
118. Al-Hawawreh M, Sitnikova E, Aboutorab N. X-IIoTID: a connectivity-agnostic and device-agnostic intrusion data set for industrial internet of things. IEEE Internet Things J. 2022;9(5):3962–77. https://doi.org/10.1109/JIOT.2021.3102056.
119. Thota S, Menaka D. Botnet detection in the internet-of-things networks using convolutional neural network with pelican optimization algorithm. Automatika 2024;65(1):250–60. https://doi.org/10.1080/00051144.2023.2288486

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.