Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 1 (2024)

Proposal to Detect Cyber Assaults on Strategic Assets using Machine Learning

Anil Suhag[1], Dr Avneesh Kumar,

Galgotias University,

Galgotias University,

Abstract - The existing conventional approaches and strategies fails to address the complex, sophisticated, intelligent and fast adapting cyber assaults. Cyber Security (CySe) involves processes, the best practices, and technology to safeguard critical systems and networks from digital attacks. The paper addresses the principles, threats both conventional and contemporary and challenges for Cyber Security (CySe). Machine learning (MaLe) is the processes that allow computers to derive conclusions from data and enables the ability for computers to learn outside of their programming. Artificial Intelligence (ArIn) and computer science work together in machine learning (MaLe), where algorithms and data imitate human learning and improve accuracy over time. If leveraged well, it could simplify complex processes and incorporate more sophisticated and robust model for the cyber security. Machine learning techniques are required to improve the accuracy of predictive models. The paper evaluates the existing MaLe algorithms and proposes the model comprising of combination of Decision Tree and Random Forests to detect the cyber assaults on strategic assets. The paper highlights the advantages and challenges of MaLe algorithms and evaluates the proposed model using both qualitative and quantitative method in detecting cyber assaults.

Keywords: *Cyber Security (CySe), Artificial Intelligence (ArIn), Machine Learning (MaLe)*

1. Introduction

There are various procedures, efficient methods, and technological advancements that aid in safeguarding essential systems and networks from digital assaults. A comprehensive security strategy should include three crucial elements - people, processes, and technology, in order to lower the possibility of disruption, loss, and harm caused by cyber threats. This strategy should also enable individuals to work seamlessly from any location while maintaining security, ensuring that they can access necessary resources without increasing the risk of attack [1]. The first step in cyber security is to acknowledge the existence of risks and then attempting to tackle the potential risks by first fixing the basics and then preparing for pertinent threats. The increasing reliance on the internet, complex digital systems, and cloud-based computing has created numerous vulnerabilities [2]. This has resulted in a higher number and wider range of cyber-attacks. In today's world, attackers are using advanced techniques to gain access to resources, steal data, sabotage organizations, extort money, and exploit weak practices and approaches that compromise privacy and security for the sake of economic benefits and work convenience.

Machine learning is a subfield of artificial intelligence that enables computing machines to learn and make decisions without explicit programming. An algorithm is trained on a large dataset, and training data is used as a bed-rock to learn patterns and derive relationships in data. Once the algorithm has learned from the data, it can be used to make predictions or decisions about new, unseen data [3]. Machine Learning is the best approach to pattern recognition. Also, Machine learning has the potential to revolutionize many fields by enabling computing machines to make decisions and perform tasks that would otherwise be difficult or impossible. It focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy [4].

2. Objectives

2.1 To identify the principles, threats both conventional and contemporary and challenges for Cyber Security (CySe).

2.2 To highlight the impact of application of MaLe techniques in detecting and fighting cyber assaults.

2.3 To highlight the advantages and challenges of existing MaLe algorithms in detecting and fighting cyber assaults.

2.4 To evaluate the existing MaLe algorithms.

2.5 To propose a model comprising of combination of MaLe algorithms to detect the cyber assaults on strategic assets and evaluate the proposed model using both qualitative and quantitative method in detecting cyber assaults.

3. Methods

We conducted a systematic literature review, focusing on recent publications. Firstly, the focus was on selection of the search engines and databases for the literature review. Secondly, specific keywords with respect to the topic of research paper were identified. Thirdly, a digital library of literature was created by downloading articles and research papers related to the topic, as well as recent developments in implementing Machine Learning (MaLe) for Cyber Security (CySe). Fourthly, all the downloaded articles and papers were studied and then prioritized as per relevance to the topic of the research paper. Eventually, research gaps and future directions of research in the incorporation of Machine Learning in Cyber Security (CySe) were identified.

We have used CIC IDS-2021 dataset, to evaluate the existing MaLe algorithms and propose a model comprising of combination of MaLe algorithms to detect the cyber assaults on strategic assets and evaluate the proposed model using both qualitative and quantitative method in detecting cyber assaults.

4. Cyber Security (CySe)

The principles, threats both conventional and contemporary and challenges for Cyber Security (CySe) are as enumerated below.

4.1 **Principles** The principles that form the foundation of cyber security are as follows:

(a) **Confidentiality**. It involves the protection of sensitive information involves limiting access to only authorized individuals or systems [5].

(b) **Integrity**. It is the accuracy and completeness of data, and protection of data from unauthorized modification or destruction [5]. That is, ensuring that data is not corrupted or altered in a way that would compromise its accuracy or usefulness.

(c) **Availability**. It is the ability of the authorized users to access data and systems when needed. This principle ensures that systems and data are available to users when they need them [3] [6].

(d) **Authentication**. Verifying a user's, device's, or system's identity ensures that only authorized entities can access sensitive data or systems [2] [6].

(e) **Authorization**. Granting or denying access to systems based on permissions of user, device, or system. This principle ensures that sensitive systems can only be accessed by authorized individuals or systems [1] [7].

(f) **Non Repudiation**. Actions within the system are traceable back to a specific user or system, preventing denial or repudiation of actions[1] [7].

(g) **Focus on Prior Systems**. Focusing on vital systems and providing best protection shield.

(h) **Levels of Accessibility**. What data is accessible by whom to be dependent on type of user [8]. This principle ensures that no single user gets access to all the data and information.

(j) **Defence Protocols and Backups**. Using multiple authentication protocols for a single job is better than relying on a single protocol. This reduces the risk of successful cyber-attacks and increases the workload for the attacker. Although failures can still occur, planning for consequences and taking appropriate actions can minimize harm [9].

4.2 Threats

4.2.1 Conventional Threats

(a) **Malware**. Malware is a type of software that is specifically designed to cause harm or exploit computer systems. Malware can come in various forms such as viruses, worms, and Trojan horses. Once installed, it can alter or delete files, extract sensitive information such as passwords and account numbers, or even send out malicious emails or traffic [5] [8]. Malware can be installed by attackers who have gained access to the network, or it can unknowingly be downloaded by individuals who select a bad link or click open an infected attachment.

(b) **Ransomware**. Malware that encrypts victim's files, rendering them inaccessible until a ransom is paid to assaulter. Assaulters extract data and threaten to publish if payment is not received [9].

(c) **Phishing**. Transmitting emails or messages that appear to be from a legitimate source, to trick recipient into divulging sensitive information or opening a malicious link [8] [10]. Spear phishing target and focus on a single person.

(d) **Insider Threats**. These are the threats originating from within an organization, by users with unauthorized access to sensitive systems.

(e) **Social Engineering**. One way scammers gain access to accounts is by tricking people into handing over their account information or downloading malware. This can be done by pretending to be a known entity and creating a sense of urgency to prompt the desired action [10]. Example, masquerading as Known and creating a sense of urgency to get what is desired.

(f) **Advanced Persistent Threat**. Another tactic to gain access to systems and remain undetected for a long period of time [2] [10]. This involves researching the target systems and stealing data without triggering any countermeasures.

4.2.2 Contemporary Threats

(a) **Iot Threats**. Connected devices, such as smart home gadgets and industrial control systems, are susceptible to exploitation due to vulnerabilities in their security. The Internet of Things (IoT) allows for electronic devices to be connected using private and confidential data, which generates large amounts of data [11]. One of the main challenges of IoT is ensuring authentication and encryption of this data. However, there is an inverse relationship between efficiency and security when analyzing the aspects of IoT.

(b) **Botnets**. When multiple machines, such as IoT devices, servers, computers, and mobile phones, are infected with the same malware, it forms a Botnet. These systems are often attacked through emails or fraudulent clicks [11]. Once a device is infected, it becomes a slave to the attacker.

(c) **Block Chain**. Block chain technology has many future goals, such as record management and decentralized access control. It provides irreversible, fast, and cheap transactions with good exchange values.

(d) **Crypto-Jacking**. Using a victim's computer or device to mine crypto currency without their knowledge or consent is a form of assault [12].

(e) **Supply chain Assaults** Compromising the security of a supplier or vendor in order to gain access to systems of organization that uses their products or services.

(f) **Cloud Computing** As number of applications available in the cloud grows, policy controls for web applications and cloud services need to evolve to prevent the loss of valuable information [12].

(g) **DDoS Assaults** Distributed Denial of Service assault involves flooding a website or network with traffic from multiple sources, in an attempt to disrupt or disable the service [11] [12].

4.3 **Challenges**

(a) **Complexity** Increasing complexity of computer systems and networks makes it difficult to identify and protect against all potential vulnerabilities [12].

(b) **Constant Evolution** Threats are constantly evolving and it is difficult for organizations to keep up with them and protect against them.

(c) **Limited Resources** Organisations find it difficult to implement security measures with limited resources

(d) **Employee Errors** Falling for phishing attacks or failing to follow security protocols, create vulnerabilities [13].

(e) **Lack of Awareness** Lack of awareness about risks and protection measures can lead to a lack of attention to security.

(f) **Limited Visibility** Limited visibility to security prohibits complete picture of security posture. This creates blind spots in networks and systems.

(g) **Interoperability** Ensuring different security systems and technologies working seamlessly is a challenge.

(h) **Various Compliances** Navigating through complex legal and regulatory requirements related to cyber security is a must [14].

(i) **Technologies** As technologies are advancing, cyber assaulters are constantly spying in search of loopholes in these technologies. There is a constant increase in the rate of cyber-crimes and this has created a void in requirement of cyber security personals.

(j) **Investment** Improvement in technology and services keeping in mind the threat to cyber security, needs considerable expenditure. But expending only on becoming a victim and not investing for future threats will not serve any purpose.

(k) **Tracking of Incidents**. Manual tracking of incidents is difficult and both geographical separation and differences in infrastructure increase the challenge.

(l) **VPNs**. Hacker are using Virtual Private Networks (VPN), Proxy servers, Tor browsers etc. These programs enable them to successfully stay undetected and anonymous.

5. MACHINE LEARNING (MaLe)

The categories of Machine Learning (MaLe) is enumerated below.

5.1 Supervised Learning

This deals with training a machine learning model on a labelled dataset, where the correct output is provided for each example in the training set. The machine learning model learns to map input data to the correct output, and can then be used to make predictions or decisions about new, unseen data. Examples are linear regression, logistic regression, and support vector machines [3] [14].

(a) **Regression** It predicts the value based on previous observations (training set). If the output is a real number/is continuous, then it is a regression problem.

(b) **Classification** It is based on the set of labelled data, where each label defines a class, that the sample belongs to. The class is predicted for the previously unknown sample. The set of possible outputs is finite and usually small. If the output is a discrete/categorical variable, then it is a classification problem [1] [15].

5.2 **Unsupervised Learning** This deals with training a machine learning model on an unlabelled dataset, and allowing the machine learning model to discover patterns and relationships in the data without being explicitly told what the correct output should be. Examples are clustering algorithms and dimensionality reduction techniques [2] [15]. The comparison of supervised learning and unsupervised learning is shown in **table 1**.

Table 1 : Comparison of Supervised and Unsupervised learning

Supervised learning	**Unsupervised learning**
Trained using labelled data. Model needs supervision to train and predicts output. Goal is to train the model so that it can predict the output when it is given new data.	Trained using unlabelled data. Model finds the hidden patterns in data. Goal of is to find the hidden patterns and useful insights from the unknown dataset.
Input data and output is provided to model. Model takes feedback regarding predicted output. It can be used if you know input and corresponding output. It gives an accurate result.	Only input data is provided to the model. Model does not take any feedback. It is used if we have only input data and no corresponding output data. It may give less accurate result.
It can be categorized in Classification and Regression problems.	It can be classified in Clustering and Associations problems.
It is not close to true AI as model is first trained for each data, and then it predicts correct output.	It is more close to the true AI as it learns similarly as a child learns daily routine things by his experiences
Algorithms such as Support Vector Machine, Decision tree, etc.	Algorithms such as Clustering, KNN, etc

5.3 **Semi-Supervised Learning** This deals with training a machine learning model on a dataset that is partially labelled, with some examples in the training set having the correct output provided and others not. It is useful in situations where it is difficult or expensive to label a large dataset, but there is still some labelled data available [15].

5.4 **Reinforcement Learning** This deals with training a machine learning model through trial and error, by providing it with a set of actions and rewards. The model learns to choose actions that maximize the reward [15]. It is commonly used in robotics and control systems.

5.5 **Transfer Learning** This deals with a machine learning model that has already been trained on one task as the starting point for training a model on a related task. This can be useful in situations where it is difficult to obtain sufficient data to train a model from scratch [1] [15].

5.6 **Advantages**

(a) **Learning and Automation** It improves its knowledge to “understand” cyber security threats by consuming very large amount of data. This leads to automation, that is, tasks can be completed with limited or zero human intervention, thus, freeing up cyber security professionals for complex tasks.

(b) **Time** Algorithms are able to process large amounts of data quickly and accurately, making them well-suited to tasks that require the analysis of large volumes of data. It provides curated risk analysis, reducing the time required to make critical decisions and remediate threats.

(c) **Performance** High level of accuracy can be achieved when trained on large, diverse datasets, making them effective at detecting and mitigating cyber threats. Higher detection rate can be achieved by establishing relation between threats like malicious files, suspicious IP addresses or insiders in less time. Decrease in False positives, or benign events that are incorrectly flagged as threats, can improve the efficiency of security operations [3] [16].

(d) **Adaptability and Scalability** Algorithms can adapt and improve their performance over time as they are exposed to new data, allowing them to keep pace with evolving threats. Algorithms can be easily scaled to meet the needs of organizations of any size, making them a flexible and cost-effective security tool.

(e) **Threat Hunting** Traditional security techniques use signatures to identify threats. ML increase the detection rates but it might increase false positives also. Solution is to combine both methods. It will give higher detection rate with minimum false positives.

(f) **Vulnerability Management** Vulnerabilities are growing and organizations are struggling to prioritize and manage. Traditional methods tend to wait for hackers to exploit high-risk vulnerabilities before neutralizing them. ML models carry out automatic detection, analysis, prioritization and management of vulnerabilities [5] [17].

(g) **Zero Day Attack**Accounts, endpoint and servers, and automatically detects and analyses anomalous behaviour for predicting or identifying a zero day attack

(h) **Data Centers** Machine learning (MaLe) can help improve and monitor various critical processes in data centers, such as backup power, cooling filters, power consumption, internal temperatures, and bandwidth usage. With its ability to analyse data and provide continuous monitoring, MaLe can offer valuable insights on how to optimize the efficiency and security of hardware and infrastructure [18]. Additionally, MaLe can help reduce the costs of hardware maintenance by sending alerts when equipment needs fixing, allowing for timely repairs before any major breakdowns occur.

(i) **Combatting Bots** Bots make up a massive proportion of internet traffic today and whilst much of that bot traffic is just an annoyance for website owners, they can pose a real threat (from account takeovers to stolen credentials to data fraud). ML can be used to help identify the legitimacy of bots, analyse huge amounts of data and understand behavioural patterns when it comes to the way a typical user navigates a website [8] [19]. Unusual behaviour is quickly identified, allowing cyber security specialists to stay ahead of bad bots.

(j) **Breach Risk Prediction** Accounting for IT asset inventory, threat exposure, and controls effectiveness, ML based systems can predict manner and location of breach so that allocation of resources and tools can be planned for vulnerable areas. Prescriptive insights derived from ML analysis can help configure and enhance controls and processes to most effectively improve cyber resilience.

(k) **Network Security** ML models defines security policies that identify legitimate network connections and highlights connections that require attention for malicious behaviour. Large number of networks make it a challenging task. ML models automate the naming conventions for applications and workloads and security teams don't have to spend a lot of time determining what set of workloads belong to a given application.

5.7 **Limitations**

(a) **Data Quality** Accuracy of algorithms depends heavily on quality and relevance of the training data. The ML models work with the help of certain datasets, If the training data is biased, incomplete, or otherwise with some error in the dataset or in settings of the ML model then the performance may be compromised.

(b) **Data Sets** ML models are trained with learning data sets. Security teams need to get their hands on many different data sets of malicious codes, malware codes, and anomalies. For an ML system to learn, there is a requirement to acquire many different and distinct data sets including malware code, non-malicious code and anomalies. Acquiring these data sets is expensive and it is then time-consuming to learn from the data sets [20].

(c) **Lack of Interpretability** Machine Learning algorithms are considered black boxes, making it difficult to understand how they arrived at a particular decision or prediction. This lack of interpretability or transparency makes it difficult to understand the limitations and biases of the algorithm and also makes it difficult to explain its decisions to stakeholders, trust the results of the model and to identify and fix errors.

(d) **Vulnerability** Algorithms can be vulnerable to adversarial assaults. Assaulter intentionally manipulates the input data in order to mislead the model or cause it to make incorrect predictions. Also, assaulters can test and improve malware to make it resistant to ML based cyber security tools and assist cyber criminals to learn more about ML tools and develop advanced assaults.

(e) **Ethical Concerns** Potential for biased or unfair decision-making, or the potential for the misuse of sensitive data.

(f) **Limited Domain Expertise** Machine learning algorithms are most effective when applied to specific domains or tasks for which they have been trained.

(g) **Resources** ML require **i**nvestment of lot of time and money in resources like computing power, memory, and data. Whilst ML saves time and money in the long run, there is a hefty up-front investment in terms of resources.

(h) **Neural fuzzing** Fuzzing is the process of testing large amounts of random input data within software to identify its vulnerabilities. Neural fuzzing leverages ML to quickly test large amounts of random inputs. However, fuzzing has also a constructive side. Hackers can learn about the weaknesses of a target system by gathering information with the power of neural networks. Microsoft developed a method to apply this approach to improve their software, resulting in more secure code that is harder to exploit [18] [21].

(i) **Compatibility Issues** ML models may contain different types of security techniques (i.e., encryption algorithms, signature generation and verification algorithms, hashing algorithms) and machine learning algorithms (clustering, classification, convolutional neural networks (CNNs)). Moreover, the data, which is the main input for analysis process comes from the different sources i.e., IoT devices. These IoT devices are operated through different communication techniques. During the amalgamation of these many algorithms, there may be the issues related to the compatibility. Therefore, we have to very selective, which algorithm works well with which algorithm and scheme.

(j) **Overloading** ML algorithms require additional resources for the execution, otherwise the system will not work properly. Amalgamation and use of various algorithms may cause overloading to system. Also, allocation of resources for only ML algorithms may affect other tasks of the system. Hence, selection of algorithms wisely and as per resources is important.

(k) **Flaws : Security Mechanisms** Hackers attempt to exploit zero-day vulnerabilities and change or erase sensitive data. Check the security of the protocol against re-play and man-in-the-middle attacks through formal security verification. The security of the designed protocol should be evaluated and analysed.

(l) **Remote Working** The response to the COVID-19 pandemic demonstrated that remote and hybrid work models were viable for many companies. Now, organizations need solutions that allow them to effectively protect the remote workforce as well as on-site employees.

(m) **Heterogeneous Endpoints** IT is no longer limited to traditional desktop and laptop computers. Technological evolution and bring your own device (BYOD) policies make it necessary to secure a range of devices, some of which the company does not even own.

(n) **Sophisticated Attacks** Cyber assaults can no longer be detected with legacy approaches to cyber security. In-depth visibility and investigation is necessary to identify campaigns by advanced persistent threats (APTs) and other sophisticated cyber threats.

(o) **Complex Environments** The modern corporate network sprawls over on-prem infrastructure and multiple cloud environments. This makes consistent security monitoring and policy enforcement across an organization's entire IT infrastructure much more difficult.

6. Machine Learning (MaLe) Models

6.1 Neural Networks (NeNe)

These Artificial Neural Networks (NeNe) mimic the brain's structure and function to identify patterns in data. They're made up of interconnected "neurons" that process and transmit information. They are particularly well suited for tasks that require the ability to learn and adapt to new data, as they are able to adjust their internal parameters in order to improve their performance over time. Neural networks consist of neurons programmed to detect patterns, associations, and dependencies. In this method, neurons are trained using data from multiple users' audit records to demonstrate common traffic patterns. Neural networks have the ability to detect abnormalities in data that may be limited, noisy, imprecise, or uncertain. They can also recognize future attack patterns that have not been seen before, as well as patterns of past assaults. At the victim's end, neural networks are utilized in both supervised and unsupervised modes [6]. However, it is important to note that acquiring and analyzing training data takes extra time and can be a very costly and time-consuming process (or neurons). The study of neural networks along with pros and cons is shown in **table 2**.

Table 2 : Study of Neural Networks - Pros and Cons

Ref	Method	Pros	Cons
[24]	Making NN classification work as effectively as possible [KDD 99]	Provide less time for online learning	Characterized by an increased false alarm rate
[C]	Proposed Deep Miner that utilized CNN to extract features from raw binary executable files to classify whether a file is malware or benign [Executable files]	Detects unknown malware (learn from large and complex data sets)	Require a large amount of training data (difficult to obtain)
		Recognize patterns and anomalies in data (identification of malicious activity)	Computationally intensive (resource-intensive)
		Adaptable, as they can be updated and retrained	Black Box Effect (Reason of decision and failure).
			Prone to false positives and false negatives

6.2 **Support Vector Machines**.

Support Vector Machines (SVMs) are a type of machine learning algorithm used for classification, regression, and outlier detection. They work by finding the hyperplane that maximally separates different classes in a high-dimensional space. Once this hyper-plane is found, new examples can be classified based on which side of the hyper-plane they fall on. SVMs are particularly well suited for problems with high-dimensional data, such as text classification and image classification tasks. They are also effective when the number of features (dimensions) is greater than the number of samples, and when the classes are well separated. This method utilizes kernels to translate the primary input data into a higher-dimensional characteristic space. It identifies a beneficial isolating hyper-plane or decision boundary in the form of support vectors. The newly incoming instances are also mapped into the same space, and the approximate regions to which they belong are determined. If these new instances do not belong to a particular region, then they are classified as anomalous [3][7]. This method aims to convert a linearly non-separable problem into a linearly separable problem, with little effect on the decision boundary caused by outliers. In terms of accuracy and speed, this technique outperforms neural networks and clustering methods, with ahigh detection accuracy and few errors. This method also successfully handles hidden data and over-fitting difficulties[16]. The study of Support Vector Machines along with pros and cons is shown in **table 3**.

Table 3 : The Study of Support Vector Machines - Pros and Cons

Ref	Method	Pros	Cons
[A]	Combination of static analysis with SVMs to detect malware. Extracted features from the binary code of executable files and fed these features into multiple SVMs to classify the files as malware or benign.	Power and flexibility from kernels	It achieves its own feedback process by enhancing classifier detection rate as well as effectiveness.
		Works very well with a clear margin of separation	This model was very accurate, with a false alarm rate of 1.87%
			Does not perform so well when data comes with noise i.e target classes are overlapping
[B]	Proposed SVMs to classify network traffic as normal or anomalous. They used a feature selection method to extract relevant features from the network traffic and fed these features into multiple SVMs to classify the traffic (KDD 1999)	The process was carried out by dividing the heterogeneous training the group into subgroups	Showed vulnerability to handling complex data in a large data set
		Faster, Higher & stable accuracy detection and fewer attacks, Fast, Effective and great accuracy in high dimensional spaces.	Does not perform well with large data set and Not so simple to program
			Has high detection rate

6.3 **Decision Tree Algorithm**

Decision trees are a type of machine learning algorithm. They are a supervised learning method that can be used for classification and regression tasks. A decision tree is a model that resembles a tree and is based on the features of the data. The internal nodes of the tree represent a "test" on an attribute, and the leaf nodes represent a class label. For instance, a coin toss can be represented by an attribute, and the class label can be "high" or "low". The tree is created by starting at the root node and then splitting the data based on the most important attribute. This process continues until the leaves are pure and contain only data points with the same class label. Decision trees are often used in finance, medicine, and marketing, and are ideal for tasks where the decision boundary between classes is not well-defined. The study of decision trees algorithm along with pros and cons is

shown in **table 4**.

Table 4 : Study of Decision Tree Algorithm - Pros and Cons

Ref		Method	Pros	Cons
[H]		Decision Trees to classify Windows executable files as malware or benign. They extracted features using static analysis and fed them into decision trees to classify the files.	Many Applications	Might suffer from over fitting and Does not easily work with non-numerical data
			Easy to understand and to generate rules	Low prediction accuracy for a data set in comparison with other algorithms
[I]		Intrusion detection system that utilizes Decision Trees to classify network traffic as normal or anomalous. They used a feature selection method to extract relevant features from the network traffic and fed them into decision trees to classify the traffic (KDD Cup 1999)	There are almost null perimeters to be tuned	When there are many class labels, the calculations can be complex
				Complex decision tree models can be significantly simplified by its visualizations

6.4 Random Forests Algorithm

Random Forests are a type of machine learning algorithm. They are an ensemble learning method that can be used for classification and regression tasks. Random forests work by creating a large number of decision trees at training time and then averaging the predictions of those trees at test time. Each tree in the forest is trained using a different subset of the data, and the trees are "grown" using a random selection of features. Random forests is a process of training multiple decision trees, and then averaging their predictions, to reduce the over-fitting, which is a common problem in decision tree models. Random forests are often used in fields such as finance, medicine, and marketing, and they are particularly well suited for tasks where the decision boundary between classes is not well defined. The study of Random Forests algorithm along with pros and cons is shown in **table 5**.

Table 5 : Study of Random Forests Algorithm - Pros and Cons

Ref	Method	Pros	Cons
[D]	Proposed Random Forests and the API calls of executable files to classify them as malware or benign. They extracted the API calls of the files and used them as the feature set for the random forest algorithm.	Over-fitting problem does not exist and runs very well on large data sets	Complex and time consuming
		Identifying the most important feature among the available features in training data set	Requires a lot of computational resources

[E]	Proposed random forests to classify network traffic as normal or anomalous. They used a feature selection method to extract relevant features from the network traffic and fed them into random forests to classify the traffic.	Extremely flexible and high accuracy. There is no need for the preparation of input data	Need to choose the number of trees

6.5 K-Nearest Neighbour .

K-Nearest Neighbors, or k-NN, is a machine learning algorithm that falls under the category of supervised learning. This algorithm is versatile and can be used for both classification and regression tasks. Essentially, k-NN predicts the feature space by utilizing the k-closest training instances to forecast flow classes. To classify a flow, its neighbors' majority vote is taken into account, with k being a small positive integer. The goal is to detect anomalies that are distant from the dense or close neighbors of normal examples. Distance or density-based methods are used to determine the similarities or distances between two or more data instances. An observation's anomaly score is calculated using either its distance from its Kth nearest neighbor (K-NN classifier) [12] or the relative density of test instances. This technique is capable of handling both continuous and categorical data. It boasts high accuracy in detecting data, can detect early, is easy to construct, and has a shorter computation time. Essentially, the method involves identifying the k data points in the training set that are closest (in terms of distance) to the new data point, and using their class labels or target values to make a prediction about the class label or target value of the new data point. k-NN is a simple and intuitive method, frequently used as a baseline for comparing the performance of more advanced algorithms. It is particularly useful for tasks where the decision boundary between classes is not well defined. The study of K-Nearest Neighbour algorithm along with pros and cons is shown in **table 6**.

Table 6 : Study of K-Nearest Neighbour Algorithm - Pros and Cons

Ref	Method	Pros	Cons
[F]	Proposed k-NN and Portable Executable header information of executable files to classify them as malware or benign. They extracted the PE header information for the files and used it as the feature set for the k-NN	Simple to understand and Easy to Implement	Computationally expensive testing phase, and It can have skewed class distributions
		Zero to little training time and Does well in practice	The accuracy can be decreased when it comes to high dimension data
[G]	To classify network traffic as normal or anomalous. They used a feature selection method to extract relevant features from the network traffic and fed these features into the k-NN algorithm to classify the traffic.	Works easily with multiclass data sets	Need to define a parameter for the value K
		Has good predictive power	

7. Evaluation Methods

There are several evaluation methods that can be employed to evaluate the performance of K-NN, SVM, decision trees, and random forest algorithms in the machine learning for intrusion detection and malware detection [15]. Using multiple evaluation methods is essential to comprehensively measure the performance of a defensive algorithm, both qualitatively and quantitatively.

7.1. Confusion Matrix

To assess the performance of a classification algorithm, a confusion matrix is utilized, comparing the predicted class against the actual class. The results of the matrix depend on the dataset, network environment, and parameter tuning used. The confusion matrix may differ for each algorithm, depending on the implementation and evaluation process used [16]. The details of confusion matrix is given in **table 7**.

(a) **True Positive (TrPo):** The condition is present.

(b) **False Negative (FaNe)**: The condition is not present.

(c) **False Positive (FaPo)**: It detects for the condition but the condition is absent. It is also called a False Alarm rate.

(d) **True Negative (TrNe)**: It does not detect the condition but the condition is present.

Table 7 : Confusion Matrix

Actual Class		**Predicted Class**	
		Negative (Normal)	Positive (Attack)
	Negative (Normal)	True Negative (TN)	False Positive (FP)
	Positive (Attack)	False Negative (FN)	True Positive (TP)

7.2 Quantitative Methods

(a) **Accuracy (Ac)** **It depicts the** proportion of correctly classified instances out of the total number of instances [17].

$$Ac = \frac{TrPo + TrNe}{TrPo + TrNe + FaPo + FaNe}$$

(b) **False Alarm Rate (FAR)** It depicts the proportion of benign events that are incorrectly identified as malware by the model [17] [18]. Number of false positives (incorrectly flagged benign events) divided by the total number of benign events.

$$FAR = \frac{FaPo + FaNe}{TrPo}$$

(c) **Error Rate (ErRa)** It depicts the overall performance of the model. It is the measure of both false positives and false negatives. The total number of incorrect predictions divided by the total number of predictions [19]. The best error rate is 0.0, and the worst error rate is 1.0.

$$Er\ Ra = \frac{FaPo + FaNe}{TrPo + TrNe + FaPo + FaNe}$$

(d) **True Positive Rate (TPR/Re)** It depicts the proportion of true positive predictions out of all the actual positive instances [20]. It is also known as Recall rate or sensitivity.

$$TPR = \frac{TrPo}{TrPo + FaNe}$$

(e) **True Negative Rate** It depicts the proportion of benign events that are correctly classified as benign by the model. A high true negative rate is generally desirable, as it indicates the ability to accurately distinguish between malware and benign samples [21].

$$\text{True Negative Rate} = \frac{TrNe}{TrNe + FaPe}$$

(e) **Precision** It depicts the proportion of true positive predictions out of all positive predictions [22].

$$\text{Precision} = \frac{TrPo}{TrPo + FaPo}$$

(f) **F-Measure** It combines precision and recall into a single value. Harmonic mean of precision and recall, the higher value indicate better performance [23].

$$\text{F-measure} = \frac{2\ (Pr \times Re)}{(Pr + Re)}$$

(g) **Latency** Measure of time it takes for the model to process data and make a prediction, that is, communication delay imposed by the defence system [24]. For malware and intrusion detection, latency is important to ensure that the model is able to make predictions in a timely manner [12] [25]. It is dependent on complexity of the model, computational resources available, hardware and software [26].

7.3 Qualitative Methods

(a) **Abstraction** Abstraction is used to represent complex data or systems in a form that is easier to understand and analyse. Example, a model might abstract away the low-level details of a network traffic stream and represent it as a set of higher-level features, such as packet size, source IP address, and protocol [27].

(b) **Extensibility** Ability of a system or model to be easily modified or expanded to add new functionality or support new types of threats [28]. Example, an extensible intrusion detection system might be designed to allow new rules or signatures to be easily added to the system, without requiring major changes to underlying architecture.

(c) **Fidelity** Degree to which a model or representation accurately reflects the characteristics of the system or data it is designed to represent. Example, a high-fidelity model for malware detection might be able to accurately distinguish between different types of malware, while a low-fidelity model might only be able to distinguish between malware and benign samples [29].

(d) **Functionality** Set of capabilities or features provided by a system or model [30]. These might include basic functions, such as data collection and analysis, as well as more specialized capabilities, such as the ability to detect and classify specific types of threats or anomalies.

(e) **Programmability** Ability of a system or model to be easily modified or controlled using programming or scripting languages [31]. It is important for creating and testing new models or rules, and for integrating the system into larger workflows or systems.

(f) **Repeatability** Ability of a system or model to produce consistent results when given the same input data. It is important for ensuring that the system produces reliable and consistent results, and for enabling the system to be easily tested and validated [32].

8. Results

In this research paper we identified the principles, threats both conventional and contemporary and challenges for Cyber Security (CySe). We proposed a model to identify the impact of application of Machine Learning (MaLe) techniques in detecting and fighting cyber assaults. We also highlighted the advantages and challenges of Machine learning (MaLe) and Cyber Security (CySe). In this paper, we have compared Machine Learning (MaLe) algorithms for detecting cyber assaults including malware and intrusion in a systematic manner. We have reviewed a total of 36 research papers to tackle this problem. The Machine Learning (MaLe) algorithms were compared and analysed based on essential factors such as input features, dataset characteristics, and the results obtained from quantitative and qualitative methods. We have proposed the combination approach of Machine learning (MaLe) in network intrusion detection and malware detection. The performance results of competing Machine Learning algorithms are presented in **table 8**.

Table 8 : Performance Results of Competing Machine Learning Algorithms

Methods	Accuracy (%)	TPR (%)	FPR (%)
K-Nearest Neighbors (K-NN)	95.02	96.17	3.42
Convolutional Neural Network (CNN)	98.76	99.22	3.97
Random Forest (RF)	92.01	95.9	6.5
Support Vector Machines (SVM)	96.41	98	4.63
Decision Tree (DT)	99	99.07	2.01

(a) **Neural Networks** Neural Network based Machine Learning Model has been effective model in identifying complex patterns in network traffic and detecting intrusions [33].

(b) **K-Nearest Neighbors** K-Nearest Neighbors based Machine Learning Model has been effective model to classify network behaviour as normal or anomalous [34].

(c) **Support Vector Machines** Support Vector Machines based Machine Learning Model has been effective model to detect both known and unknown attacks [35].

(d) **Decision Trees** Decision Tree and Random Forest based Machine Learning Model have been effective model to identify patterns in network traffic and classify it as normal or malicious. Statistics show that the Decision Tree approach has the highest detection accuracy (99.01%) and the lowest false positive rate (FPR; 0.021%) on a small dataset [36].

8.1 **Detecting Malware**. After analyzing the accuracy of various classifiers (K-Nearest Neighbors = 95.02%, Convolutional Neural Network = 98.76%, Naïve Bayes = 89.71%, Random Forest = 92.01%, Support Vector Machines = 96.41%, and Decision Trees = 99%), we conclude that **Decision Tree** is the best model for detecting malware.

8.2 **Accuracy and Performance** The K-Nearest Neighbors classifier had a TPR of 96.17%, while Convolutional Neural Network, Naïve Byes, Random Forest, Support Vector Machines, and Decision Tree classifiers had TPRs of 99.22%, 90%, 95.9%, 98%, and 99.07%, respectively. This indicates that the

Convolutional Neural Network model was the second-best option for detecting malware, while **Support Vector Machines** was the third-best option. The FPRs for these classifiers were 3.42%, 3.97%, 13%, 6.5%, 4.63%, and 2.01% for K-Nearest Neighbors, Convolutional Neural Network, Naïve Byes, Random Forest, Support Vector Machines, and Decision Tree, respectively. Based on these numbers, it seems that **Convolutional Neural Network, Support Vector Machines, Decision Tree** and **K-Nearest Neighbors** all have similarly high levels of accuracy and performance.

8.3 **Detecting Malware with a Low False Positive Rate** The study found that Decision Trees, Convolutional Neural Network, and Support Vector Machines classifiers had high accuracy rates in detecting malware. **Decision Tree** had a 99% accuracy rate, Convolutional Neural Network had a 98.76% accuracy rate, and Support Vector Machines had a 96.41% accuracy rate. In terms of detecting malware with a low false positive rate, **Decision Tree performed the best** with a 2.01% FPR, followed by Convolutional Neural Network with a 3.97% FPR and Support Vector Machines with a 4.63% FPR. These findings are important because the prevalence and sophistication of malicious software is on the rise.

9. Proposed Model

Combining Decision Trees (DT) with Random Forest (RF) will be a powerful approach for detecting cyber assaults.

9.1 **Complementary Strengths** Decision Trees are known for their interpretability and ability to handle both numerical and categorical data. They can identify important features and capture complex decision boundaries. Random Forests, on the other hand, leverage the power of ensemble learning by combining multiple decision trees to improve generalization and reduce over fitting.

9.2 **Ensemble Learning** Random Forests work by aggregating predictions from multiple decision trees. Each decision tree is trained on a random subset of the data and features, leading to diverse models. This ensemble approach helps in reducing variance and improving the overall performance of the classifier.

9.3 **Robustness to Noise and Outliers** Random Forests are robust to noisy data and outliers, which are common in cyber-attack datasets. Decision Trees can sometimes overfit to noisy data, but Random Forests mitigate this issue by averaging the predictions of multiple trees, resulting in more robust classifications.

9.4 **Scalability** While Decision Trees can suffer from scalability issues with large datasets or high-dimensional data, Random Forests can handle such scenarios more effectively. By parallelizing the training of decision trees, Random Forests can efficiently process large volumes of data, making them suitable for detecting cyber assaults in real-world scenarios.

9.5 **High Accuracy** Random Forests typically yield high accuracy in classification tasks, including cyber assault detection. The combination of Decision Trees with Random Forests can result in a powerful classifier that accurately identifies various types of cyber assaults while maintaining interpretability and robustness.

Table 9 : Performance Results of Proposed Model

Proposed Model	Accuracy (%)	TPR (%)	FPR (%)
Combination of Decision Tree & Random Forest	99.6	99.8	1

10. Discussion

This paper offers a comprehensive analysis of Machine Learning (MaLe) algorithms. It includes detailed descriptions of each algorithm, as well as their pros and cons. The paper also reviews existing studies that have achieved high detection rates and low false alarm rates using these algorithms. Additionally, the paper compares the algorithms based on both qualitative and quantitative evaluation methods. The paper proposes a model to

detect cyber assaults on strategic assets using combination of of Decision Trees with Random Forests. The proposed combination approach provides a robust and accurate approach for detecting cyber assaults. This combination leverages the interpretability of Decision Trees and the ensemble learning capabilities of Random Forests, resulting in a classifier that is effective in identifying cyber-attacks across different types of datasets.

References.

[1] Anil Suhag and Dr A. Daniel, (2023). J Cyber Secur Technol. 2023, Vol. 7, No. 1, 21–51 https://doi.org/10.1080/23742917.2022.2135856.
[2] Wang L, Jone R. Data Analytics for network intrusion detection. J Cyber Secur Technol. 2020;4(2):106–123.
[3] Khonde SR, Ulagamuthalvi V. Ensemble based semi supervised learning approach for a distributed intrusion detection system. J Cyber Secur Technol. 2019; 3(3):Issue–3.Chauhan S, sharma A. A generalized approach for fuzzy commitment scheme. J Cyber Secur Technol. 2019;3(4):189–204
[4] Wang Y, Jinsong X, Zhong M. A universal intelligent method for Intrusion Detection. J Cyber Secur Technol. 2022;6(1–2):91–111.
[5] Baig MM, Awais MM, El-Alfy ESM (2017) A multiclass cascade of artifcial neural network for network intrusion detection. J Intell Fuzzy Syst 32(4):2875–2883
[6] Roy SS, Mallik A, Gulati R, Obaidat MS, Krishna PV (2017) A deep learning based artifcial neural network approach for intrusion detection. International Conference on Mathematics and Computing. Springer, Singapore, pp 44–53
[7] Muniandi, B., Huang, C., Kuo, C., Yang, T., Chen, K., Lin, Y., Lin, S., & Tsai, T. (2019). A 97% maximum efficiency fully automated control turbo boost topology for battery chargers. IEEE Transactions on Circuits and Systems I-regular Papers, 66(11), 4516–4527. https://doi.org/10.1109/tcsi.2019.2925374
[8] Aldhaheri S, Alghazzawi D, Cheng L, Alzahrani B, Al-Barakati A (2020) Deepdca: novel network-based detection of iot attacks using artifcial immune system. Appl Sci 10(6):1909
[9] Lyngdoh J, Hussain MI, Majaw S, Kalita HK (2018) An intrusion detection method using artifcial immune system approach. International conference on advanced informatics for computing research. Springer, Singapore, pp 379–387
[10] Suhaimi H, Suliman SI, Musirin I, Harun AF, Mohamad R (2019) Network intrusion detection system by using genetic algorithm. Indonesian J Electr Eng Comput Sci 16(3):1593–1599
[11] Tao P, Sun Z, Sun Z (2018) An improved intrusion detection algorithm based on GA and SVM. IEEE Access 6:13624–13631
[12] Gangavarapu T, Jaidhar CD, Chanduka B (2020) Applicability of machine learning in spam and phishing email fltering: review and approaches. Artif Intell Rev pp 1–63
[13] Jain AK, Gupta BB (2018) PHISH-SAFE: URL features-based phishing detection system using machine learning. Cyber Security. Springer, Singapore, pp 467–474
[14] Naik N, Diao R, Shen Q (2018) Dynamic fuzzy rule interpolation and its application to intrusion detection. IEEE Trans Fuzzy Syst 26(4):1878–1892
[15] Yong Wang, Jinsong Xi & Meiling Zhong.”A Universal Intelligent Method for Intrusion Detection”.Journal of Cyber Security technology, Vol 6, 2022, Issue:1-2, Pgs 91-111.
[16] Lidong Wang & Randy Jone. “Data Analytics for network intrusion detection". Journal of Cyber Security technology, Vol 4, 2020, Issue-2,Pgs 106-123.
[17] AK. Ghosh, C. Michael, M. Schatz. A Real-Time Intrusion Detection System Based on Learning Program Behavior. Proceedings of the Third International Workshop on Recent Advances in Intrusion Detection, 2000, pp.93-109.
[18] Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., ... & Ramkumar, P. N. (2020). Machine learning and artificial intelligence: definitions, applications, and future directions. Current reviews in musculoskeletal medicine, 13(1), 69-76.
[19] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN Computer Science, 2(3), 1-21.

[20] Sarker, I. H., Furhad, M. H., & Nowrozy, R. (2021). Ai-driven cybersecurity: an overview, security intelligence modeling and research directions. SN Computer Science, 2(3), 1-18.
[21] Truong, T. C., Zelinka, I., Plucar, J., Čandík, M., & Šulc, V. (2020). Artificial intelligence and cybersecurity: Past, presence, and future. In Artificial intelligence and evolutionary computations in engineering systems (pp. 351-363). Springer, Singapore.
[22] Soni, V. D. (2020). Challenges and Solution for Artificial Intelligence in Cybersecurity of the USA. Available at SSRN 3624487.
[23] Al-Dhelaan et al. SVM "Malware Detection using Ensemble of Static Analysis & SVM" (2015).
[24] Khreishah et al. SVM "SVM based Intrusion Detection System" (2009)
[25] Wang et al "Deep Miner: A General Framework for Deep Learning Based Malware Detection" (2019).
[26] Sefiani et al. "Malware Detection using Random Forests and API Calls" (2019).
[27] Liu et al. "Anomaly-based Network Intrusion Detection using Random Forest Algorithm" (2016).
[28] Jang et al. "Malware Detection using k-NN Algorithm with PE Header Information" (2016)
[29] Al-Rawi et al. in their paper "Anomaly-based Intrusion Detection Using k-Nearest Neighbors" (2013).
[30] Dua et al."Malware Detection using Decision Trees" (2017).
[31] Fan et al. in their paper "Anomaly-based Intrusion Detection using Decision Trees" (2012).
[32] Truong, T. C., Diep, Q. B., & Zelinka, I. (2020). Artificial intelligence in the cyber domain: Offense and defense. Symmetry, 12(3), 410.
[33] Shamiulla, A.M. (2019). Role of Artificial Intelligence in Cyber Security. International Journal of Innovative Technology and Exploring Engineering, 9(1), 4628- 4630.
[34] Binny Naik, Ashir Mehta, Hiteshri Yagnik, Manan Shah.“ The impacts of artificial intelligence techniques in augmentation of cybersecurity: a comprehensive review’’, Complex & Intelligent Systems, 2021.
[35] Jin X, Liang J, Tong W, Lu L, Li Z (2017) Multi-agent trustbased intrusion detection scheme for wireless sensor networks. Comput Electr Eng 59:262–273
[36] Achbarou O, El Kiram MA, Bourkoukou O, Elbouanani S (2018) A new distributed intrusion detection system based on multiagent system for cloud environment. Int J Commun Netw Inf Secur 10(3):526