MACHINE LEARNING IN CYBERSECURITY

Abstract

Information science is the driving figure behind the foremost later noteworthy shifts in cybersecurity operations and innovation. Finding designs or contemplations approximately security events in cybersecurity information and developing suitable data-driven models is essential for automating and improving a security architecture. Information science is the consider of real-world occasions utilizing information. It is regularly alluded to as numerous logical strategies, machine learning methods, forms, and frameworks. Since of their extraordinary qualities like adaptability, adaptability, and the capacity to fast alter to novel and obscure challenges, machine learning procedures have been utilized in a wide extend of logical areas. Numerous machine learning algorithms have successfully tackled such a wide variety of computer security challenges. Machine learning algorithms can be attacked during both the training and assessment stages, which often result in notable performance declines and security lapses. In contrast, not much research has been done to comprehend the nature and extent of ML approaches' vulnerabilities against safety risks and associated defence mechanisms. To interest scholars' curiosity, scientists and engineers, it is essential to arrange contemporary cybersecurity-related studies utilising machine learning. Therefore, we provide a comprehensive review of the latest machine learning research in cybersecurity in this article, including the fundamentals of cyberattacks and their accompanying defences, the fundamentals of the most widely used machine learning algorithms, and suggested ML

Keywords: Cybersecurity, machine learning, intrusion detection, spam detection

Authors

Chiman Saini
Department of CSE,
World Collage of Technology
and Management, Farukh Nagar
Gurgoan-122506, Haryana, India.
Chimansaini1994@gmail.com

Poonam Kukana
Department of CSE
UIE, Chandigarh University
Mohali-140413, Punjab.
poonamkukana@gmail.com,

Sukhjinder Kaur
Department of CSE
Rayat Bahra University
Mohali-140103, Punjab.
skaur29100@gmail.com,

Ashima
Department of CSE
Rayat Bahra University
Mohali-140103, Punjab.
ashimamockoul@gmail.com

Bhupinder Kaur
Department of CSE
UIE, Chandigarh University
Mohali-140413, Punjab.
erbhupinderkaur@gmail.com

I. INTRODUCTION

With the introduction of technology ranging from cell phones to extensive communication networks, society has become incredibly digitally linked and internet usage has skyrocketed. Globally, there are currently around 5 billion smart devices and an estimated 3 billion internet users. One This cyber association is broadly utilized for a assortment of purposes, such as online keeping money and shopping, mail, exchanging records or basic data, video conferencing, gaming, and more. Terabytes of information are made, prepared, traded, and put away each moment as a result of the Web of Things (IoT) and other applications. In reality, it is evaluated that the final two a long time alone have delivered 90% of the information within the globe nowadays. The yearly number of cyberattacks has grown as more people use the internet and its related services.

Cyberattacks are the most lethal and damaging weapons, even if they don't utilise any physical weapons. They may reveal sensitive personal information through phishing or the most classified information from the government agencies through espionage. Experts in cybersecurity claim that just in 2017, hackers may have caused damages of up to $5 billion USD, and that damage might increase to US$6 trillion yearly by 2021.

In recent decades, a number of defences against cyberattacks have been developed; they are commonly referred to as intrusion detection systems (IDSs). Computational intelligence methods, such as data mining (DM), (DL), and (ML), have been applied recently to guarantee cybersecurity. Despite the fact that the use of methods from intelligence computation has greatly progressed, improving performance and resilience to cyber attacks. Along with tackling several obstacles, such zero-day attacks, computational intelligence in cyber security still has to make major progress in understanding hostile samples and attacks.

Concern concerning the security and attack-proneness of machine learning algorithms is also developing.

The majority of earlier survey articles often did not include every aspect of machine learning and cybersecurity, including information on commonly used algorithms, cyberattacks, and corresponding defence strategies, as well as information about cybersecurity datasets, problems, and adversarial machine learning. This study varies greatly from the other publications in that ot offers a thorough review of the fundamentals of ML algorithms, cybersecurity research, adversarial ML, datasets, current issues, and future research prospects.

II. TYPES OF CYBER ATTACKS AND THEIR DEFENCE

Cyberattacks are simply methods to compromise computer activity on the network of a victim or to get past security measures to access a victim's computer without authorization. "A cyber-attack is any attack on a computing device that decreases the accessibility, safety, or quality of the data stored within," the Dartmouth college institute for security technology studies states. From a variety of angles, cyber attacks may be divided into several groups according to the impact they have on a system or its design. The essential concepts of cyberattacks and defences are introduced in this section.

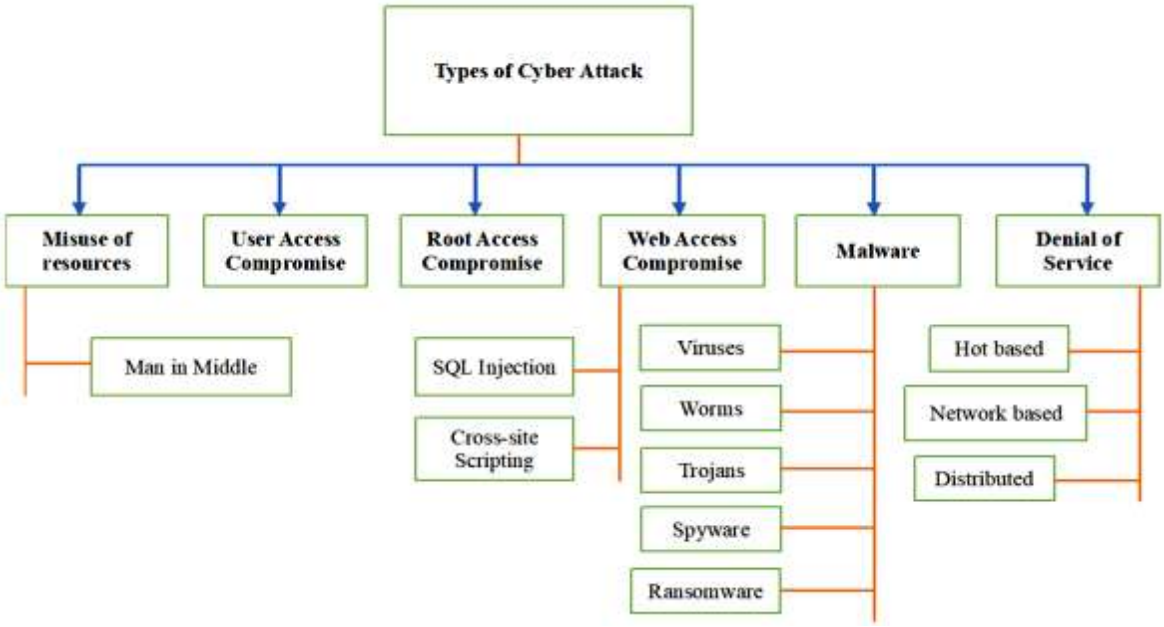

Figure 1: Cyber-attacks Types

1. Misuse of Resource Attack

Employees who are unintentionally careless or overconfident lead to security lapses and provide hackers access to company data. Sometimes well-meaning staff members use workplace network resources to access the virtual private network (VPN) send emails, or wire transfers to others, creating a backdoor for attackers to cause widespread damage to the firm. Even though external attacks make news all the time, internal resource exploitation is still a big problem for companies and organisations all over the world. Employee mismanagement of resources and internal system flaws account for about 25% and 28% of global security breaches, respectively, as to the Ponemon Institute. Organisations were responsible for at least 53% of last year's worldwide security breaches.

a. **Man in Middle Attack:** Figure 2 appears an outline of a Man in Center assault. A programmer can dispatch this assault by interferometer with a trusted client's and server's communication. One common illustration of a MitM assault is session capturing. This sort of assault includes the aggressor taking control of or breaking into a session between the casualty, who is the server's trusted client. By utilizing their claim Web Convention (IP) rather than the victim's, the assailant proceeds to communicate with the server, which considers the attacker’s IP to be a solid client. In the process, the victim's PC is disconnected by the attacker's computer, which also takes over the victim's business information and sequence number.

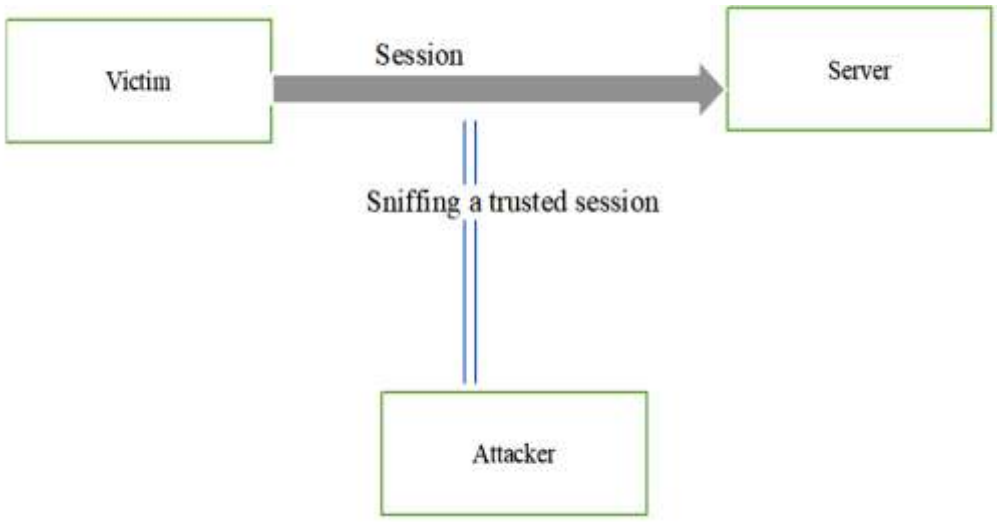

Figure 2: An illustration of a man-in-the-middle assault

b. User Access Compromise: A typical attack type is the compromising of client information, such as a password. Common techniques for getting a user's personal information include brute-force guessing and dictionary attacks, attempts to enter the password database using social engineering, and network connection sniffing to get plaintext passwords. Additional popular techniques for compromising user data are appear- phishing and phishing attacks. Phishing attacks are a technique used to lure people into believing an email that asks for personal information or coerces them into taking certain activities.

In contrast with spear phishing, that is a more focused assault than phishing, the strategies include technical fraud or social engineering, such as attaching a link to a trustworthy website to the mail in order to obtain malware and provide the accessing website personal credentials.

Attackers spend time researching the targets of this kind of assault and crafting believable and intimate communication. Email faking is among the most straightforward methods of spear-phishing attacks. In this case, the attacker poses as one of the victims' management representatives or business associates in order to send them emails. Attackers steal personally identifiable information by cloning trustworthy websites in order to increase their credibility. Figure 3 provides the entire taxonomy of phishing attacks.

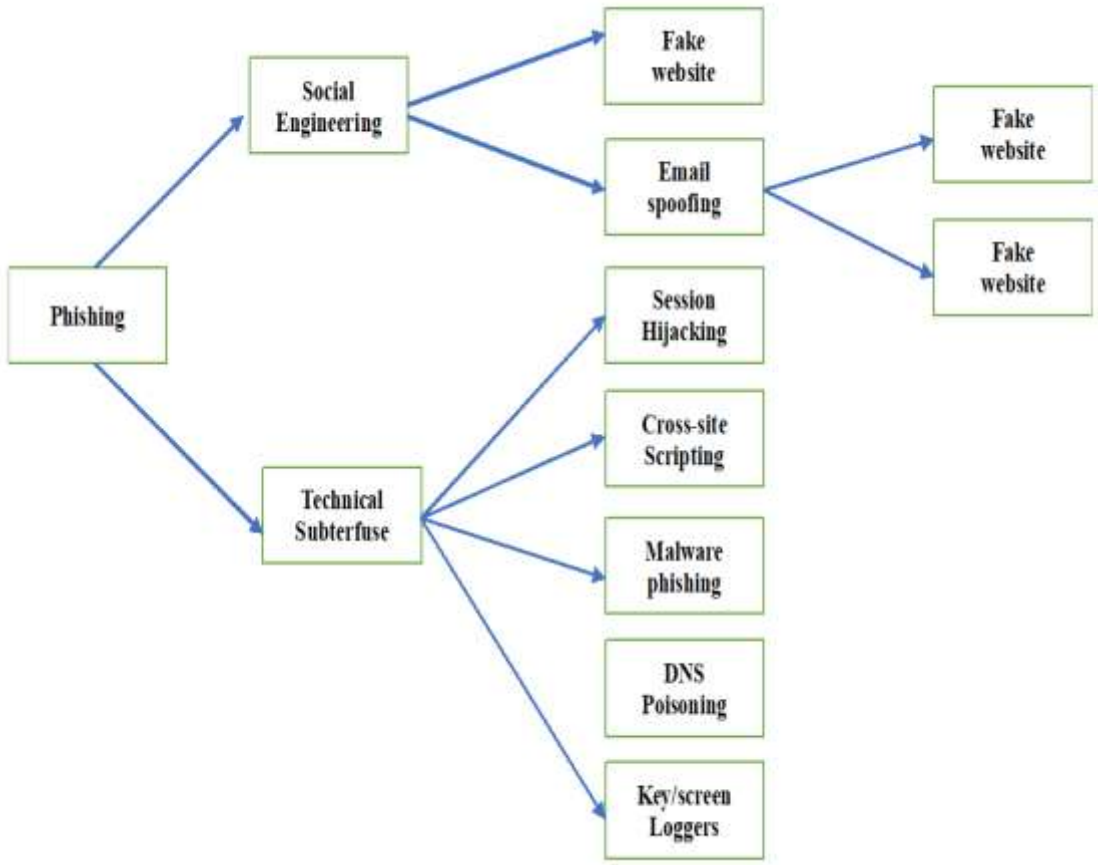

Figure 3: An example of the taxonomy of phishing attacks

c. Root Access Compromise: This assault is distinct from a user compromising attack since attackers get control to the administrators into account, which has greater rights than any other user on the system, instead of only one host.

d. Web Access Compromise: This attack is executed via taking benefit of deficiencies in websites. Two well-known techniques for online compromise are SQL injections and cross-site scripting (XSS).

- **SQL Injection Attack:** This attack occurs on websites that rely on databases, and it involves the attacker entering SQL queries into the database using input data (such login credentials) that is passed from the client to the server. Thus, attackers execute a prepared SQL statement for a post request rather than the anticipated data. In the event that the database lacks read-only access, this command can now and then modify delicate information and/or conduct regulatory assignments in expansion to executing and perusing secret information from the database. To recover related data for an account from the database, for case, a energetic SQL inquiry for an internet site can inquire a client for their account number. "SELECT * FROM users WHERE account = "+ AccountNumber," +";'" F for example Even if an authorised account number is used with this command, it still gives attackers a way in. Such as if an intruder enters "" or "2" = "two" in this way, the resulting SQL query will like this: “SELECT FROM users WHERE account =”2” =”2”,”

Since "2" equals "2," This always returns true, therefore the database will supply the information for every user rather than just one account.

- ***XSS Attack:*** In this kind of attack, the victim's web browser is forced to run or downloading script from the attacker's (third party) websites. Attackers typically add JavaScript code to various websites along with a payload. One of these websites executes the attacker's scripts using the web browser of the victim payload when the victim seeks or requests details from it.

  The attacker's malicious script has the ability to unlock the victim's session cookie and take control of it in order to collect data, including keystroke logs. Additionally, it gives the attacker remote access to the victim's computer.

  Figure 4 provides an outline of XSS.

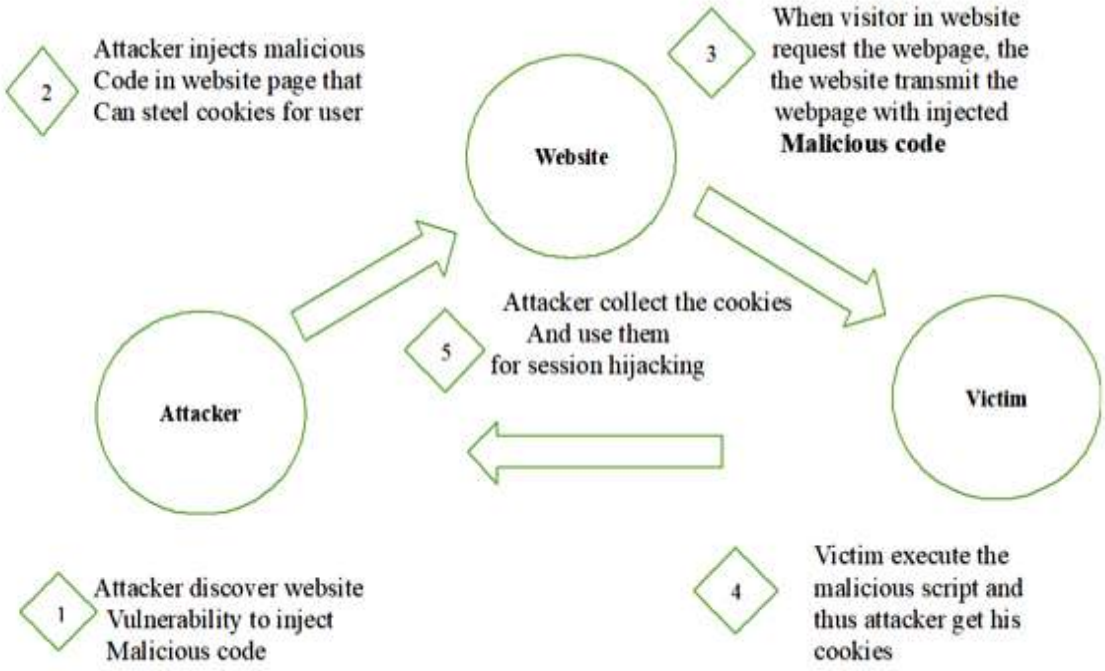

Figure 4: The framework of an attack using cross-site scripting

e. Malware Attacks: Malware an acronym for awful program, is fair an undesirable piece of program. Malware has long been utilized by cybercriminals to attain their objectives, which incorporate deleting or halting a cyber-physical framework, getting huge amounts of private data, tainting a arrange or frameworks, presenting pernicious scripts, and more. Based on the objectives of the gatecrashers and the recurrence of their transmission, malware can be categorized into a few bunches.

- **Viruses:** Malicious malware, viruses, and other host programs can damage files on the host computer and a shared network, much like a live infection in the

human body. Two examples of dangerous viruses are the Melissa and Creeper viruses.

- **Worms:** There are differences between the ways that viruses and worms spread. A host computer is not necessary for worms to propagate, in contrast to viruses. Self-replicating worms often contain email attachments. Additionally, no files on the host computer are harmed by worms. Worms can use the available network resources to replicate themselves to every email contact of the victim in order to carry out a denial-of-service attack." MyDoom, CodeRed, Love Gate, SQL Slammer, and worm are examples of common worms.

- **Trojan:** Based on their intended uses, Trojan horses are entirely distinct from viruses and worms. Attackers use social engineering tactics to trick victims into planting a trojan on their machine. Unlike viruses and worms, trojans provide a backdoor that enables attackers to launch harmful software whenever necessary, rather than replicating themselves or infecting host computer data. Zeus, Dark Comet, and the Shedun android virus are examples of Trojan horses.

- **Spyware*:*** Spyware is used for tracking user activity rather than immediately initiating an assault. Without the user's awareness or agreement, this program is used for stealing private user data, including login passwords and keystroke data.

- **Ransom Ware:** This kind of malware is unique in that its payload not only tampers with the victim's data but also starts a ransom demand procedure. Usually, ransomware assaults are executed via trojans. Torrent Locker, WannaCry, and others are instances of ransomware.

f. Denial of Service: The main objective of such an attack is to completely destroy the regular operation of a system or network. DoS assaults may be divided into three main categories.

- **Host-based:** Malware or worms are installed on the host computers in host-based assaults, where they run their payloads or function to bombard the whole network with an endless stream of host requests, beginning with the host.

- **Network Based:** To carry out their payload, attackers target a whole network rather than a single host computer, hence stopping the network's regular operations.

- **Distributed:** In order to totally take down the victim's network, a Distributed DoS assault is typically launched from both a host computer and a network.

III. BASICS OF MACHINE LEARNING

The collective name for computing techniques that seek to mimic the ways humans learn on computers in order autonomously find and acquire new information is machine learning. It is an interdisciplinary topic of study that includes statistics, psychology, neurology, and computer science. Learning algorithms have advanced significantly in practice as a result of

recent advancements in large data and processing performance. To provide readers with some background information, we include distinct sections for neural network-based techniques and traditional machine learning. Three general categories of machine learning algorithms can be distinguished based on learning approaches: Supervised, unsupervised and reinforcement learning (RL). In order to map the provided real output labels and determine the connection with their corresponding feature value, models are created using supervised learning approaches. Neural networks, decision tree, support vector machine, and other methods are examples of supervised learning techniques. Conversely, unsupervised learning algorithms use the entire training dataset to learn the data and create clusters without being aware of the results of each input. Using training data lacking category labels is how unsupervised learning varies from supervised learning. RL is a trial-and-error approach to learning that uses a particular agent to learn an environment. K-means clustering and k-NN are examples of unsupervised learning algorithms. Training data for reinforcement learning combines supervised and unsupervised techniques. RL investigates behaviours until they are accurate rather than giving data from training with the appropriate label.

Here, we provide a brief synopsis of some well-known machine learning (ML) and cybersecurity techniques. First, we shall discuss the traditional machine learning techniques and their application. We will then examine neural network-based techniques and their associated applications A brief summary of the algorithms' advantages and disadvantages as well as how they work is also included in Table 1.

1. Traditional ML Algorithms

a. Decision Tree: Each vertex (node) in a decision tree represents an attribute, and each branch establishes a maximum value that the attribute may have. DTs are rule-based tree-structured categorisation models. In order to best segment the training data, the root, the highest vertex in a tree, stores the bulk of information gained (variations in entropy) over all features. The lowest nodes are the leaves. Each class is represented by a leaf. To complete the instance that has to be categorised, the DT works top-down during the classification process. A DT uses a tree structure and the information gain equation below to partition instances as effectively as possible:

$$\text{Gain}(P, Q) = \text{Entropy}(P) - \sum_{v \in Dq} \frac{|P_v|}{|P|} \text{Entropy}(P_v) \quad (1)$$

Gain (P, Q) is the term used to describe the decrease in entropy when sorting P based on attribute Q. A top-down method is used to choose nodes from characteristics with greater data gain value. In order to keep the model from becoming over- or under-fitted, researchers54 suggested a number of essential components (pre-pruning, post-pruning, etc.) throughout DT development. In order to categorize or predict fresh events, Tree structure is ultimately transformed into a set of rules.

The DT algorithm's primary benefits are its ease of use and excellent accuracy of classification. The complexity of the computation of the DT classifier is one of its

primary drawbacks. The DT is utilised as a single classifier in security-related applications, including in intrusion detection, it is used as a collaborative classifier.

b. Support Vector Machine: The Support Vector Machine (SVM) is a supervised learning technique that is frequently utilised in cybersecurity, searches the feature space for a dividing hyperplane between its classes. The hyperplane is chosen to be as far away from the closest data point as feasible.

c. The classifier of Naïve Bayes: An approach for probabilistic supervised learning is the Naive Bayes classifier, that given all features as input, provides the likelihood of a class. Bayes' rule serves as the foundation for the Naïve Bayes classifier. Another name for it is the generative model. Given a class, this classifier employs a conditional probability of every characteristic, p(ajb), and the previous probability of all classes, p(b), to derive the following probability of a class, p(bja). Because each characteristic contributes independently to determining a class's posterior probability, the word "naıve" is used:

$$P(b/a) = \frac{P(a.b)}{P(a)} = \frac{P(a/b)P(b)}{P(a)} \quad (2)$$

Where the input vector is denoted by a and the class vector by b. The Naive Bayes classifier's primary benefit is its resilience to noisy training data. Low training samples don't affect performance because the classifier depends on the probabilistic value of every feature. However, this algorithm's primary drawback is that, despite the fact that this seldom occurs in practice, all characteristics are assumed to be independent.

d. k-means Clustering: To find unique clusters in the dataset, k-means clustering, an unsupervised machine learning technique, uses k as a value of cluster groups. Clusters are formed based on the commonalities among all of the data points in the dataset. First, out of m data points, an estimated k number of centroids are found. Equation (3) then assigns m data points, x1, x2, ::, xm to their closest centroids using Euclidean distance measures:

$$\text{Distance} = \sum_{i=1}^{m} d(x_{i,}\, centroid\ (x_{i,})) \quad (3)$$

The centroid that the xi data point corresponds to is denoted here by centroid(xi). In succeeding stages, the centroids are updated using the average distance among each data point supplied to those centroids. These procedures are carried out continuously unless no point of data has the capacity to change any cluster centroids. The objective is to reduce the separation among each centroid as well as the cluster's associated data points. When data labelling becomes challenging, clustering algorithms are mostly employed to identify patterns and clusters in large data settings. One problem with k-means clustering is determining the starting k values. Feature similarity estimates in security applications have made considerable use of K-means clustering.

2. Neural Network-based Algorithms

a. ***Artificial Neural Network:*** Artificial neurone networks. The nodes (perceptron's) that make up artificial neural networks (ANNs) are modelled after brain neurones. An ANN is composed of three layers: the input layer, output layer and the hidden layer. Several hidden layers may exist, contingent on the algorithmic architecture. In a similar manner, each layer transmits its results to the layer below it, which then transmits the outcome. The input layer sends its output to the hidden layer Prior to the development of the SVM in 1990, ANNs were used extensively. In the field of cybersecurity, the ANN has once again become popular with the creation of recurrent, feed-forward, including convolutional neural network.

With an output label of y, the ANN employs (x1, x2, ::, xn) inputs, utilising a weight vector (w1, w2, :: wn), weight the input data during the learning phase. Adjusting the weights during the learning minimizes the learning error, which is E= Pni=1 jdi ~ yij, where the error id is the difference between the neuron's actual output (yi) and its planned output (di). By repeatedly iterating the learning process until the model's error falls below its threshold value, the gradient algorithm known as back-propagation achieves this adjustment. The weight vector can be altered using the formula below:

$$w_{i,j}\ w_{i,j} + \Delta w_{i,j} \tag{4}$$

where j is the hidden node and $\sim w_{i,j} = \eta \delta j x_{i,j}$, i is the input node.

b. ***Conventional Neural Network:*** Large training datasets are mostly handled by DL, a subset of ML methods, which employ hierarchical features abstraction and representation. Traditional machine learning methods perform worse when dealing with large datasets and complex data. DL uses graphics processing units (GPUs) to compute massive volumes of data in order to solve this problem. The most popular deep learning technique in cybersecurity applications is the convolutional neural network (CNN). The CNN's two main layers are the pooling layer and the convolution layer. The input data is convoluted by the convolution layer using several kernels of the same size. If the necessary feature is present at a given point, the convolution process returns a high value for that position, and vice versa, in order to extract features of the input data. For example, the convolution kernel computes every kernel cell value and the associated overlapping image pixels value image data using element-wise multiplication. The convolution and pooling layers are the two primary layers.

Two different pooling techniques—maxpooling and average pooling—are utilised in the pooling layer to down sample the feature sizes. Specifically, average pooling takes the average values from the features computed in the preceding layer, whereas max-pooling selects the largest value. To put it briefly, the pooling mechanism with a kernel output the largest value of the supplied input that is under the kernel at a certain place.

c. Restricted Boltzmann Machines: An improved Boltzmann machine (BM) that lessens the BM's complexity is called a restricted Boltzmann machine (RBM). Stated differently, the RBM speeds up the algorithm's process of learning by limiting the connections between every unit in the same level (visible and buried layers). Reconstruction and training are the two primary goals of the RBM. Energy functions for visible and hidden units of an RBM network with biases (a,b), wij weights between the ith visible layer and the jth hidden layer, and vi visible layer and hj hidden layer might resemble this.

$$E(v,h)=-\sum_{i-1,i\in V}^{n} a_{i,} v_{i,} - \sum_{j-1,j\in V}^{m} b_{j,} h_{j,} - \sum_{i-1}^{n} v_i\, h_j\, w_{i,j} \quad (5)$$

In this case, the binary states vi and hj stand for the jth hidden component in m hidden units and the ith visible part of n visible units, respectively.

d. Deep Belief Network: A generative model with several layers of hidden variables, the deep belief network (DBN) is another well-liked DL approach in cybersecurity. The DBN's architecture makes use of the RBM. The DBN is made up of stacked RBMs that use training data to carry out greedy learning in an unsupervised setting, layer after layer. The RBM has been taught over the previous trained layer of the DBN. Layer by layer, the DBN carries out its training.

Table 1: An overview of the most widely used deep learning and machine techniques in cybersecurity

Method	Working principal	Advantages	Disadvantages
Decision tree	A tree-structures classification model with rules that is developed using the information gained from each feature in the training set	It is simple to implement and has a lower computational cost.	All of the trained model's data must be saved. There is a lot of space complexity.
Support vector machine	Finds a separating hyperplane between its classes in the feature space and maximises the distance between each hyperplane and its nearest data points.	Ideal for big feature dimensions but small sample sizes	The ideal kernel size (k-value) selection
Naïve Bayes classifier	Uses Bayes' method to compute the posterior probability of a class input.	Easy to use, resilient to noisy data for training, and performance unaffected by small sample sizes	Assumes that each feature contributes separately throughout the learningprocedure, however this is rarely the case in reality.
ANN	Consists of a layer or layers across the input and output that are concealed. uses the backpropagation technique to	Ideal for very accurate pattern recognition problems	When compared to other methods, the computational complexity is

	store input data as weights in the hidden layer		substantial.
k-means clustering	Creates groupings or clusters from training data points using similarity metrics.	Simple to put into practice. Ideal for issues where data labelling is extremely challenging	Initial k-value selection necessitates domain expertise.
CNN	CNN's convolution layer uses a number of hidden layers and a pooling layer to generatively extract features from training input and use that knowledge to predict output.	really helpful for pattern recognition and picture categorisation	Computationally challenging. Performance deteriorates when the sample size is small.
Restricted Boltzmann machine (RBM)	With an unsupervised generative learning model, nodes in the same layer—that is, the visible and hidden layers—can only communicate with one another.	Through a feedback mechanism, the RBM is able to extract important characteristics in an unsupervised context.	The cost of computation is really expensive.
Seep belief network (DBN)	To get dependable results, the DBN is composed of layered RBMs that employ greedy layer-by-layer training.	Given that the RBM is applied to each layer of the DBN's training data, it performs better than the RBM.	Considering how many parameters it requires, the computational cost is really significant.

IV. CONCLUSION

Due to the omni-connectivity of digital technology and the pervasiveness of small (like smart watches) to large (like smart metering systems) computing devices, cyber-enabled networks are creating, processing, storing, and exchanging enormous amounts of data, ranging from the general population to private and governmental entities. As a result, protecting data and internet connections has become crucial for both people and whole countries, as well as for small and large businesses. These days, ML has demonstrated significant advancements in cyberspace security by guaranteeing network resilience and data integrity. However, adversaries have also discovered how to use machine learning to skew the effectiveness of cybersecurity measures, such as the malware detection procedure, intrusion detection system, cyber identity detection, etc. In contrast, not many research has been done to examine ML vulnerability problems and the accompanying defence strategies. In response to that demand, we compiled the most current cybersecurity-related research that apply machine learning in one document. By taking into account the fundamentals of the algorithms, the DM strategies employed in these algorithms, and the application, we provided the most widely used machine learning algorithm in cybersecurity in this thorough review. Additionally, efforts on adversarial machine learning have been thoroughly explained, including how resilient DL is to assaults.

REFERENCES

[1] Dasgupta, D., Akhtar, Z., & Sen, S. (2022). Machine learning in cybersecurity: a comprehensive survey. *The Journal of Defense Modeling and Simulation*, *19*(1), 57-106.
[2] Bharadiya, J. (2023). Machine learning in cybersecurity: Techniques and challenges. *European Journal of Technology*, *7*(2), 1-14.
[3] Handa, A., Sharma, A., & Shukla, S. K. (2019). Machine learning in cybersecurity: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(4), e1306.
[4] Al-Garadi MA, Mohamed A, Al-Ali A, et al. A survey ofmachine and deep learning methods for Internet of Things(IoT) security. arXiv.org, vol. arXiv:1807.11023, 2018,pp.1–42.
[5] Thomas J. Individual cyber security: empowering employeesto resist spear phishing to prevent identity theft and ransomwareattacks. Int J Business Manag 2018; 13: 1–24.
[6] Abdulkareem Al-Enezi K, Alshaikhli I, Alkandari A, et al.A survey of intrusion detection system using case studyKuwait Governments entities. In: proceedings of the 20143rd international conference on advanced computer scienceapplications and technologies (ACSAT ’14). IEEE ComputerSociety, USA, Amman, Jordan, 29–30 December 2014,pp.37–43.
[7] Liu L, Vel OD, Han Q, et al. Detecting and preventing cyberinsider threats: a survey. IEEE Commun Surv Tutorial 2018;20: 1397–1417.
[8] Arief B, Adzmi MAB and Gross T. Understanding cybercrimefrom its stakeholders’ perspectives: Part 1–attackers.IEEE Secur Priv 2015; 13: 71–76.
[9] Gupta S, Singhal A and Kapoor A. A literature survey onsocial engineering attacks: Phishing attack. In: 2016 internationalconference on computing, communication and automation (ICCCA), IEEE, Noida, India, 29–30 April 2016, pp.537–540.
[10] Shaikh AN, Shabut AM and Hossain MA. A literaturereview on phishing crime, prevention review and investigationof gaps. In: 2016 10th international conference on software, knowledge, information management & applications (SKIMA), IEEE, Chengdu, India, 15–17 December 2016, pp.9–15.
[11] Gupta BB, Arachchilage NAG and Psannis KE. Defendingagainst phishing attacks: taxonomy of methods, currentissues and future directions. TelecommunSyst 2018; 67:247–267.
[12] Al-Enezi KA, Al-Shaikhli IF, Al-Kandari AR, et al. A surveyof intrusion detection system using case study KuwaitGovernments entities. In: 2014 3rd international conferenceon advanced computer science applications and technologies, IEEE, Amman, Jordan, 29–30 December 2014, pp.37–43.
[13] Ye Y, Li T, Adjeroh D, et al. A survey on malware detectionusing data mining techniques. ACM ComputSurv 2017; 50:41:1–41:40.
[14] Tahir R. A study on malware and malware detection techniques.Int J Educ Manag Eng 2018; 8: 20–30.
[15] Bontupalli V and Taha TM. Comprehensive survey on intrusiondetection on various hardware and software. In: 2015national aerospace and electronics conference (NAECON), IEEE, Dayton, OH, USA, 15–19 June 2015, pp.267–272.
[16] Liu L, Vel OD, Han Q, et al. Detecting and preventing cyberinsider threats: a survey. IEEE Commun Surv Tutorial 2018;20: 1397–1417.
[17] Liu H and Lang B. Machine learning and deep learningmethods for intrusion detection systems: a survey. MachLearn Cybersecur Threat Chall Opportun 2019; 9: 1–48.
[18] 49. Jordan MI and Mitchell TM. Machine learning: trends, perspectives, and prospects. AmAssoc Adv Sci 2015; 349: 255–260.
[19] Qiu J, Wu Q, Ding G, et al. A survey of machine learning forbig data processing. EURASIP J Adv Sign Proc 2016; 2016:67:1–67:16.
[20] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level controlthrough deep reinforcement learning. Nature 2015; 518:529–533.
[21] Alpaydin E. Introduction to machine learning. Cambridge, Massachusetts: MIT Press, 2020.
[22] Kotsiantis SB. Decision trees: a recent overview. ArtifIntellRev 2013; 39: 261–283.
[23] Hastie T, Tibshirani R and Friedman J. The elements of statisticallearning: data mining, inference, and prediction.New York, USA: Springer Science & Business Media, 2009.
[24] Bhuyan MH, Bhattacharyya DK and Kalita JK. Networkanomaly detection: methods, systems and tools. IEEECommun Surv Tutorial 2014; 16: 303–336.
[25] Buczak AL and Guven E. A survey of data mining andmachine learning methods for cyber security intrusion detection.IEEE Commun Surv Tutorial 2016; 18: 1153–1176.
[26] Duda RO, Hart PE and Stork DG. Pattern classification. New York, USA: John Wiley Sons, 2012.
[27] Fischer A and Igel C. Training restricted Boltzmannmachines: an introduction. Pattern Recognit 2014; 47: 25–39.