Dijlah Journal of Engineering Science (DJES) Vol. 3, No. 1, March, 2026, pp.164-172
ISSN: Printed: **3078-9656, Online: 3078-9664**, paper ID: 120

Detect cyber-attacks using machine learning

Hayder Sameer Noori
Southern Federal University, Department of software engineering

Article Info

Article history:

Received Nov.,10, 2025
Revised Dec.,17, 2025
Accepted Jan.,21, 2026

Keywords:

Cyber-attacks
Machine Learning
Detection

ABSTRACT

A key component of the workable Machine learning (ML) is the solution for physical object systems. which can be characterized as the ability of an intelligent smart device to alter its state and behavior. Both human and machine-generated data can contain useful information that machine learning (ML) can comprehend. Machine learning algorithms can be applied to a variety of tasks, including regression and classification. Additionally, in an Internet of Things network, security services are provided by Machine Learning (ML). The application of machine learning (ML) in cyberattack detection challenges is becoming more and more popular. ML may be applied to a variety of cyber-security prosecutions. Assuming that the summary's various study types were combined with Machine Learning (ML) techniques to determine the most accurate ways to detect offensive, individual poor groundwork exists on effective detection techniques advisable for practical physical object system environments

Corresponding Author:

Hayder Sameer Noori
Southern Federal University, Department of software engineering
Baghdad, Iraq

1. Introduction

In a world dominated by social media, online payments, cloud computing, and machine operations, devices and technology are improved fastly. However, as technology improves, cyber attack also improves, thus creating and developing new types of attacks, in addition to new methods, and techniques that enable attackers to breach more sophisticated or strictly regulated networks, inflect more harm, and perhaps escape detection [Ben15]. Moreover, customers' and businesses' data are now at risk of attacks, making it vital to specify ways to preserve data stability and protection on a worldwide scale [Al-+21]. It is increasingly crucial for countries and large corporations to have cybersecurity solutions in their networks to protect data security and privacy from cyberattacks. This paper offers a thorough evaluation of previous research on cyberattacks utilizing machine learning algorithms as a solution in order to comprehend this concept, its manifestations, and the strategies employed to counteract it. An overview of the cyberattack and an example of how machine learning models may identify cyberattacks are the goals of this article. In a culture that uses big data, social media, and other technologies more and more, information security and data integrity are often threatened., online payments, data recorded or handled over the web, and machine operations performed out via IT systems. The number of attacks and intensity of the harm done to targets by cyber attacks are both steadily rising thanks to the development of new strategies and techniques [Ben15]. Threats to cybersecurity are a worrying topic right now that has to be addressed globally. The majority of people's personal information is utilized and shared because almost everyone has smart technologies that are linked to the internet and many of them use social network platforms. There is a widespread lack of knowledge regarding attacks (types, characteristics, and potential effects), which makes it extremely difficult to guarantee adequate information security worldwide [Cho16]. It is difficult for the research community to identify novel assault kinds. The majority of earlier research papers surveyed attack detection [AL21]. The categorization models use machine learning techniques to identify the different types of attacks [Al +21]. This paper's major goal is to present a thorough analysis of prior research on cybersecurity and cyber attacks in order to understand this idea, its manifestations, and the methods used to detect it. This paper will discuss some machine learning models that are used to detect cyber attacks such as SVM, XGboost.
introduces the background for the types of cyber-attacks. The related work of detecting cyber-attacks using machine learning are outlined in Section 3. Section 4 introduces the metrics, Background in this section we define cyber-attacks and why machine learning methods are important for detecting cyber-attacks. In addition to the definition of several types of cyber-attacks. Define cyber-attacks in the worldwide reviews of literature, there are many

explanations of the phrases cyber-attacks, hackers, and cyber threat, all of which contain the same goal of jeopardizing the confidentiality, privacy, security, and accountability of the information [Ben15]. Cyber-attacks are evolving along with the innovation of new smart devices, thus new strategies and algorithms are built to detect these attacks, getting to even more difficult destination. Despite advancements in cybersecurity, traditional threats remain the most prevalent and frequently encountered. This paper outlines and analyzes several types of cyberattacks, as illustrated in Figure 1.Man-in-the-Middle Attack A man-in-the-middle (MitM) attack occurs when a malicious actor intercepts communication between two parties. In this scenario, every data packet transmitted from source A to destination B is first captured by the attacker, granting them access to the information being exchanged. This intrusion poses serious risks, including unauthorized access to sensitive data and the possibility of altering the contents of the packets before they reach their intended recipient [Cho16]. Man-in-the-middle (MitM) attacks include an intruder interfering with a two-party transaction [Ben15]. The attacker can quickly filter and grab data once he has interrupted the stream. Attacks against eavesdropping are another name for it. The MITM attack can take many different forms, such as collecting credentials or passwords, among others. These attacks typically take place on an insecure public Wi-Fi network where the attackers place themselves in between the victim's device and the open network. Without the user's knowledge, the visitor gives the attacker access to all information [Ben15]. In some instances, the attacker installs software to use malware to acquire information about the victim.

2- Research Question

Which algorithms are the most accurate and effective in cyberphysical systems contexts, and how might machine lea rning approaches outperform conventional methods in the detection of different cyberattacks, including malware, ph ishing, and DDoS attacks?

3- **Research Hypotheses**

Compared to conventional rulebased or signaturebased detection systems, machine learning algorithms are more acc urate at detecting and classifying several kinds of assaults, including DDoS, malware, phishing, and spam.In terms o f detection accuracy, F1 score, and robustness against false positives, collective learning techniques (such Random F orest, AdaBoost, and Majority Voting) perform better than singlemodel classifiers (like SVM and Decision Tree).Be cause deep learning architectures (such MLP and LSTM) can recognize nonlinear and temporal patterns in network t raffic data, they perform better when detecting sophisticated and complicated cyberthreatsto evaluate and categorize cyberthreats (such as malware, phishing, DDoS, and bruteforce attacks) and comprehend how they affect network pe rformance and information security.to look into and contrast how well various machine learning algorithmssuch as S VM, decision trees, random forests, XGBoost, AdaBoost, MLP, and LSTM—detect and mitigate online threats.
to assess how well crowdsourced learning techniques work to increase the precision and dependability of cyberattac k detection systems.

4- Research Objectives

to evaluate and categorize cyberthreats (such as malware, phishing, DDoS, and bruteforce attacks) and comprehend how they affect network performance and information security.to look into and contrast how well various machine l earning algorithms—such as SVM, decision trees, random forests, XGBoost, AdaBoost, MLP, and LSTM—
detect and mitigate online threats.to assess how well crowdsourced learning techniques work to increase the precisio n and dependability of cyberattack detection systems.

Research Approaches

To identify the most effective detection methods, such as datadriven analysis, algorithmic assessment, and comparis on modeling, this study uses a quantitative and experimental approach utilizing machine learning models.

Research Methodology

Attack types

2.1.2 The brute force A node's class and probability are used to determine each branch's Gini value, which shows which branch is more likely to occur. In this instance, C denotes the total number of categories, while Pi indicates the observed relative frequency of each class within the dataset.Brute force attacks involve systematically trying numerous combinations to unlock secured information—such as passwords or encrypted data—until the correct credentials are discovered, granting access to the protected content.Phishing is a deceptive technique used to obtain sensitive user information by posing as a trustworthy entity, such as a legitimate-looking website (often using unsecured HTTP protocols).The classification and probability of a node are used to calculate the Gini index for each branch, which helps identify the branch with the highest likelihood of occurrence. In this context, C stands for the number of distinct classes, and Pi reflects the proportion of each class within the dataset.Brute force methods rely on repeated attempts to breach secured systems by guessing credentials until successful access is achieved.

2.1.3 Malware

Malware encompasses various types of malicious software designed to compromise data confidentiality. Advanced systems utilizing gradient boosting techniques—particularly those tailored for unstructured data—employ strategies

such as parallel processing, tree pruning, hardware-level optimization, regularization, sparsity handling, and cross-validation to enhance performance and detection capabilities.It is tailored for unstructured data and uses the gradient boosting technique. It makes use of regularization, sparsity awareness, hardware optimization, parallelization, tree trimming, and cross-validation to The term "malware" describes a variety of destructive software programs that an attacker employs to compromise confidentiality. number of background communications that were categorized as background communications information. There are also several types of malware such as computer viruses and ransom ware. A network or a specific system can be destroyed by malware, which is malicious software. Malware includes both traditional forms of malicious software like viruses, Trojans, and worms and more recent forms like spyware and ransom ware. When a user clicks a dubious unsafe link, opens an email attachment, or installs software from an unreliable source, the virus goes into action and infects the network.

2.1.4 Phishing is the technique of gaining access to and stealing important customer information by pretending to be a trustworthy source (for instance, a website on http). Phishing is a fraudulent practice in which imposters send emails to a large number of recipients pretending to be someone else in order to make the message appear to be from a reputable or well-known source. The purpose of this is to obtain the recipients' personal information, such as credit card numbers. These users are persuaded to visit phony websites by their large discounts and attractive offers. Customers are drawn to this and click on the links in these emails, providing critical information such banking details, credit card information, login credentials, and other personal information without realizing that theft has taken place. One of the most prevalent and growing types of online threats today is phishing.

2.1.5 Spam

Spam is unwanted email messages are referred to as spam. In addition to taking up the recipient's time, these spam emails frequently contain Java applets that run when the message is opened. Cybercrime is the illicit dissemination of spam through the email transmission of immoral or illegal product marketing.

2.1.6 DoS and DDoS A denial-of-service attack floods servers, networks, or systems with traffic in order to deplete their information and resource capacity. The system cannot process the valid requests as a result. A denial-of-service assault (DoS) typically entails the coordinated efforts of a person or an organization to stop an Internet site, server, or service from operating effectively or at all. Its methods of execution, goals, and targets might vary. A distributed-denial-of-service (DDoS) assault is one in which the aggressor launches numerous attacks against the victim as opposed to initiating a single attack from a single compromised device. Since there are numerous attack categories and no research publication has yet utilized this dataset for multi-class classification for different forms of DDoS attack detection, using it for multi-class classification for AI and ML algorithms is a difficult issue. Furthermore, the dataset is nearly 26 GB in size, making it impossible to use on conventional computers for this study because greater processing power is needed to handle such a large amount of data. Reducing the dataset while preserving the distribution and data quality is crucial to overcoming this difficulty. The brute force Brute force attacks entail repeatedly attempting to access protected data (passwords, encryption, etc.) until the correct key is found, at which time the data was compromised. The attacker bombards systems, servers, or networks with more traffic and requests than they can manage, initiating multiple attacks at once, exhausting the bandwidth resources and bringing the system to a halt. For this attack, attackers also make use of numerous compromised devices.

2.2 Machine learning for cyber attacks

Machine learning is the act of teaching a model set of data with important attributes so that it can understand the underlying ideas and forecast the behavior of new data. It can categorize previously unknown data points or forecast what future data might look like. It can be used to create complex and focused phishing emails in the context of cybercrimes. Instead of being used in cybercrime to avoid filters, get through tests, and create tailored phishing emails, machine learning is utilized in cyber security to recognize similar malware and harmful connections. Comparing the two, it seems that machine learning is used considerably more consistently in cyber security. Future developments in evasive malware and phishing, however, could seriously endanger the cyber security sector. 3 Related Work Cyber-attacks harm companies of all sizes as well as people on a daily basis, but there is no much understanding about what is a cyber-crime [Al-+21]. Numerous studies focused on the mechanisms of cyber-attacks and, consequently, the defensive or mitigation measures against them. These tactics comprise deep learning and machine learning methods, Attacks that cause. Cyber threats can lead to various types of attacks, such as denial of service (DoS), distributed denial of service (DDoS), privilege escalation from user to root (U2R), and unauthorized access from remote systems to local machines (R2L). Additionally, malicious software encompasses a wide range of techniques including botnets, zero-day exploits, deceptive phishing schemes, SQL injection attacks, ransomware, cross-site scripting (XSS), harmful applications, and probing activities aimed at discovering system vulnerabilities [ASL20] [Par+19] [Hus+21].

3.1 Defensive machine learning strategies Support Vector Machine (SVM), naive Bayes, decision trees, random forests, logistic regression, neural networks, and hybrid approaches.

Defensive machine learning techniques SVM, clustering, association rules mining, neural networks, and mining for frequent patterns. 3.2 Random Forest Classifier (RFC) RFC [AL21] uses three primary parameters: The Gini index serves as a primary metric for determining how decision tree nodes split, evaluating the effectiveness of each division. It calculates the likelihood of each branch by considering the class label and its associated probability. Specifically, C denotes the number of distinct classes, while Pi reflects the relative frequency of each class within the dataset. An alternative method for guiding node splits based on probability is the entropy criterion, as shown in Equation (2). Random Forest Classifier (RFC) is a technique that constructs predictive models by utilizing randomly selected subsets of the data. The prediction result from each decision tree constructed is then acquired, and the prediction result with the most votes is the final prediction result [AAA20]. 4 Metrics used Some criteria must be chosen to assess the effectiveness of the approaches in order to compare the outcomes of various algorithms. The typical approach is to consider number of false positives and false negatives (i.e., the number of background communications classified as botnets) (i.e. the number of botnets labeled as background communications). Four scores can be applied as shown in figure .Support Vector Machines (SVMs) are supervised learning techniques commonly applied to tasks involving classification and regression. These models function by identifying a hyperplane that effectively separates data points belonging to different categories, and they are recognized for their strong generalization performance on unseen data. In this research, the LinearSVC implementation from the Sci-Kit Learn library was utilized, configured with the "ovr" (one-vs-rest) strategy to manage multiclass classification. The square-hinge loss was selected due to its computational advantages, and the regularization (or cost) parameter was fixed at 1 to mitigate overfitting. L2 regularization was employed during penalization, as L1 led to overly sparse model coefficients.Decision trees represent another form of supervised learning, where data is organized from the root node down to leaf nodes that assign class labels. These models divide the feature space into multiple segments using axis-aligned rectangles or hyperplanes. Known for their robustness against outliers and minimal preprocessing requirements, decision trees often serve as a baseline for evaluating other machine learning algorithms. In this research, the minimum number of samples required to split an internal node was set to 3, and the Gini index was selected as the criterion for determining splits. To preserve cost complexity, pruning was omitted, and the model was configured to consider all available features when identifying the optimal split.

3.3 Deep learning algorithms the author Rehak have created CAMNEP in 2008 and are used as network intrusion detection system and is defined as Cooperative Adaptive Mechanism for Network Protection (CAMNEP). The CAMNEP system employs a series of algorithms to detect anomalies that keep track of a method of anticipated network traffic and do comparison with actual traffic to spot differences that could be signs of attacks. Anomaly aggregators, trust models, and anomaly detectors are its three main layers that analyze the traffic. Each anomaly detection technique used by the anomaly detector layer uses a unique set of attributes to examine the Net Flows. Events are created from the output and transmitted to the trust models. The trust model groups traffic based on the Net Flows. Net flows that exhibit comparable behavioral tendencies group together. The composite output is produced by the aggregator layer by combining the unique findings of various anomaly detectors. In order to identify trends and anomalies that might point to new or developing assaults, machine learning algorithms can be trained on historical data of recognized cyberthreatsEnsemble Learning Algorithms3.5 XGBoost The gradient boosting framework is used by XGBoost [AL21], an ensemble machine learning technique, to predict outcomes from machine learning. The speed and scalability of XGBoost are widely known [KG20]. Extreme gradient boosting, a hybrid of gradient boosting and XGBoost, is the source of XGBoost. To obtain closer approximations, it makes advantage of sophisticated regularization and second-order gradients.XGBoost is a powerful algorithm It is tailored for unstructured data and uses the gradient boosting technique. It makes use of regularization, sparsity awareness, hardware optimization, parallelization, tree trimming, and cross-validation to improve the performance of the gradient boosting method. It is highly scalable and has higher computational speed on memory-restricted systems. The algorithm uses "Friedman mse" as a measure of the quality of a split and for classification utilizing probabilistic outputs, deviance is a loss function that is comparable to logistic regression.Adaptive Boosting (commonly referred to as AdaBoost) is an ensemble learning strategy designed to enhance the performance of a base model—such as decision trees—by iteratively focusing on the errors made by weaker classifiers and transforming them into stronger ones. Unlike algorithms that rely on random predictions, AdaBoost offers improved accuracy. However, it tends to be slower in computation compared to methods like XGBoost and is more susceptible to noise and outliers. The effectiveness of splits within AdaBoost is assessed using the Gini index.The Majority Vote Classifier (MV-4) is another ensemble approach that aggregates the outputs of several classifiers to produce a final decision. In this research, the four most effective models—Random Forest, AdaBoost, Decision Tree, and XGBoost—were integrated using majority voting to form the MV-4 classifier, which was then applied to identify various types of DDoS attacks.

3.4 Data mining techniques the author Ertoz have created MINDS in 2004 and are used by the Minnesota Intrusion Detection System to automatically identify threats. The amount of Net Flow having identical source IP address as the evaluated Net Flow, the amount of Net Flows toward the same target host, the amount of Net Flows towards the same target host and having identical source port, and the amount of Net Flows from the same source host having identical target port are used to generate data for each evaluated Net Flow. A Net Flow's anomaly value is determined by how far it is from the typical sample. Xu et al devised this algorithm. All of the Net Flows having identical source IP are used to generate the context for each Net Flow that needs to be assessed. Some categorization rules that categorize the traffic into regular and abnormal flows are used to identify anomalies. The Multi-Layer Perceptron (MLP) is a supervised learning algorithm that employs the back propagation learning method to learn a function that maps input features to output class labels in real-time., but requires proper optimization algorithms and hyper parameter tuning. In this study, the optimizer used is ad delta and the loss function is categorical cross-entropy. Different kinds of cyberattacks, including malware and DoS/DDoS attacks, can be identified and categorized using machine learning techniques. To attain the highest level of accuracy, the output layer uses the Soft ax activation function, while the hidden layers employ the rule activation function. Long Short-Term Memory (LSTM) networks are a specialized form of recurrent neural networks (RNNs) widely utilized in deep learning applications, particularly for analyzing and forecasting sequential or time-dependent data. It was developed to overcome the vanishing and exploding gradient problem encountered in traditional RNNs. In this study, the model uses the Adam optimizer and categorical cross-entropy as the loss function. Optimal classification performance was obtained by employing a two-layer LSTM architecture, each containing eight units, followed by a softmax output layer tailored for multiclass prediction.

4.1 Accuracy The accuracy ratio is the proportion of accurate predictions made across all samples. Since the accuracy value is still high even with poor prediction performance on the minority class, this metric is inappropriate in cases with unbalanced data sets.

4.2 Precision is the ratio of correctly predicted positive outcomes to all positive outcomes. Precision reflects the quantity of false positives.

4.3 Recall is the percentage of accurate predictions made out of all positive samples. Recall is a good indicator of how many false negatives there were

4.4 F1-Score The F1-Score, which is the harmonized average of Precision and Recall, is frequently used since it allows for a better understanding of the prediction performance of the model by balancing the previous measures.

Remember that having a poor recall is worse than having a low precision because it means that while most detected attacks are botnets (precision), most 7 botnet attacks go unnoticed (recall). Naturally, it is simple to have an excellent memory because all you have to do is categorize every communication as belonging to a botnet. Maximizing the f1 score while ensuring that the recall is not too weak is the compromise that has been decided upon. The Receiver Operating Characteristic (ROC) curve is an effective graphical method for assessing the performance of classification models. It displays the True Positive Rate (TPR) along the vertical axis and the False Positive Rate (FPR) along the horizontal axis. The TPR represents the fraction of actual positive cases that the model correctly classifies, while the FPR indicates the fraction of negative cases that are incorrectly labeled as positive.An ideal classifier would have a TPR of 1 and an FPR of 0, meaning that it correctly identifies all positive instances and no negative instances are incorrectly identified. This point is located at the top left corner of the ROC curve. A higher Area Under the Curve (AUC) value indicates stronger classification performance by the model. AUC ranges from 0 to 1, where 1 represents a perfect classifier and 0 represents a classifier that performs no better than random guessing.

3. **Implementation**

Different kinds of cyberattacks, including malware and DoS/DDoS attacks, can be identified and categorized using machine learning techniques These algorithms can be trained to recognize patterns and anomalies in network traffic, and can be used to send notifications to security engineers when an attack is detected. However, it is important to note that these algorithms can also be evaded by attackers who craft malware to evade detection. Additionally, Machine learning can be used to detect cyber-attacks in electric vehicles using LSTM based models, this research is still ongoing.

Table 1 shows The computational demands of different AI and machine learning algorithms frequently applied in detecting DDoS cyber threats are typically expressed based on two factors

Algorithm	Training Complexity	Prediction Complexity
SVM	$O(n^2 f + n^3)$	$O(nsv f)$
Decision Tree	$O(n^2 f)$	$O(f)$
XGBoost	$O(n f ntrees)$	$O(f ntrees)$
Random Forest	$O(n^2 f ntrees)$	$O(f ntrees)$
AdaBoost	$O(n f)$	$O(f ntrees)$
Neural Network	-	$O(fn_i1 + n_i1 n_i2 + ...)$

the volume of training data (n) and the number of input features (f), and other relevant parameters specific to each algorithm. This information can be used to choose an algorithm with a lower computational complexity when the performance of different algorithms is similar. The selection of a low computational complexity algorithm is important in situations where computational resources are limited.Distributed Denial of Service (DDoS) is a type of cyber-attack that uses a large number of compromised devices or systems to target a single or multiple victims. These attacks are well coordinated and can cause significant damage to the targeted systems in terms of bandwidth and power consumption, as well as loss of confidential data. Due to the high cost of these attacks, it is important to develop better Algorithms designed to identify various forms of DDoS attacks aim to achieve high detection accuracy while maintaining computational efficiency.Most studies in the literature have focused on the detection of DDoS threats as a binary classification problem, which only determines whether or not an attack was attempted. However, understanding the particular kind of DDoS assault that is aimed at the network or system is crucial for successful defense against it. By transforming the problem into a multi-label classification problem, this study introduces an Ensemble Classifier that efficiently detects various DDoS threats by combining the performance of the top four performing algorithms and comparing it with various Artificial Intelligence and Machine Learning (AI and ML) algorithms. The CICDDoS2019 dataset is a dataset that contains benign and the most up-to-date common DDoS attacks, which closely resembles real-world data in the form of PCAPs (packet capture files). The dataset also includes the results of network traffic analysis using CICFlowMeter-V3, which are labeled flows based on the timestamp, source and destination IPs, source and destination ports, protocols, and attack types The goal of the dataset is to provide realistic background traffic for use in DDoS attack detection research.The dataset was generated using the B-Profile system, which is a system that profiles the abstract behavior of human interactions and generates naturalistic benign background traffic in a test bed environment a test bed environment. The dataset includes the abstract behavior of 25 users based on the HTTP, HTTPS, FTP, SSH, and email protocols.Using whether or not an attack was attempted. However, understanding the particular kind of DDoS assault that is aimed at the network or system is crucial for successful defense against it. By transforming the problem into a multi-label classification problem, this study introduces an Ensemble Classifier that efficiently detects various DDoS threats by combining the performance of the top four performing algorithms while also evaluating its performance alongside a range of Artificial Intelligence and Machine Learning (AI/ML) techniques. The network traffic analysis findings from CICFlowMeter-V3 are also included in the collection. These results are labeled flows according to the date, protocols, attack kinds, source and destination IPs, and source and destination ports. The abstract nature of human relationships and produces background traffic that is benign and realistic. To address this challenge, it is important to reduce the dataset while maintaining the integrity of data along with the distribution. One way to achieve this is by using the Scikit-learn Python library, which provides various techniques for dimensionality reduction and feature selection. This can help to effectively train AI and ML models by reducing the dataset size without compromising the quality of the data.

The results from testing the AI and ML algorithms on the test dataset are presented in Table 2.

Algorithm	Accuracy Score	Design Parameters
SVM	92.75%	C = 1.0, penalty='l2', loss='squared hinge'
Decision Tree	98.99%	criterion='gini', samples split=3
Random Forest	99.24%	criterion='gini', samples split=2
XGBoost	98.59%	criterion='friedman mse', loss='deviance'
AdaBoost	99.01%	criterion='gini', splitter='best'
MJV-4	99.01%	Random Forest, AdaBoost Decision Tree, XGBoost
MLP	97.33%	activation='relu', optimizer='adadelta'
LSTM	98.16%	activation='softmax', optimizer='adam'

The data shows the accuracy score for each algorithm.Among the Among machine learning methods, the SVM classifier has the lowest accuracy score of all, with an accuracy score of 92.75% on a multi label dataset. With an accuracy score of 98.99%, the Decision Tree algorithm outperforms the SVM. When it comes to AI algorithms, LSTM performs marginally better than Multilayer Perceptron (MLP) since it has a higher accuracy score on the test dataset. The Multi-Layer Perceptron (MLP) model achieved an accuracy of 97.33%, outperforming the LSTM model by 0.84%, which recorded an accuracy of 98.17%. The ensemble learning approach incorporated four algorithms, among which Random Forest delivered the highest accuracy at 99.24%, exceeding the performance of Adaptive Boosting by a margin of 0.23%. With an accuracy score of 98.59%, the XGBoost algorithm performs worse than Random Forest and Adaptive Boosting. Furthermore, the classification accuracy of the Majority Voting (MJV-4) algorithm is 99.01%. Except for Random Forest, The accuracy obtained from the test dataset surpasses that of all other methods outlined in Section 4. Furthermore, the accuracy metrics for each algorithm are consistent with their corresponding F1-Scores, as presented in Table 3.

Table 3. F1-Score of Artificial Intelligence and Machine Learning Models

Threats	SVM	Decision Tree	Random Forest	XGBoost	AdaBoost	MJV-4	MLP	LSTM
BENIGN	0.90	0.90	0.95	0.46	0.90	0.90	0.92	0.89
DNS	0.90	0.99	0.99	0.98	0.99	0.99	0.97	0.98
LDAP	0.95	0.99	0.99	0.98	0.99	0.99	0.98	0.98
MSSQL	0.89	0.99	0.99	0.98	0.99	0.99	0.99	0.99
NTP	0.98	0.99	1.00	0.99	0.99	0.99	0.99	1.00
NetBIOS	0.93	0.99	0.99	0.99	0.99	0.99	0.97	0.97
SNMP	0.97	0.99	0.99	0.99	0.99	0.99	0.97	0.97
SSDP	0.80	0.99	0.99	0.98	0.99	0.99	0.97	0.98
UDP	0.89	0.99	0.99	0.99	0.99	0.99	0.95	0.97
Syn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
UDP-lag	0.95	0.99	1.00	0.99	0.99	0.99	0.97	0.98
WebDDoS	0.52	0.38	0.53	0.00	0.32	0.38	0.00	0

Analysis

The results obtained from testing the AI and ML algorithms on the CICDDoS2019 dataset were analyzed and discussed in this section.Among the evaluated algorithms, Support Vector Machine (SVM) recorded the lowest accuracy at 92.75%, trailing Decision Tree by approximately 6.24% and XGBoost by around 5.84%, as indicated in Table 2. The F1-Score for SVM was also comparatively lower for certain attack categories, such as WebDDoS. However, it achieved a perfect F1-Score of 1.00 for detecting SYN-based threats, demonstrating its effectiveness in that specific scenario. The ROC curve for SVM aligns with its F1-Score, showing high AUC values for attacks like DNS, NetBIOS, and SNMP, while others such as LDAP and SSDP exhibited lower AUC values.

In contrast, the Decision Tree algorithm achieved an accuracy of 98.99%, outperforming SVM. Its F1-Scores were consistently strong across most attack types, with the exception of WebDDoS. The ROC curve further confirms its superior performance, with an AUC of 0.99 for the majority of threats.XGBoost yielded an accuracy of 98.59%, placing it above SVM but slightly below Decision Tree. While its F1-Score was lower for WebDDoS and benign traffic, it successfully identified SYN attacks with perfect precision—something neither SVM nor Decision Tree accomplished. The ROC curve for XGBoost revealed a notably low AUC of 0.45 for benign traffic and 0.00 for WebDDoS, indicating poor detection capability for those categories.Overall, the Decision Tree algorithm demonstrated the most reliable performance, achieving the highest accuracy and F1-Scores across the majority of attack types.Within the scope of this study, the Random Forest algorithm demonstrated the highest classification accuracy, reaching 99.24%. This performance surpasses that of SVM by approximately 6.50% and exceeds XGBoost by around 0.65%. Among all evaluated models, Random Forest also achieved the most favorable F1-Score, particularly excelling in differentiating between benign and malicious traffic. Its F1-Score for benign samples

was 0.95, outperforming all other algorithms. Furthermore, the ROC curve for Random Forest confirms its effectiveness, showing an area under the curve (AUC) of 1.00 across all attack categories, indicating near-perfect detection capability.Among all the AI and ML models evaluated in this study, Random Forest achieved the highest accuracy, reaching 99.24%. This result is approximately 6.50% greater than that of the SVM model and about 0.65% higher than XGBoost. Random Forest also recorded the top F1-Score among the tested algorithms, demonstrating strong capability in distinguishing between benign and malicious traffic. Notably, it achieved an F1-Score of 0.95 for benign traffic, outperforming all other models in this category. The ROC curve further supports its effectiveness, showing an area under the curve (AUC) of 1.00 across all attack types, indicating near-perfect classification performance. Additionally, Adaptive Boosting (AdaBoost) was employed as another ensemble learning method in this study.By integrating the strengths of Random Forest, AdaBoost, Decision Tree, and XGBoost, the Majority Voting Classifier (MV-4) achieves an impressive accuracy of 99.01%. This performance is nearly equivalent to that of AdaBoost and slightly below the accuracy of Random Forest. According to the F1-Scores presented in Table 3, the classifier perfectly identifies SYN attacks with a score of 1.00, while most other threat categories maintain an F1-Score of 0.99. The ROC curve further confirms MV-4's effectiveness, showing an area under the curve of 1.00 for all attack types, along with micro and macro average scores of 1.00, indicating that MV-4 serves as an optimal classifier for this dataset [2].Compared to machine learning algorithms, artificial intelligence algorithms perform marginally worse. MLP's accuracy score is 97.33%, which is 4.58% greater than SVM's and almost 2% lower than Random Forest's. Additionally, the MLP algorithm's F1-Score is lower for certain attack types, like WebDDoS cyber threats. The area under the curve for WebDDoS is 0.52, which influences the area under the curve for the macro average to 0.96, according to the RoC Curve for MLP. With a score of 1.00, it can identify additional risks, though. Overall, it can be concluded that Random Forest algorithm performs the best among all the algorithms as it has the highest accuracy score, F1-Score, and ROC Curve for most of the attack types. The other algorithms such as Decision Tree, Adaptive Boosting, and Majority Voting Classifier

Conclusion and future work

It is concluded that the MV-4 and Random Forest classifiers perform the best among all the algorithms evaluated in this study for detecting multiple types of DDoS cyber threats. The results obtained from the F1-Score and RoC curve also support this conclusion. In future work, solutions can be developed and deployed to prevent these critical cyber threats by using AI and ML algorithms.In addition, the performance of the algorithms can be further improved by using different feature selection techniques and by using more advanced architectures for neural networks. Another avenue for future work would be to test these algorithms on other datasets for DDoS detection to see if the results are consistent. Furthermore, the performance of the algorithms could be compared against traditional methods for DDoS detection such as rule-based systems and signature-based systems. This would provide a more complete understanding of The performance of AI and machine learning techniques is evaluated in contrast to conventional approaches to highlight their relative effectiveness. Overall, the use of AI and ML algorithms for DDoS detection has the potential to provide more effective and efficient solutions for protecting networks from DDoS attacks.There is a widespread lack of knowledge regarding attacks (types, characteristics, and potential effects), which makes it extremely difficult to guarantee adequate information security worldwide [Ben15]. In this article we show several types of attacks that are been spread among all the companies. We also show methods used to detect these attacks using machine learning methods. In addition, we give an overview in the metrics that can be used for evaluating the detection of these attacks.

References

[1] [AAA20] Hasan Alkahtani, Theyazn HH Aldhyani and Mohammed Al-Yaari introduced a framework for adaptive anomaly detection targeting digital entities within cyberspace. This work was published in the journal Applied Bionics and Biomechanics in the year 2020.

[2] [Al-+21] Belal Al-Fuhaidi et al. "Literature Review on Cyber Attacks Detection and Prevention Schemes". In: 2021 International Conference on Intelligent Technology, System and Service for Internet of Everything (ITSS-IoE). IEEE. 2021, pp. 1–6.

[3] [AL21] Rafa Alenezi, A., & Ludwig, S. A. (2021). Explainability of cybersecurity threats data using SHAP. 2021 IEEE Symposium Series on Computational Intelligence (SSCI), 1–10. IEEE.

[4] [ASL20] Aiyanyo, I. D., Samuel, H., & Lim, H. (2020). A comprehensive survey of machine learning applications in both defensive and offensive cybersecurity strategies. Applied Sciences, 10(17), 5811.

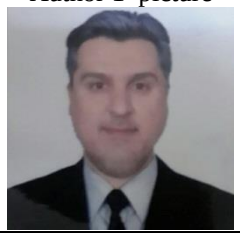
[4] [Ben15] Andreea Bendovschi. "Cyber-attacks–trends, patterns and security countermeasures". In: Procedia Economics and Finance 28 (2015), pp. 24–31.

[5] [Cho16] Abdullahi Chowdhury. "Recent cyber security attacks and their mitigation approaches–an overview". In: International conference on applications and techniques in information security. Springer. 2016, pp. 54–65.

[6] Husák, M., et al. (2021). Predictive approaches in cyber defense: Present applications and future research directions. Future Generation Computer Systems, 115, 517–530.
[7] [KG20] Rajesh Kumar and S Geetha. "Malware classification using XGboostGradient boosted decision tree". In: Adv. Sci. Technol. Eng. Syst 5 (2020), pp. 536–549. 8
[8] Park, M., et al. (2019). Evaluating threats in Android systems connected to IoT devices through the lens of situational awareness. Wireless Communications and Mobile Computin

BIOGRAPHIES OF AUTHORS (10 PT)

The recommended number of authors is at least 2. One of them as a corresponding author.
Please attach clear photo (3x4 cm) and vita. Example of biographies of authors:

Author 1 picture 	Hayder Sameer Noori is an engineer based in Baghdad, who is married and holds Iraqi nationality. Hayder completed his university education in 2002 with a Bachelor of Software Engineering from Al-Mansour University College. In 2017, he earned a Master's degree with an "Excellent" grade from Southern Federal University. Professionally, Hayder has been working as a Chief Engineer at the Ministry of Education since August 15, 2005.
2 Author 2 picture	Mini cv
Author 3picture	Mini cv
Author 4 picture	**Mini cv**
Author 5picture	**Mini cv**