**Chapter 2**

**Machine Learning in Cyber Security**

Chiman Saini1, Poonam Kukana2, Sukhjinder Kaur3, Ashima4, Bhupinder Kaur5

1Department of CSE, World Collage of Technology and Management,

Farukh Nagar, Gurgoan-122506, Haryana, India

Chimansaini1994@gmail.com

2,5Department of CSE, UIE, Chandigarh University, Mohali-140413, Punjab

poonamkukana@gmail.com, erbhupinderkaur@gmail.com

3,4Department of CSE, Rayat Bahra University, Mohali-140103, Punjab

skaur29100@gmail.com, ashimamockoul@gmail.com

**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 building appropriate data-driven models are the keys to mechanizing and intellectuals upgrading a security framework. 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 quickly 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.Both the training and evaluation stages of machine learning algorithms are susceptible to assaults, 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 attract the interest of academics, scientists, and engineers, it is essential to arrange contemporary cybersecurity-related studies utilising machine learning.Thus, in this article, we present a thorough analysis of the most recent research on machine learning in cybersecurity, 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

**1. 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. As the use of the internet and its related services has grown, so too has the annual number of cyber attacks.

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 of government agencies through espionage. Cybersecurity experts estimate that cyberattacks may have cost US$5 billion in damage in 2017 alone, 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), deep learning (DL), and machine learning (ML), have been applied recently to guarantee cybersecurity. Despite the fact that the use of computational intelligence techniques has greatly progressed, improving performance and resilience to cyber attacks. In addition to overcoming many challenges, such as 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 it offers a thorough review of the fundamentals of ML algorithms, cybersecurity research, adversarial ML, datasets, current issues, and future research prospects.

**2. Types of Cyber attacks and their Defence**

Cyberattacks are simply methods to compromise computer operations on a victim's network or to get past security measures to access a victim's computer without authorization. Any attack on a computer system that compromises the availability, confidentiality, or integrity of the data stored within is considered a cyber attack, according to the Dartmouth College Institute for Security Technology Studies. 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

**2.1 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.Despite the frequent headlines about external assaults, internal resource misuse continues to be a major issue for businesses and organisations worldwide.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.

**2.1.1 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 .The attacker's computer disconnects the victim's PC during that process, and it also assumes the victim's sequence number and corporate information.

****

Figure 2. An example of a Man-in-the-Middle attack

**2.2 User access compromise**

A typical attack type is the compromise of user personal information, such as a password. Common techniques for getting a user's personal information include brute-force guessing and dictionary attacks, social engineering attempts to gain access to the password database, and network connection sniffing to get plaintext passwords. Additional popular techniques for compromising user data are spear-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.

The techniques include social engineering or technical fraud, such as attaching a link to a reliable website to the email in order to download malware and provide the visiting website with personal credentials, in contrast to spear phishing, which is a more targeted attack than phishing.

 Attackers spend time researching the targets of this kind of assault and crafting believable and intimate communications. Email faking is among the most straightforward methods of spear-phishing attacks.In this case, the attackers pose 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

**2.3 Root Access Compromise**

Since the attackers get access to the administrator's account, which has more privileges than other users on the system, rather than only one host, this attack is different from a user compromise attack.

**2.4 Web Access Compromise**

This attack is carried out via taking advantage of website weaknesses. Cross-site scripting (XSS) and structured query language (SQL) injections are two popular methods of internet compromise.

**2.4.1 SQL Injection Attack:** This type of 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 information for an account from the database, for case, a energetic SQL inquiry for an internet site can inquire a client for their account number. For example, "SELECT \* FROM users WHERE account = ''+ AccountNumber +'';.'" Even if this command works well for a valid account number, it still gives attackers a way in. For instance, if an attacker inputs the number "" or "2" = "two" in this manner, 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.

***2.4.2 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 with payload in the victim's web browser 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 overview of XSS.



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

**2.5 Malware attack**

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.

Infections, worms, trojans, spyware, ransomware, scareware, bots, and rootkits are among the common dangers they incorporate.

**2.5.1 Viruses:** Similar to a living infection in the human body, malicious malware, viruses, and other host programs can corrupt files on the host computer and a shared network. The Melissa and Creeper viruses are two instances of harmful viruses.

**2.5.2 Worms:** There are differences between the ways that viruses and worms spread. Worms do not require a host computer to spread, in contrast to viruses. Self-replicating worms often contain email attachments. Additionally, worms do not damage the host computer's files. 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. Common worms include Love Gate, CodeRed, SQL Slammer, MyDoom, and Sorm worm.

**2.5.3 Trojan**: Based on their intended uses, Trojan horses are entirely distinct from viruses and worms. Attackers deceive victims into installing a trojan on their computer by using social engineering techniques. In contrast to viruses and worms, trojans do not replicate themselves or infect host computer data; instead, they establish a backdoor that allows attackers to run malicious software as needed. Zeus, Dark Comet, and the Shedun Android virus are examples of Trojan horses.

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

***2.5.5 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.Typically, trojans are used to carry out ransomware attacks. Torrent Locker, WannaCry, and others are instances of ransomware.

**2.6 Denial of service**

The main objective of this type of cyberattack is to completely destroy the normal functioning status of a system or network. DoS assaults can be divided into three main categories.

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

***2.6.2 Network Based****:* Attackers target an entire network to execute their payload rather than a single host computer, hence stopping the network's regular operations.

**2.6.3 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.

**3. Basics of machine learning**

Machine learning (ML) is the collective term for computational techniques that aim to replicate human learning processes on computers in order to automatically find and pick up new knowledge. 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).
Models are developed utilizing supervised learning techniques to map the given real output labels and ascertain the relationship with their associated feature value. Supervised learning techniques include neural networks, support vector machines (SVMs), decision trees (DTs), and other approaches. 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. Unsupervised learning differs from supervised learning in that it uses training data without category labels. RL is a trial-and-error learning technique that seeks to learn an environment using a specific agent. 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 applications. 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.

**3.1 Traditional ML algorithms**

**3.1.1 Decision Tree:** A DT is a rule-based tree-structured classification model where each vertex (node) in the tree represents an attribute and each branch determines the maximum value that an attribute can have. The majority of information gain (variations in entropy) across all features is stored in the root, the topmost vertex in a tree, which is utilized to best partition the training data. The lowest nodes are the leaves. A leaf is used to symbolize each class. During the classification process, the DT moves top-down to fulfil the instance that needs to be classified. A DT uses a tree structure and the information gain equation below to partition instances as effectively as possible:

 Gain (P,Q) =Entropy(P)-$\sum\_{v∊Dq}^{}|P\_{v}|$ Entropy ($P\_{v}$) (1)

 |P|

In this case, the reduction in entropy to sort P according to attribute Q is called Gain(P,Q). Using a top-down approach, features with increasing information gain value are selected as nodes. 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 forecast new occurrences, the tree structure is ultimately transformed into a set of rules.

The DT algorithm's primary benefits are its ease of use and excellent classification accuracy. The computational complexity of the DT classifier is one of its primary drawbacks. The DT is employed as a collaborative classifier in intrusion detection in addition to being a single classifier in security-related applications.

**3.1.2 Support Vector Machine**

The Support Vector Machine (SVM), one of the most popular supervised learning algorithms 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.

**3.1.3 The classifier of Naïve Bayes**

The Naive Bayes classifier is a probabilistic supervised learning method that, given all features as input, provides the likelihood of a class. The Naive Bayes classifier is built on top of Bayes' rule. It is also known as the generative model. The posterior probability of a class p(bja) is determined by this classifier using the conditional probability of all characteristics given a class, or p(ajb), and the prior probability of all classes, p(b). Because each characteristic contributes independently to determining a class's posterior probability, the word "naıve" is used:

 P(b|a) = P(a,b) = P(a|b)P(b) (2)

 P(a) P(a)

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.

**3.1.4 K-means clustering**

To find unique clusters in the dataset, k-means clustering, an unsupervised machine learning technique, uses k as the 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:

 Distance = $\sum\_{i=1}^{m}d(x\_{i, }centroid (x\_{i, }))$ (3)

Here, centroid(xi) represents the centroid to which the xi data point belongs. In later phases, the average distance between each data point provided to those centroids is used to update the centroids. These steps are performed throughout the process until no data point has the ability to alter any cluster centroids. The goal is to minimize the distance between each centroid and the related data points inside a cluster. Clustering techniques are mostly used to find data patterns or clusters in huge data contexts when data labeling becomes difficult. An issue with k-means clustering is figuring out the initial k value. In security applications that use feature similarity calculations, K-means clustering has been used extensively.

**3.2 Neural network-based algorithms**

***3.2.1 Artificial neural network:***networks of artificial neurones. The nodes (perceptrons) that make up artificial neural networks (ANNs) are modelled after brain neurones. The input layer, hidden layer, and output layer are the three layers that make up an ANN. There could be multiple hidden layers, depending on the algorithmic design. Likewise, each layer sends its output to the one behind it, and the output layer finally sends the result. The hidden layer receives the output from the input layer. ANNs were widely utilized before the SVM was developed in 1990. With the development of recurrent, feed-forward, and convolution neural networks, the ANN has once again gained popularity in the cybersecurity sector.

With an output label of y, the ANN employs inputs (x1, x2, :::, xn). A weight vector (w1, w2, ::, wn) is used to weight the input data during the learning phase. Adjusting the weights during the learning process minimizes the learning error, which is E= Pni=1 jdi ~ yij, where the error is the difference between the neurone'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:

 wi, j wi, j +Δwi, j (4)

 where j is the hidden node and ~wi, j = ηδjxi, j, i is the input node.

***3.2.2 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 each kernel cell value and the associated overlapping image pixel value for image data using element-wise multiplication. The two main layers are the pooling layer and the convolution layer.

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 within a kernel outputs the largest value of the supplied input that is under the kernel at a certain place.

**3.2.3 Restricted Boltzmann machines**

The restricted Boltzmann machine (RBM) is an enhanced version of the Boltzmann machine (BM) that reduces the complexity of the BM. In other words, by restricting the connections between all units in the same layer (visible and buried layers), the RBM accelerates the algorithm's learning process. Training and reconstruction are the two main objectives of the RBM. An RBM network with vi visible layer, hj hidden layer, wij weights between the ith visibility layer and the jth hidden layer, and biases (a,b) may have the following energy functions for its visible and hidden units.

 E(v,h)=-$\sum\_{i-1,i∊V}^{n}a\_{i, }v\_{i, }$ - $\sum\_{j-1,j∊V}^{m}b\_{j, }h\_{j, }$ - $\sum\_{i-1}^{n}v\_{i }h\_{j }w\_{i,j }$ (5)

In this case, the jth hidden unit of m hidden units and the ith visible unit of n visible units are represented by the binary states vi and hj.

**3.2.4 Deep belief network**

Another popular DL technique in cybersecurity is the deep belief network (DBN), a generative model with multiple layers of hidden variables. The RBM is used in the architecture of the DBN. The DBN is made up of stacked RBMs that use training data to carry out layer-by-layer greedy learning in an unsupervised learning environment. 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 machine learning and deep learning techniques in cybersecurity.

|  |  |  |  |
| --- | --- | --- | --- |
| **Method** | **Working Principal**  |  **Advantages**  | **Disadvantages** |
| Decision tree | A tree-structured 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 | seeks to identify a separating hyperplane among its classes in the space of features in order to maximise the distance between the hyperplane and its closest data points. | Ideal for big feature dimensions but small sample sizes | The ideal kernel size (k-value) selection is |
| 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 store input data as weights in the hidden layer. | Ideal for very accurate pattern recognition problems | When compared to other methods, the computational complexity is 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. |
| RestrictedBoltzmann 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. | The RBM can extract significant features in an unsupervised learning environment thanks to a feedback mechanism. | The cost of computation is really expensive. |
| Deep beliefnetwork (DBN) | The DBN is made up of stacked RBMs that use greedy layer-by-layer training to get reliable results. | Since the DBN uses the RBM on top of every layer of its training data, it performs better than the RBM. | Considering how many parameters it requires, the computational cost is really significant. |

**4. Conclusion**

Because of digital omni-connectivity 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 public to classified by individuals and government organizations. 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 (ML) 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 applications, we provided the most widely used machine learning algorithms 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.

3. Bharadiya, J. (2023). Machine learning in cybersecurity: Techniques and challenges. *European Journal of Technology*, *7*(2), 1-14.

4. 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.

5. 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.

6. Thomas J. Individual cyber security: empowering employeesto resist spear phishing to prevent identity theft and ransomwareattacks. Int J Business Manag 2018; 13: 1–24.

7. 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.

8. Liu L, Vel OD, Han Q, et al. Detecting and preventing cyberinsider threats: a survey. IEEE Commun Surv Tutorial 2018;20: 1397–1417.

9. Arief B, Adzmi MAB and Gross T. Understanding cybercrimefrom its stakeholders’ perspectives: Part 1–attackers.IEEE Secur Priv 2015; 13: 71–76.

10. 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.

11. 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.

12. Gupta BB, Arachchilage NAG and Psannis KE. Defendingagainst phishing attacks: taxonomy of methods, currentissues and future directions. TelecommunSyst 2018; 67:247–267.

13.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.

14. Ye Y, Li T, Adjeroh D, et al. A survey on malware detectionusing data mining techniques. ACM ComputSurv 2017; 50:41:1–41:40.

15. Tahir R. A study on malware and malware detection techniques.Int J Educ Manag Eng 2018; 8: 20–30.

16. 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.

17. Liu L, Vel OD, Han Q, et al. Detecting and preventing cyberinsider threats: a survey. IEEE Commun Surv Tutorial 2018;20: 1397–1417.

18. 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.

19. 49. Jordan MI and Mitchell TM. Machine learning: trends, perspectives,and prospects. AmAssoc Adv Sci 2015; 349: 255–260.

20. 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.

21. Mnih V, Kavukcuoglu K, Silver D, et al. Human-level controlthrough deep reinforcement learning. Nature 2015; 518:529–533.

22. Alpaydin E. Introduction to machine learning. Cambridge,Massachusetts: MIT Press, 2020.

23. Kotsiantis SB. Decision trees: a recent overview. ArtifIntellRev 2013; 39: 261–283.

24. Hastie T, Tibshirani R and Friedman J. The elements of statisticallearning: data mining, inference, and prediction.New York, USA: Springer Science & Business Media, 2009.

25. Bhuyan MH, Bhattacharyya DK and Kalita JK. Networkanomaly detection: methods, systems and tools. IEEECommun Surv Tutorial 2014; 16: 303–336.

26. 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.

27. Duda RO, Hart PE and Stork DG. Pattern classification.New York, USA: John Wiley Sons, 2012.

28. Fischer A and Igel C. Training restricted Boltzmannmachines: an introduction. Pattern Recognit 2014; 47: 25–39.