energies

MDPI

Review

Performance Comparison and Current Challenges of Using Machine Learning Techniques in Cybersecurity

Kamran Shaukat [1,2,*], Suhuai Luo [1], Vijay Varadharajan [1], Ibrahim A. Hameed [3,*], Shan Chen [1], Dongxi Liu [4] and Jiaming Li [4]

[1] School of Electrical Engineering and Computing, The University of Newcastle, Newcastle 2308, Australia; suhuai.luo@newcastle.edu.au (S.L.); vijay.varadharajan@newcastle.edu.au (V.V.); Shan.Chen@newcastle.edu.au (S.C.)
[2] Punjab University College of Information Technology, University of the Punjab, Lahore 54590, Pakistan
[3] Department of ICT and Natural Sciences, Norwegian University of Science and Technology, 7491 Trondheim, Norway
[4] Data61, Commonwealth Scientific and Industrial Research Organization, Canberra 3169, Australia; dongxi.liu@data61.csiro.au (D.L.); Jiaming.Li@CSIRO.au (J.L.)
* Correspondence: Kamran.shaukat@uon.edu.au (K.S.); ibib@ntnu.no (I.A.H.); Tel.: +61-401-754746 (K.S.)

Received: 3 April 2020; Accepted: 11 May 2020; Published: 15 May 2020

Abstract: Cyberspace has become an indispensable factor for all areas of the modern world. The world is becoming more and more dependent on the internet for everyday living. The increasing dependency on the internet has also widened the risks of malicious threats. On account of growing cybersecurity risks, cybersecurity has become the most pivotal element in the cyber world to battle against all cyber threats, attacks, and frauds. The expanding cyberspace is highly exposed to the intensifying possibility of being attacked by interminable cyber threats. The objective of this survey is to bestow a brief review of different machine learning (ML) techniques to get to the bottom of all the developments made in detection methods for potential cybersecurity risks. These cybersecurity risk detection methods mainly comprise of fraud detection, intrusion detection, spam detection, and malware detection. In this review paper, we build upon the existing literature of applications of ML models in cybersecurity and provide a comprehensive review of ML techniques in cybersecurity. To the best of our knowledge, we have made the first attempt to give a comparison of the time complexity of commonly used ML models in cybersecurity. We have comprehensively compared each classifier's performance based on frequently used datasets and sub-domains of cyber threats. This work also provides a brief introduction of machine learning models besides commonly used security datasets. Despite having all the primary precedence, cybersecurity has its constraints compromises, and challenges. This work also expounds on the enormous current challenges and limitations faced during the application of machine learning techniques in cybersecurity.

Keywords: cybersecurity; machine learning; malware detection; intrusion detection system; spam classification

1. Introduction

In this age, the cyberspace is growing faster as a primary source for a node to node information transfer with all its charms and challenges. The cyberspace serves as a significant source to access an infinite amount of information and resources over the globe. In 2017, the internet usage rate was 48% globally, later it increased to 81% for developing countries [1]. The broad spectrum of the cyberspace embraces the internet, users, the system resources, the technical skills of the participants and much more, not just the internet. The cyber-world also plays a significant role in causing limitless vulnerabilities to cyber threats and attacks. Cybersecurity is a set of different techniques, devices, and methods used to defend cyberspace against cyber-attacks and cyber threats [2]. In the modern world of computer and information technology, the cybercrimes are growing with faster steps as compared to the current cybersecurity system. The weak system configuration, unskilled staff, and scanty amount of techniques are some factors that rise to vulnerabilities in a computer system to threats [3]. Because of the growing cyber threats, more headway needs to make when developing cybersecurity methods. The outdated and conventional cybersecurity methods have a substantial downside because these methods are ineffectual in dealing with unknown and polymorphic security attacks. There is a need for robust and advanced security methods that can learn from their experiences and detect the previous and new unknown attacks. Cyber threats are increasing in a significant way. It is becoming very challenging to cope with the speed of security threats and provide needful solutions to prevent them [4].

Machine learning: One of the primarily used advanced methods for cybercrime detection is machine learning techniques. Machine learning techniques can be applied to address the limitations and constraints faced by conventional detection methods [5]. Researchers have addressed the advancements, limitations, and constraints of applying machine learning techniques for cyberattack detection and have provided a comparison of conventional methods with machine learning techniques. Machine learning is a sub-field of artificial intelligence. ML techniques are built with the abilities to learn from experiences and data without being programmed explicitly [6]. Applications of ML techniques are expanding in different areas of life, such as education [7,8], medical [9–11], business and cybersecurity [12–14]. Machine learning techniques are playing their role on both sides of the net, i.e., attacker-side and defender-side. On the attacker side, ML techniques are employed to pass through the defense wall. In contrast, on defense side, ML techniques are applied to create prompt and robust defense strategies.

Cyber threats: Machine learning techniques are playing a vital role in fighting against cybersecurity threats and attacks such as intrusion detection system [15,16], malware detection [17], phishing detection [18,19], spam detection [20,21], and fraud detection [22] to name a few. We will focus on malware detection, intrusion detection system, and spam classification for this review. Malware is a set of instructions that are designed for malicious intent to disrupt the normal flow of computer activities. Malicious code runs on a targeted machine with the intent to harm and compromise the integrity, confidentiality and availability of computer resources and services [23]. Saad et al. in [24] discussed the main critical problems in applying machine learning techniques for malware detection. Saad et al. argued that machine learning techniques have the ability to detect polymorphic and new attacks. Machine learning techniques will lead to all other conventional detection methods in the future. The training methods for malware detections should be cost-effective. The malware analysts should also be able to keep with the understanding of ML malware detection methods up to an expert level. Ambalavanan et al. in [25] described some of the strategies to detect cyber threats efficiently. One of the critical downsides of the security system is that the security reliability level of the computing resources is generally determined by the ordinary user, who does not possess technical knowledge about security.

Another threat to computer resources is a spam message. Spam messages are unwanted and solicited messages that consume a lot of network resources along with computer memory and speed. ML techniques are being employed to detect and classify a message as spam or ham. ML techniques have a significant contribution to detect spam messages on computer [26,27], SMS messages on mobile [28], spam tweets [29], or images/video [30,31].

An intrusion detection system (IDS) is a protection system to computer networks from any malign intrusions for scanning the network vulnerabilities. Signature-based, anomaly-based and hybrid-based are considered major classifications of an intrusion detection system for network analysis. ML techniques have a substantial contribution to detecting different types of intrusions on network and host computers. However, there are numerous areas such as detection of zero-day and new attacks are considered significant challenges for ML techniques [32].

Threats to validity: For this review, we included the studies that (1) deal with anyone of the six machine learning models in cyber security, (2) target cyber threats including intrusion detection, spam detection, and malware detection, and (3) discuss the performance evaluation in terms of accuracy, recall, or precision. We have used multiple combinations of strings such as 'machine learning and cyber security' and 'machine learning and cybersecurity' to retrieve the peer-reviewed articles of journal, conference proceedings, book chapters, and reports. We have targeted the six databases, namely Scopus, ACM Digital Library, IEEE Xplore, ScienceDirect, SpringerLink, and Web of Science. Google Scholar was also used for forward and backward search. We have focused on recent advancement in the last ten years. In total, 2852 documents were retrieved, and 1764 duplicated items were removed. The title and abstract were screened to identify potential articles. The full text of 361 studies were assessed to find the relevancy with the inclusion criteria. We have excluded the articles that were discussing (1) the cyber threats other than intrusion detection, spam detection, and malware detection, (2) threats to cyber-physical systems, (3) threats to cloud security, IoT devices, smart grids, smart cities, and satellite and wireless communication. With backward and forward search, 19 more studies were retrieved. In total, 143 studies were finally selected for data extraction purpose. Figure 1 depicts the process of article selection. The previous survey and review articles were used in addition to these included papers to provide a comprehensive performance evaluation.

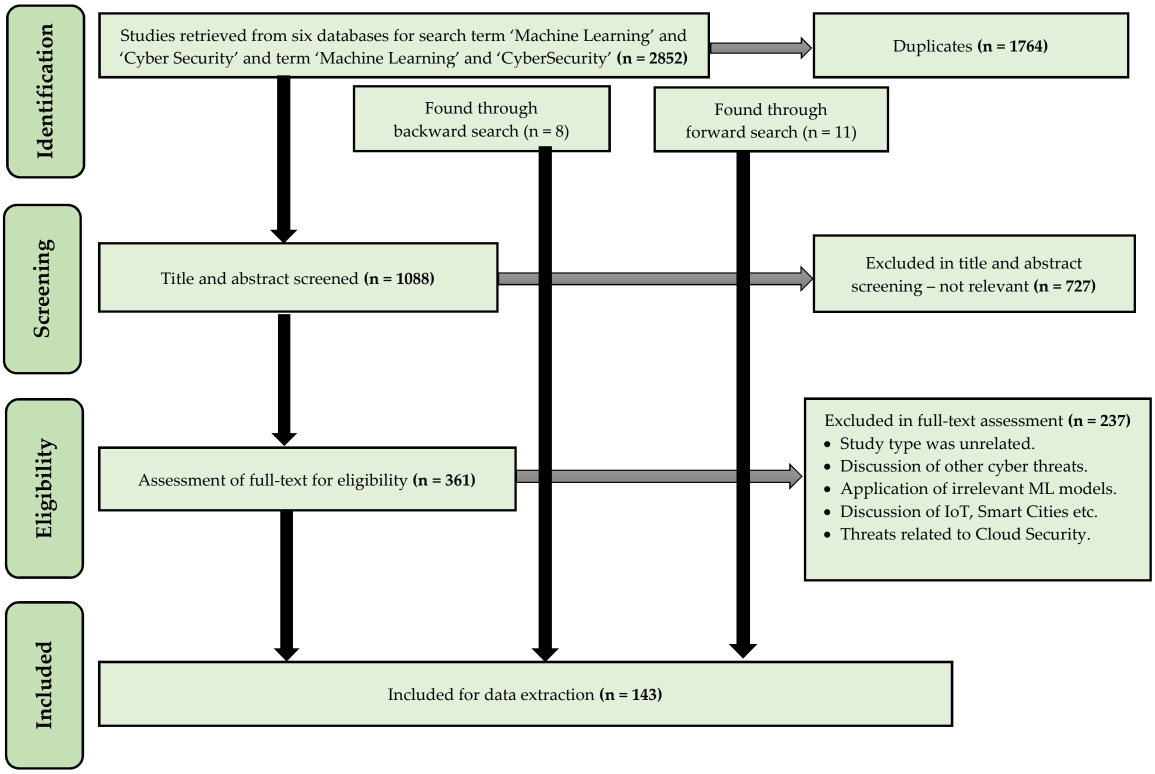
Identification
Screening
Eligibility
Included
Studies retrieved from six databases for search term 'Machine Learning' and 'Cyber Security' and term 'Machine Learning' and 'CyberSecurity' (n = 2852)
Duplicates (n = 1764)
Found through backward search (n = 8)
Found through forward search (n = 11)
Title and abstract screened (n = 1088)
Excluded in title and abstract screening – not relevant (n = 727)
Assessment of full-text for eligibility (n = 361)
Excluded in full-text assessment (n = 237)
• Study type was unrelated.
• Discussion of other cyber threats.
• Application of irrelevant ML models.
• Discussion of IoT, Smart Cities etc.
• Threats related to Cloud Security.
Included for data extraction (n = 143)

Figure 1. Flow Diagram Showing Article Selection Process for This Review.

Xin et al. in [33] reviewed the critical challenges faced by machine learning techniques and their solutions in a network intrusion detection system. Each ML technique has its pros and cons. No ML technique could be declared as the best technique with no limitations. One of the biggest challenges faced by ML techniques is that data collection is a lengthy and laborious procedure. Most of the publicly available datasets are outdated, have missing or redundant values [33]. In contrast, this paper covers other cybersecurity threats and the evaluation of ML models in those areas.

Gandotra et al. in [34] provided a classification of malware in the static, dynamic and hybrid analysis. Moreover, he provided a review of several research papers that applied machine learning techniques to detect malware. However, he only targeted a cyber threat, i.e., malware. Moreover, critical analysis and performance evaluation of machine learning techniques are missed. There is no description of the state-of-the-art malware datasets. In contrast, our paper has targeted sever cyber threats and provided the description of commonly used datasets. Moreover, the performance evaluation of significant machine learning techniques on a frequently used dataset is also presented.

Bhavna et al. in [35] reviewed several papers applying machine learning techniques to detect cyber threats. However, they have focused and described more on instruction detection. Performance evaluation of machine learning techniques and benchmark datasets are also not provided.

Ford et al. in [36] presented a survey on the application of machine learning techniques in cybersecurity. This survey addressed the crucial challenges in applying machine learning technique in cybersecurity. ML techniques are efficiently fighting against the cyberattacks and threats. However, machine learning classifiers themselves are exposed to various cyber and adversarial attacks. There is an immense amount of work needed to improve the safety of ML from adversarial cyberattacks. Jiang et al. in [37] examined the various publications on using machine learning techniques in cybersecurity from 2008 to early 2016. The authors also described that, despite the growing role of machine learning techniques in cybersecurity, the selection of appropriate and suitable machine learning technique for a specific underlying safety problem is still a challenging matter of grave concern.

Hodo et al. in [38] assessed the performance of machine learning techniques in anomaly detection and measured the usefulness of feature selection in ML IDS. They claimed that although convolutional neural network (CNN) classifier could have served as a satisfactory classifier in cybersecurity, still it has not been used to its full potential. Moreover, machine learning models are unable to adequately detect the attacks because of the missing and incorrect signatures in the signature list of an intrusion detection system. Besides, further work is needed to explore the knowledge-based and behavioral-based approaches.

Apruzzese et al. in [39] presented an analysis of machine learning techniques in cybersecurity to detect the spams, malware and intrusions. It asserted that the machine learning techniques are vulnerable to cyber threats and all the methods are still struggling to overcome all the limitations and obstacles. The biggest challenge is that most same classifier used for different kind of safety problems. It is highly required to find suitable classifier for a particular safety issue. It also emphasized that all the shortcomings of machine learning techniques should be handled as a matter of deep concern as cyber attackers are leveraging all their resources.

The communication technologies used by smart grids are leading to cybersecurity deficits. Yin et al. in [40] developed a method to gauge the vulnerable area of distributed network protocol 3 (DNP3) protocol based on IoT-based smart grid and SCADA system. The obtained vulnerability measures were used to develop an attack model for the data link layer, transport layer, and application layer. Furthermore, they developed two algorithms by applying machine learning techniques to transform the data. Authors showed by experimental results that the proposed system classified intrusive fields with detailed information about DNP3 protocol. Peter et al. in [41] discussed three types of malware and central measures that are crucially needed to overcome the security threats. They suggested that cybercrimes can be reduced by continuously updating the cybersecurity policy, decreasing the reaction time and robust segmentation. Ndibanje et al. in [42] presented a classification method for obscure malware detection by using API call as malicious code. They applied similarity-based machine algorithms for feature extraction and claimed to have effective results for obscure detection methods.

Torres et al. in [43] discussed the utilization of machine learning classification techniques applied in cybersecurity. They provided a review of different alternatives to using machine learning models to reduce the error rate in intrusion and attack detection. However, this paper describes the significant challenges and considers several other cyber threats to cybersecurity. Ucci et al. [44] focused on achieving malware detection using key machine learning techniques. They analyzed malware detection using a feature extraction process. They also emphasized that there is an urgent need to update the currently used datasets as most of the publicly available datasets are outdated. In contrast, this paper provided an overview of commonly used ML models, their complexities and evaluations criteria based on several datasets in multiple cyber domains. Table 1 presents a comparison of this paper with the existing survey and review papers. It can be observed that most of the review papers have not presented a comprehensive review of significant cyber threats. Moreover, none of the paper provided the performance evaluation of famous machine learning techniques. Secondly, we have not just provided the performance evaluation. Instead, we have compared them based on benchmark datasets. Our comparisons in Tables 4–9 depict the performance of each machine learning technique on the detection of significant cyber threats based on frequently used datasets. We have also described the current challenges of using machine learning techniques in cybersecurity that open new horizons for future research in this direction.

Contributions: In this review paper, we build upon the existing literature of applications of ML models in cybersecurity and provide a comprehensive review of ML techniques in cybersecurity. The following are significant contributions to this study:

(1) To the best of our knowledge, we have made the first attempt to provide a comparison of the time complexity of commonly used ML models in cybersecurity. We have also described the critical limitations of each ML model.
(2) Unlike other review papers, we have reviewed applications of ML models to common cyber threats that are intrusion detection, spam detection and malware detection.
(3) We have comprehensively compared each classifier performance based on frequently used datasets.
(4) We have listed the critical challenges of using machine learning techniques in the cybersecurity domain.

This review paper is organized as follows: Section 2 describes an overview of cybersecurity threats, commonly used security datasets, basics of machine learning, and evaluation criteria to evaluate the performance of any classifier. Section 3 provides a comprehensive comparison of frequently used ML classifiers based on different cyber threats and datasets. Section 4 concludes this study and points out the critical challenges of ML models in cybersecurity.

Table 1. Comparison of Existing Review Papers (Legend: √ means covered; ≈ means partially covered; × means not covered.).

Sr#	Cite	Year	No of References	Cybersecurity				Machine Learning			Performance Evaluation	Challenges of ML in Cybersecurity
				Spam Detection	Malware Detection	IDS	Datasets	Techniques	Metrics	Time Complexity		
1	[34]	2014	51	×	√	×	×	×	×	×	×	×
2	[35]	2014	18	×	×	√	≈	√	≈	×	×	×
3	[36]	2014	24	≈	×	√	×	×	×	×	×	×
4	[45]	2015	113	×	×	√	√	√	√	×	×	×
5	[37]	2016	164	√	√	√	×	×	×	×	×	×
6	[46]	2017	21	×	×	√	≈	×	√	×	×	×
7	[38]	2017	154	×	×	√	√	√	√	×	×	×
8	[44]	2017	107	×	√	×	√	×	×	×	×	≈
9	[47]	2018	68	×	√	√	×	√	×	×	×	√
10	[48]	2018	107	×	√	×	≈	×	×	×	×	×
11	[39]	2018	40	≈	≈	√	×	√	≈	×	×	×
12	[49]	2018	14	≈	≈	≈	×	√	×	×	×	×
13	[50]	2019	12	×	≈	≈	×	×	×	×	×	×
14	[51]	2019	45	×	×	√	√	√	×	×	×	×
15	[52]	2019	174	≈	×	√	√	√	√	×	×	×
16	[43]	2019	200	√	√	×	×	√	√	×	×	×
17	[24]	2019	38	×	√	×	×	×	×	×	×	×
18	[25]	2020	14	≈	×	√	×	≈	×	×	×	×
19	**Our Paper**	**2020**	**170**	**√**	**√**	**√**	**√**	**√**	**√**	**√**	**√**	**√**

2. Cybersecurity and Machine Learning

This section is divided into four parts. First part provides the basics of cyber threats and attacks. The second part describes the commonly used security datasets for computer networks and mobile. The third part presents the fundamentals of machine learning and various machine learning algorithms. The fourth section describes different metrics to evaluate a classifier.

2.1. Basics of Attacks and Threats

The malicious attacking technologies are gaining faster progression than the defending techniques. Cybersecurity aims to maintain data protection, resource protection, data privacy, and data integrity [53,54]. There are various threats and attacks on cyberspace. Common threats to cyberspace are fraud detection, malware detection, spam classification, phishing, disabling firewall and antivirus, logging of keystrokes, malicious URL, and probing to name a few.

Phishing and malware are considered as critical threats to cyberspace. Phishing is the method to get unauthorized access to the data by pretending as a legitimate user. Sending a link of a web page posing as a legitimate page that navigates to other links to enter personal information is an example of phishing. In contrast, malware is malicious software that is developed intentionally to get unauthorized access on the target computer and disrupt the normal flow of activities [55]. Malware detection has three sub-classes, namely static, dynamic, and hybrid detections. In the case of static malware detection, the applications are examined for the malicious pattern without executing them. However, dynamic detection is performed while the applications are running. Hybrid detection is a mixture of both detection techniques. Virus, Trojan horse and worms are sub-categories of malware. A virus is a piece of malicious code that destroys the data on the system by unwitting the user. A worm is malicious software that illegally consumes the system resources by replicating itself. A trojan horse obtains unauthorized access to the data by professing itself as legitimate software. A Trojan horse does not replicate itself [56,57].

A spam message via email or SMS is another critical threat to the computer and network resources. Spam messages consume a lot of computer memory and network resources. Spam message affects both mobile and computer networks. Spam can be found in the form of email, images, videos, tweets, and spam blogs on mobile and computer networks.

Several defense mechanisms have been installed on network systems to detect unauthorized intrusion and probing. Cybercriminals can scan computer networks for vulnerabilities. There are three categories of intrusion detection based on network analysis such as signature-based, anomaly-based and hybrid-based. Signature-based techniques are used to detect the known attacks, whereas anomaly-based detection detects any unusual behavior within the network. Hybrid-based detection is a combination of both detection techniques. There are four categories of cyber-attacks, namely user to root (U2R), remote to local (R2L), probing, and denial of service (DOS). If a user tries to get access rights of a root/admin user, then this attack is called U2R. In contrast, if a remote user tries to gain access as a local user, then the attack is classified as R2L. Whereas, if a legitimate user is denied to the system access by making the network resources busy, then the phenomena is called DOS. However, in the case of probing, cybercriminals only scan the network to find weak areas for future attacks.

2.2. Commonly Used Security Datasets

Machine learning techniques produce better results if the datasets have diversity and collected real-time data. In this sub-section, we will discuss the most used security datasets.

Frequently used security datasets are the Defense Advanced Research Project Agency (DARPA) datasets, URL dataset, KDD Cup 99 dataset, Australian Defense Force Academy (ADFA) dataset, HTTP CSIC-2010, Android malware dataset, Android validation dataset, Spambase, and NSL-KDD. The primary outcome of the DARPA dataset is the detection of the attacks [58].

DARPA dataset is a network traffic and audit logs-based dataset. It has its limitations to handle new system variations. DARPA does not show real-world network traffic of data [59]. AFDA dataset was developed to get the better of the DARPA dataset. AFDA overcame the limitations to handle new system variations [60]. KDD Cup 99 dataset was formed using a subset of DARPA dataset. The later advancement for KDD Cup 99 dataset is NSL-KDD dataset [61]. NSL-KDD was proposed to overcome data redundancy and duplicate records. Thus, the NSL-KDD dataset has a reasonable number of records as compared to KDD Cup 99, and it performs better than KDD Cup 99 [61,62]. The primary purpose of the HTTP CSIC-2010 dataset is to detect web attacks. HTTP CSIC-2010 dataset handles a massive number of web queries. HTTP CSIC-2010 dataset is known as the most long-established and efficient dataset for attack detection in web queries [63]. The Enron dataset is composed of a massive number of emails produced by the Enron Corporation's staff. This dataset used to classify the spam emails. Spambase dataset is another commonly used dataset to ascertain and refine the spam emails. This dataset computes the different attributes of the collected observations and publicly available on the UCI ML repository [64]. The URL dataset, an internet traffic-based dataset, was proposed to blacklist malicious URLs [65]. The URL dataset consists of five different types of malicious URLs: phishing URLs, spam URLs, malware URLs, benign URLs, and defacement URLs. The Android malware dataset is an android apps-based dataset. Android malware dataset was proposed to blacklist malware android applications [66]. The Android validation dataset was generated to find various relations between 72 real apps by extracting two types of features: metadata and N-grams. The Android validation dataset shows that there are different relationships between apps, for example, siblings, false siblings, step-siblings and cousins [67].

2.3. Basics of Machine Learning

Artificial intelligence (AI) is a branch of computer science that works to find the best possible way to achieve a specific goal by simulating the human brain. Machine learning is a sub-branch of AI that takes the result from experience and uses them as future instructions without being programmed explicitly [68]. ML can be further classified into three major subtypes, namely, supervised learning, unsupervised learning and semi-supervised learning.

In supervised learning, we have prior knowledge of targeted classes and labels for the data. In unsupervised learning, there is no previous knowledge of the target classes and based on identifying patterns in the data. The combination of both supervised learning and unsupervised learning is called semi-supervised learning. Deep learning (DL) is another sub-branch of ML with more capabilities. Both ML and DL methods perform by learning from their experience. The only difference is that DL executes an action in repeat iterations to achieve the best possible outcome. DL solved the problems end-to-end, whereas ML techniques follow the concept of divide and conquer. In the last decade, an abundant amount of work has done to use both these techniques to enrich cybersecurity [69]. ML and DL techniques use the experience to generate input, but deep learning can repeatedly perform a task without any human interaction. Machine learning divides a problem into smaller pieces to generate the outcome, whereas deep learning generates end-to-end findings. The training time duration is more significant for deep learning and shorter for machine learning. In contrast, the testing time duration is shorter for deep learning and longer for machine learning. Deep learning requires powerful hardware system to perform. Machine learning performs well on the low-end hardware system. Machine learning techniques learn from prior knowledge of labels, whereas deep learning techniques learn from their past mistakes.

There are two main models for deep learning approaches: generative models and deep discriminative models [70]. A deep discriminative model can be further classified into three main classes, namely, recurrent neural networks, deep neural network, and convolutional neural networks. As the name suggests in recurrent neural network data is stored in nodes, and all the nodes establish a connection among each other in the form of loops [71]. A deep neural network is a widely used approach. There are manifold layers in deep neural networks, and the number of layers always exceeds three. A convolutional neural network is a multilayer network which processes unstructured data to generate output in the form of complex features [72]. Generative/unsupervised models are further divided into four classes, namely, deep belief networks, deep autoencoders, deep Boltzmann machines, and restricted Boltzmann machine. A restricted Boltzmann machine contains two layers. One layer is called a hidden layer, and the second layer is called a visible layer. Both the hidden and visible layers are completely connected using a set of weights, but there is no connection within the same layer [73]. A deep belief network contains more than one layer where each layer performs as a restricted Boltzmann machine. In a deep belief network, data is depicted by a visible layer, and the features are represented by a hidden layer [74]. A deep autoencoder achieves less data loss by regenerating the input neurons at the output layer such that the number of input and output neurons remain the same in both layers [75]. The deep Boltzmann machine is a multilayer network which contains multi hidden and visible layers. Each layer is connected with neighbor layer, but the connection is entirely undirected without any connection within a layer [76].

Table 2 presents a summary of frequently used machine learning models for cybersecurity. In this table, we have mentioned the time complexity of each mentioned classification model along with its brief description and limitations. Reference number column mentioned the reference of paper of time complexity value for a particular model. Computational cost, i.e., the time complexity of each model, is obtained after a rigorous literature review and web search. However, in order to have a better detection rate, there is a need to use models with lower time complexity. Generally, the models with linear complexities such as O(n) and log-logarithmic are considered best. However, quadratic, i.e., O(n^2), and cubic, i.e., O(n^3), are also acceptable for most practices. O(n^3) considered slower, but exponential and factorial time complexities are undesirable. It is crucial to use a suitable model as per the situation. There are applications, such as in military, where it is critical to have a model with a higher detection rate. However, there are medical problems such as surgical robots where there is a need for higher trustworthiness instead of a quick response.

Table 2. Summary of Significant ML Techniques in Cybersecurity.

Domain	Categories	Technique	Year	Time Complexity	Ref. No	Description	Limitations
Machine Learning	Supervised Learning	SVM	1963	O(n^2) [1]	[77]	• Can be used for classification and regression. • Plots all the data points in space containing all the number of features and then segregates the classes.	• Unable to handle large or noisier datasets efficiently. • Unable to provide direct probability estimation.
		Naive Bayes	1960	O(mn) [2]	[78]	• The probabilistic classifier used for classification. • Assumes that a feature is entirely independent of all other present features.	• Unrealistic assumption of completely independent features. • Assigns 0 probability if some category in test data set is not present in the training data set.
		Random Forest	2001	O(^{3}O(Mm log n) [3]	[79]	• Composed of many decision trees. • Every decision tree yields a prediction. The prediction with maximum votes will be the overall prediction of model.	• High computational cost • Slow prediction generator
		ANN	2000	O(emnk) [4]	[80]	• Composed of Interconnected Artificial Neurons. • Next Layer input depends on Previous Layer Output.	• High cost and time consuming. • Hard to estimate the impact of an independent variable on dependent variables.
		Decision Tree	1979	O(mn^2) [5]	[81]	• Works on an if-then rule to find the best immediate node and the process continues till the predicted class is obtained.	• Difficult to change the data without affecting the overall structure. Complex, expensive and time-consuming.
	Unsupervised Learning	K-mean	1967	O(kmni) [6]	[82]	• Starts from random centroids refines centroids in iterations till the final cluster analysis.	• High dependency on initial centroids. Inefficient clustering for varying cluster sizes
		DBN	2006	O((n + N)k) [7]	[62]	• Consists of various middle layers of Restricted Boltzmann Machine organized greedily. The last layer performs as a classifier.	• Large datasets are needed • High cost and high power eater

[1]: n = number of instances; [2]: n = number of instances, m = number of attributes; [3]: n = number of instances, m = number of attributes, M = number of trees; [4]: n = number of instances, m = number of attributes, e = number of epochs, k = number of neurons; [5]: n = number of instances, m = number of attributes; [6]: n = number of instances, m = number of attributes, k = count of clusters, i = iteration count until the threshold is reached; [7]: k = number of iterations, n = number of records, N = count of parameters.

2.4. Evaluation Criteria

A confusion matrix or error matrix is used to gauge the performance of the classification model formally. Table 3 presents an error matrix that depicts the classification result into four categories, namely TF, TN, FP, and FN. Other evaluation metrics are formed based on these four categories.

Table 3. Confusion Matrix.

		Predicted Class	
		Benign/Positive	**Malicious/Negative**
Actual Class	Benign/Positive	TP [1]	FN [3]
(Ground Truth)	Malicious/Negative	FP [2]	TN [4]

[1] True positive (TP) = Number of correctly classified normal samples. [2] False Positive (FP) = Number of incorrectly classified malicious samples. [3] False Negative (FN) = Number of incorrectly classified normal samples. [4] True Negative (TN) = Number of correctly classified malicious samples.

2.4.1. Precision

The ratio of the total number of normal correctly classified samples to the total count of all positive classified samples is called precision.

$$\text{Precision (P)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

2.4.2. Recall

The percentage of the total number of normal correctly classified samples to the total count of all positive classified samples is called recall.

$$\text{Recall (R)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

2.4.3. Accuracy

The ratio of the total number of normal correctly classified samples to all the samples in the data set is called accuracy.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{All}} \tag{3}$$

2.4.4. ROC Curve

The receiver operating characteristic curve is used to outline the overall threshold's performance with the true positive rate on the y-axis and false positive rate on the x-axis.

2.4.5. Error Rate

The ratio of the total number of misclassified samples to all the samples in the dataset is called the error rate.

$$\text{Error Rate} = \frac{\text{FP} + \text{FN}}{\text{TN} + \text{FP} + \text{FN} + \text{TP}} \tag{4}$$

3. Performance Comparison of Machine Learning Models Applied in Cybersecurity

Researchers are investigating machine learning techniques to detect different cybercrimes in cybersecurity. We have provided a detailed discussion of various cyber threats in Section 2. Furthermore, we have briefly presented an overview of frequently used security datasets in Section 2. This section provides a comprehensive survey of each ML model applied to deal with different cyber threats. Subsequent lines will explain the description of each column in Tables 4–9. The ML technique columns describe the considered machine learning model. We have considered six ML models for this study: random forest, support vector machine, naïve Bayes, decision tree, artificial neural network, and deep belief network.

We focus on three critical cyber threats, namely intrusion detection, spam detection and malware detection. The domain columns state the significant cybersecurity threats considered for this review. The reference number and year columns depict the citation number of each article and published year, respectively. The values of approach or sub-domain columns are different for each cyber threat. IDS domain has three values that are anomaly-based, signature/misuse-based and hybrid-based. Malware has three further sub-classifications that are static, dynamic and hybrid. In the case of spam, sub-domains correspond to the medium in which the authors tried to identify the spam such as image, video, email, SMS and tweets. A description of each sub-domain/approach has been provided in Section 2. Finally, the result attribute presents the evaluation of each classifier applied in a particular sub-domain of cyber threat on a specific dataset and provided in the cited paper mentioned in the reference column.

3.1. Support Vector Machine

The principle superiority of support vector machine (SVM) is that it produces the most successful results for cybersecurity tasks. SVM distributes each data class on both sides of the hyperplane. SVM separates the classes based on the notation to the margin. Support vector points are those points that lie on the border of the hyperplane. The major drawback of the support vector machine is that it consumes an immense amount of space and time. SVM requires data trained on different time intervals to produce better results for a dynamic dataset [83].

SVM showed an accuracy of 99.30% with KDD Cup 99 dataset for IDS [84]. 96.92% is the best reported accuracy for malware detection using Enron dataset [85] and 96.90% with Spambase to classify spam emails [86]. The best reported recall for SVM to detect intrusion is 82% [87], malware is 100% [88], and spam is 98.60% [89]. SVM has obtained best precision while detecting the intrusion is 74% [90], malware is 96.16% [91], and spam is 98.60% [89]. A detailed performance comparison of SVM to various cyber threats on the frequently used dataset is presented in Table 4.

Table 4. Evaluation of SVM in Cybersecurity.

ML Technique	Domain	Dataset	Reference	Year	Approach/Domain	Results		
						Accuracy	Precision	Recall
SVM	IDS	NSL-KDD	[92]	2019	Anomaly-Based	89.70%		
			[93]	2016	Anomaly-Based	98.89%	-	
			[87]	2014	Hybrid-Based	82.37%	74%	82%
		DARPA	[94]	2007	Hybrid-Based	69.80%		
			[95]	2014	Anomaly-Based	95.11%		-
		KDD CUP99	[96]	2011	Hybrid-Based	95.72%		
			[97]	2015	Hybrid-Based	96.08%	-	
			[84]	2014	Hybrid-Based	99.30%	-	
	Malware	Custom Dataset	[98]	2019	Static	95.17%	95.57%	95%
			[99]	2018	Static	89.91%	88.84%	
			[91]	2018	Dynamic	96.27%	96.16%	93.71%
		Malware Dataset	[100]	2017	Static	94.37%		
			[101]	2013	Dynamic	95%		
			[102]	2015	Dynamic	97.10%		
		Enron	[88]	2016	Static	91%	84.74%	100%
			[85]	2007	Static	96.92%	92.74%	97.27%
	Spam	SMS Collection	[89]	2014	SMS Spam	98.61%	98.60%	98.60%
		Spambase	[103]	2015	Email Spam	79.50%	79.02%	68.67%
			[86]	2011	Email Spam	96.90%	93.12%	95%
		Twitter Dataset	[104]	2018	Spam Tweets	93.14%	92.91%	93.14%
			[29]	2015	Spam Tweets	95.20%		93.60%
			[105]	2020	Spam Tweets	98.88%		94.47%

3.2. Decision Tree

Decision tree (DT) belongs to the category of supervised machine learning. DT consists of a path and two nodes: root/intermediate and leaf. Root or intermediate node presents an attribute that followed a path that corresponds to the possible value of an attribute. Leaf node represents the final decision/classification class. A decision tree is used to find the best immediate node by following the if-then rule [106]. Further, 99.96% is the reported accuracy of DT while detecting the anomaly-based IDS with KDD dataset [107]. With standard SMOTE dataset, DT shows an outstanding accuracy of 96.62% for malware detection [108]. With the Enron dataset, DT correctly classified ham emails with an accuracy of 96% [88]. The best reported recall for DT to detect intrusion is 98.10% [90], malware is 96.70% [109], and spam is 96.60% [89]. DT has obtained best precision while detecting the intrusion is 99.70% [90], malware is 99.40% [110], and spam is 98% [88]. A detailed performance comparison of decision tree to various cyber threats on the frequently used dataset is presented in Table 5.

Table 5. Evaluation of Decision Tree in Cybersecurity.

ML Technique	Domain	Dataset	Reference	Year	Approach/Domain	Results		
						Accuracy	Precision	Recall
Decision Tree	IDS	KDD	[107]	2018	Misuse-Based	99.96%		
			[90]	2005	Hybrid-Based	99.85%	99.70%	98.10%
			[111]	2017	Hybrid-Based	86.29%		78%
		NSL-KDD	[112]	2014	Anomaly-Based	99.64%		
			[113]	2017	Hybrid-Based	90.30%	91.15%	90.31%
			[114]	2019	Hybrid-Based	93.40%		
		KDD CUP99	[115]	2015	Misuse-Based	95.09%		
			[116]	2016	Hybrid-Based	99.62%		
			[117]	2018	Hybrid-Based	92.87%	99.90%	
	Malware	Custom	[110]	2016	Static	99.90%	99.40%	
			[118]	2017	Static	84.7%		
		Malware Dataset	[109]	2014	Static		97.90%	96.70%
			[119]	2013	Static	92.34%	-	93%
			[120]	2013	Dynamic	88.47%		
		SMOTE	[121]	2018	Dynamic	92.82%		
			[121]	2018	Dynamic	95.75%		
			[108]	2012	Static	96.62%		
	Spam	SMS Collection	[89]	2014	SMS Spam	96.60%	96.50%	96.60%
		Enron	[88]	2016	Email Spam	96%	98%	94%
			[88]	2016	Email Spam	96%	98%	94%
		Spambase	[122]	2014	Email Spam	92.08%	91.51%	88.08%
			[123]	2014	Email Spam	94.27%		91.02%
			[124]	2013	Email Spam	92.34%	93.90%	93.50%

3.3. Deep Belief Network

A deep belief network (DBN) consists of various middle layers of restricted Boltzmann machine (RBM) organized greedily. Every layer communicates with the layers behind it and the layers ahead of it. There is no lateral communication between the nodes within a layer. Every layer serves as both an input layer and an output layer, except the first and the last layers. The last layer functions as a classifier. The primary purpose of a deep belief network is image clustering and image recognition. It deals with motion capture data. Deep belief network has shown the accuracy of 97.50% for IDS [125], 91.40% for malware detection [126] and 97.43% for spam detection [127] with KDD, KDD CUP99, and Spambase datasets, respectively. The best reported recall for DBN to detect intrusion is 99.70% [128], malware is 98.80% [129], and spam is 98.02% [130]. DBN obtained the best precision while detecting the intrusion is 99.20% [128], malware is 95.77% [131], and spam is 98.39% [130]. A detailed performance comparison of DBN to various cyber threats on the frequently used dataset is presented in Table 6.

Table 6. Evaluation of DBN in Cybersecurity.

ML Technique	Domain	Dataset	Reference	Year	Approach/Domain	Results		
						Accuracy	Precision	Recall
DBN	IDS	KDD	[125]	2015	Anomaly-Based	97.50%		
			[132]	2015	Hybrid-Based	96.70%	97.90%	
		NSL-KDD	[51]	2017	Anomaly-Based	90.40%	88.60%	95.30%
			[128]	2019	Anomaly-Based	99.45%	99.20%	99.70%
		ISCX Dataset	[133]	2015	Misuse-Based	99.18%	-	-
	Malware	DLL	[129]	2008	Static	89.90%	87.40%	98.80%
		Custom	[131]	2016	Static	89.03%	83%	98.18%
			[131]	2016	Dynamic	71%	78.08%	59.09%
			[131]	2016	Hybrid	96.76%	95.77%	97.84%
		KDD CUP99	[134]	2015	Hybrid	91.40%	-	95.34%
	Spam	TARASSUL	[130]	2016	Email Spam	96.40%	95.31%	93.59%
			[130]	2016	Email Spam	97.50%	98.39%	98.02%
		Enron	[128]	2016	Email Spam	95.86%	96.49%	95.61%
			[85]	2007	Email Spam	97.43%	94.94%	96.47%
		Spambase	[135]	2018	Email Spam	89.20%	96%	
			[135]	2018	Email Spam	90.69%	97%	

3.4. Artificial Neural Network

An artificial neural network (ANN) classier consists of hidden neuron input and output layers and performs in two stages. The first stage is called feedforward. In this stage, each hidden layer receives some input nodes and based on the input layer and activation function, the error is calculated. In the second stage, namely feedback stage, the error is sent back to the input layer and process is continued in iterations until the correct result is gained [136]. The artificial neural network showed an accuracy of 97.53% for IDS [137], 92.19% for malware detection [138], and 92.41% for spam detection with NSL-KDD, VX Heavens, and Spambase datasets, respectively. The best reported recall for ANN to detect an intrusion is 98.94% [139], and spam is 94% [140]. ANN has obtained best precision while detecting the intrusion is 97.89% [139], malware is 88.89% [141], and spam is 95% [142]. A detailed performance comparison of ANN to various cyber threats on the frequently used dataset is presented in Table 7.

Table 7. Evaluation of ANN in Cybersecurity.

ML Technique	Domain	Dataset	Reference	Year	Approach/Domain	Results		
						Accuracy	Precision	Recall
ANN	IDS	NSL-KDD	[143]	2019	Anomaly-Based	94.50%	-	-
			[137]	2014	Anomaly-Based	97.53%	-	-
			[94]	2014	Hybrid-Based	97.06%	-	-
		DARPA	[45]	2015	Anomaly-Based	80%	-	80%
			[107]	2018	Misuse-Based	99.82%	-	-
		KDD CUP99	[139]	2009	Anomaly-Based	-	97.89%	98.94%
			[144]	2012	Anomaly-Based	62.90%	-	-
	Malware	VX Heavens	[141]	2012	Hybrid	88.89%	88.89%	-
			[138]	2012	Static	92.19%	-	-
			[145]	2013	Static	88.31%	-	-
		Enron	[136]	2018	Dynamic	82.79%	-	-
		Comodo	[146]	2016	Static	92.02%	-	-
	Spam	Spam-Archive	[140]	2011	Image Spam	93.70%	87%	94%
		Spambase	[147]	2016	Email Spam	91%	-	-
			[148]	2018	Email Spam	92.41%	92.40%	92.40%
			[142]	2013	Hybrid	93.71%	95%	-
		Twitter Dataset	[104]	2018	Spam Tweets	91.18%	91.80%	91.18%

3.5. Random Forest

Random forest (RF) follows through the task by combing different predictions generated by joining different decision trees. RF raised a hypothesis to obtain a result [127]. RF falls under the category of ensemble learning. RF also termed as random decision forest. RF is considered as an improved version of CART that is a sub-type of a decision tree.

RF has shown an accuracy of 99.95% with IDS [149], 95.60% with malware detection [150] and 99.54% for spam detection [151] with KDD, VirusShare, and Spambase datasets, respectively. The best reported recall for RF to detect intrusion is 99.95% [149], malware is 97.30% [109], and spam is 97.20% [89]. RF obtained the best precision while detecting the intrusion is 99.80% [152], malware is 98.58% [98], and spam is 98.60% [153]. A detailed performance comparison of RF to various cyber threats on the frequently used dataset is presented in Table 8.

Table 8. Evaluation of Random Forest in Cybersecurity.

ML Technique	Domain	Dataset	Reference	Year	Approach/Domain	Results		
						Accuracy	Precision	Recall
Random Forest	IDS	KDD	[149]	2019	Anomaly-Based	99.95%		99.95%
			[154]	2016	Anomaly-Based	88.65%	-	94.62%
		NSL-KDD	[152]	2019	Anomaly-Based	95.10%	92.50%	
			[155]	2019	Hybrid-Based	75.30%	81.40%	75.30%
			[156]	2017	Hybrid-Based	97.10%		
		KDD CUP99	[152]	2019	Anomaly-Based	96.30%	99.80%	
			[157]	2016	Anomaly-Based	-	98.10%	98.10%
			[156]	2017	Hybrid-Based	98.10%	-	-
	Malware	Custom Dataset	[98]	2019	Static	98.63%	98.58%	98.69%
			[91]	2018	Dynamic	96.34%	96.59%	93.46%
		Malware Dataset	[158]	2016	Dynamic	96.14%		
			[109]	2014	Hybrid		96.50%	97.30%
			[159]	2017	Hybrid	91.40%	89.80%	91.10%
		VirusShare	[150]	2009	Static	95.60%	96%	
	Spam	SMS Collection	[89]	2014	SMS Spam	97.18%	97.30%	97.20%
		Spambase	[151]	2013	Email Spam	99.54%		
			[160]	2010	Email Spam	95.43%		
			[124]	2013	Email Spam	93.89%	95.87%	94.10%
		Twitter Dataset	[161]	2011	Spam Tweets	95%	95.70%	95.70%
			[153]	2016	Spam Tweets	96.20%	98.60%	75.50%
			[104]	2018	Spam Tweets	93.43%	93.25%	93.43%

3.6. Naïve Bayes

The major limitation for Naïve Bayes (NB) classifier is that it assumes that every attribute is independent, and none of the attributes has a relationship with each other. This state of independence is technically impossible in cyberspace. Hidden NB is an advanced form of Naïve Bayes, and it gives 99.6% accuracy [162]. Naïve Bayes showed an accuracy of 99.90% with DARPA dataset for IDS [163]. 99.50% is the best reported accuracy for malware detection using NSL-KDD dataset [164]. With Spambase dataset, Naïve Bayes showed considerable accuracy of 96.46 % to classify spam or ham email [86]. The best reported recall for NB to detect intrusion is 100% [33], malware is 95.90% [164], and spam is 98.46% [86]. NB obtained the best precision while detecting the intrusion is 99.04% [163], malware is 97.50% [109], and spam is 99.66% [86]. A detailed performance comparison of NB to various cyber threats on the frequently used dataset is presented in Table 9.

Table 9. Evaluation of Naïve Bayes in Cybersecurity.

ML Technique	Domain	Dataset	Reference	Year	Approach/Domain	Results		
						Accuracy	Precision	Recall
Naïve Bayes	IDS	DARPA	[33]	2010	Anomaly-Based	91.60%		61.60%
			[163]	2007	Misuse-Based	99.90%	99.04%	99.50%
		NSL-KDD	[115]	2015	Misuse-Based	81.66%		
			[165]	2012	Anomaly-Based	36%	35%	80%
			[165]	2012	Anomaly-Based	99%	83%	78.90%
		KDD CUP99	[166]	2004	Anomaly-Based	99.27%		
			[163]	2007	Anomaly-Based		96%	99.80%
			[33]	2018	Signature-Based	99.72%		100%
	Malware	VX Heaven	[167]	2015	Static	88.80%		
		NSL-KDD	[168]	2013	Hybrid	99.50%		
			[169]	2007	Hybrid	99%		
		Malware Dataset	[119]	2013	Hybrid	89.81%	-	90%
			[164]	2015	Hybrid	95.90%	95.90%	95.90%
			[109]	2014	Hybrid		97.50%	67.40%
	Spam	SMS Collection	[89]	2014	SMS Spam	97.52%	97.50%	97.50%
		Spambase	[86]	2011	Email Spam	99.46%	99.66%	98.46%
			[103]	2015	Email Spam	76.24%	70.59%	72.05%
			[170]	2015	Email Spam	84%	89%	78%
		Twitter Dataset	[124]	2013	Spam Tweets	92%	91.60%	91.4%
			[104]	2018	Spam Tweets	92.06%	91.69%	91.96%

4. Discussion and Conclusions

Machine learning techniques have become the most integral underlying part of the modern cyber world, particularly for cybersecurity. Machine learning techniques are being applied on both sides, i.e., attacker side and defender side. On the attacker side, machine learning techniques are being used to find new ways to pass through and evade the security system and firewall. On the defender side, these techniques are helping security professional to protect the security systems from illegal penetration and unauthorized access. This paper reviews a comparative analysis of machine learning techniques applied to detect cybersecurity threats. We have considered three significant threats to cyberspace: intrusion detection, spam detection, and malware detection. We have compared six machine learning models, namely, random forest, support vector machine, naïve Bayes, decision tree, artificial neural network, and deep belief network. We have further compared these models on further sub-domain of cyber threats. The sub-domains of each cyber threat are different. Anomaly-based, signature-based, and hybrid-based are considered sub-domains for intrusion detection. For malware detection, the sub-domains are either static detection, dynamic detection or hybrid-detection. Sub-domains for the spam are the medium on which the models are applied to classify spam like images, videos, emails, SMS or calls. Section 2 described each sub-domain of threat in detail. This section is divided into two parts. First part provides the discussion on the performance of various ML models applied in cybersecurity. The second part provides the challenges of using machine models in cybersecurity and concludes the study.

4.1. *Performance Evaluation of ML Models*

Figure 2 shows the performance comparison of six machine learning techniques based on frequently used datasets to detect intrusion detection. We have picked the values from the given tables that show the maximum value for accuracy, precision and recall based on the dataset. SVM has revealed an outstanding performance of nearly 98% on KDD dataset whereas the utmost accuracy for SVM reported on NSL-KDD dataset was 83%. DBN performed outstanding nearly on all datasets and shown an accuracy above 95% to detect intrusion. On the DARPA dataset, NB and ANN performed better accuracy than other models, but ANN has given worse precision value on DARPA dataset. On NSL-KDD dataset, DBN performed best among other models concerning accuracy, precision, and recall. The SVM and DBN came up with an excellent precision on KDD-Cup 99 dataset among all other models. On KDD dataset, decision tree and random forest have shown excellent precision rate among all the models. Random forest on KDD dataset, NB on DARPA, DBN on NSL-KDD, and NB on KDD Cup99 have shown the best recall rates, respectively.

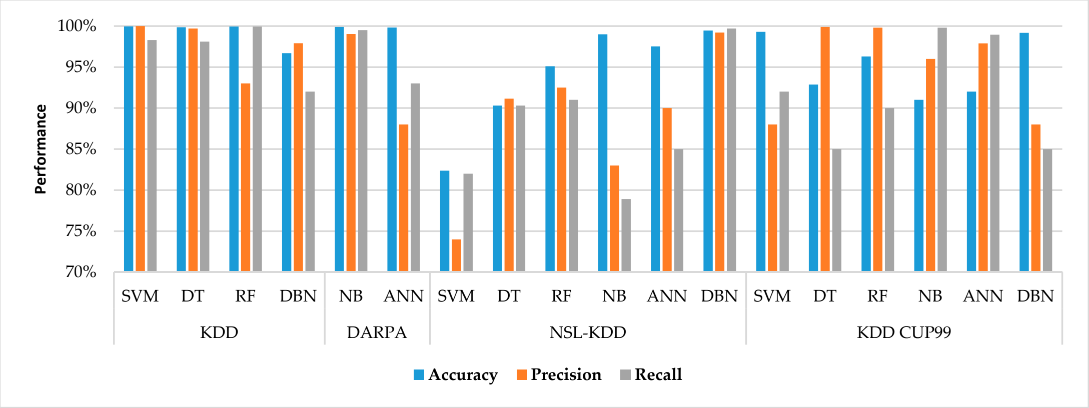

Figure 2. A comparative Analysis of Intrusion Detection using Machine Learning Techniques.

Figure 3 shows the performance evaluation of six machine learning techniques on frequently used datasets to detect malware. We have observed that there are not much benchmark datasets available for malware detection. Mostly, the researchers collected their customized datasets and applied machine learning techniques to evaluate the models. We have also noticed that machine learning techniques are often shown outstanding accuracy, precision, and recall values on the customized dataset. These proposed techniques don't show similar best performance when applied to other datasets. Classical machine learning techniques, e.g., decision tree performed better on several datasets. DBN showed an outstanding recall value almost on all datasets. DT and RF have performed with a better precision rate on VirusShare dataset. RF has shown excellent recall, precision, and accuracy values on Enron dataset. ANN has shown worst performance on Enron dataset for accuracy, recall and precision. With respect to accuracy, the NB and DT performed excellently on the VirusShare dataset compared to other models.

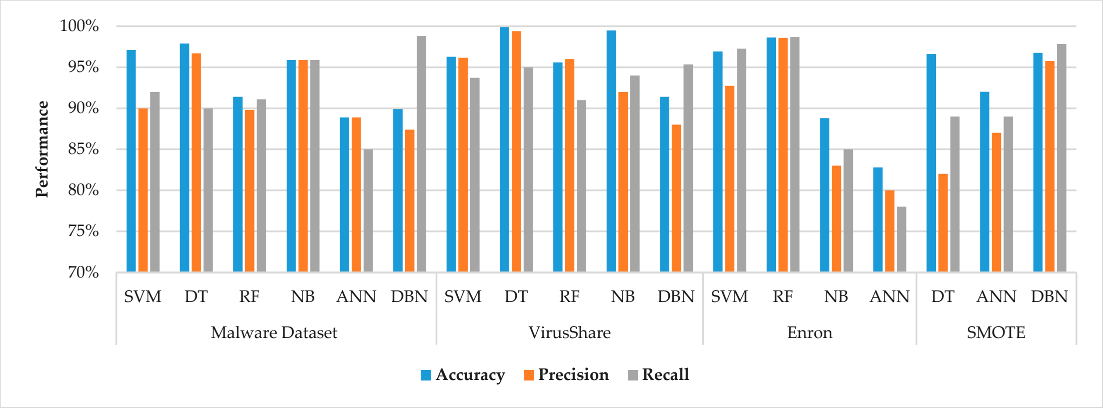

Figure 3. A comparative Analysis of Malware Detection using Machine Learning Techniques.

Figure 4 shows performance evaluation of machine learning techniques based on frequently used datasets for spam classification. Spambase is a famous spam dataset, and NB performed better among other machine learning models with respect to the accuracy, precision, and recall. Researchers have also collected dataset from Twitter, containing millions of tweets. RF has outperformed and reported more than 97% precision value. All evaluated machine learning models have shown more than 90%

accuracy to detect and classify spam. SVM and DBN have shown better accuracy, recall and precision among other models when applied to the collection of SMS to classify spam text messages. It can also be observed that SMS collection is a customized dataset collected by the researcher. Every machine learning model has given more than 95% accuracy, precision, and recall, whereas the same machine learning models have different performance values on standard datasets, i.e., Enron and Spambase.

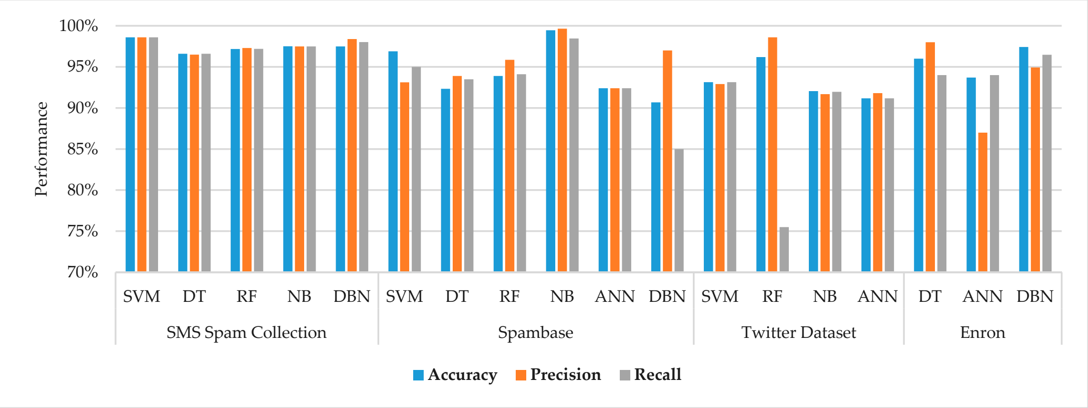

Figure 4. A comparative Analysis of Spam Detection using Machine Learning Techniques.

Figure 5 shows the comparative analysis of accuracy, precision and recall values for the detection of intrusion, spam and malware. We have taken the maximum value obtained by six machine learning models regardless of the dataset. It is depicted that SVM, DT and RF have given the maximum accuracy and precision value for intrusion detection. However, DBN and ANN reported the best recall value for the detection of intrusion. It is recommended that SVM, DT, NB, and RF should be considered for intrusion detection if accuracy is the priority for intrusion detection. DBN and ANN comparatively performed worse than other models to detect malware. However, ANN has shown exceptional recall value for the detection of malware. RF and NB have shown better accuracy for the classification of spams yet DBN again recommended in case precision and recall is the priority of situation. Keeping in view the metrics collected from reviewed papers, RF and NB are recommended for the classification of spam, and SVM and DT for the detection of malware, respectively, for better accuracy.

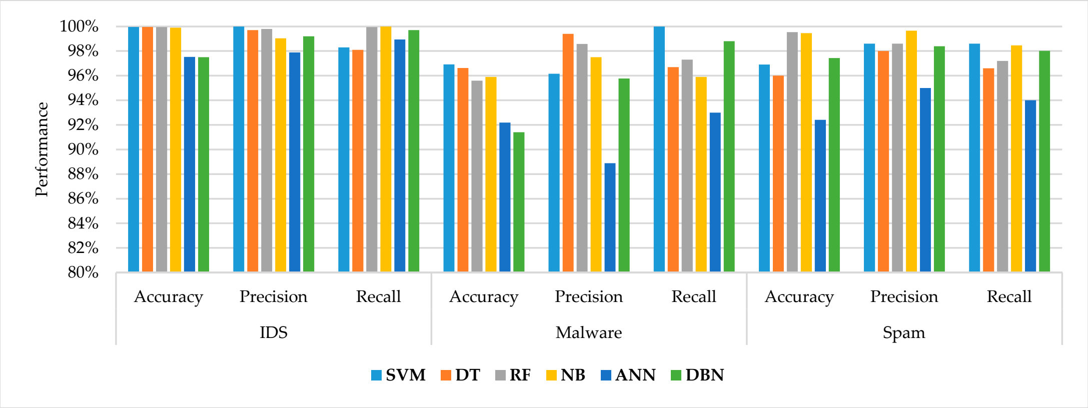

Figure 5. A Performance Evaluation of Machine Learning Techniques.

4.2. Challenges of Using ML Models in Cybersecurity and Future Directions

The application of machine learning techniques to detect several cyber threats has shown better results than conventional methods. Despite having all those improvements, machine learning techniques are still facing many challenges in the cybersecurity domain. We have presented a comparison of machine learning techniques based on frequently used datasets. The unavailability of benchmark and updated datasets for the training of machine learning models is a big challenge. Another unbalancing trend is that the same dataset is generating different results using the same techniques for the same sub-domain. In the Table 4, SVM is applied to detect anomaly-based intrusion by [92,93] but having a difference of 10% in accuracy. The same can be seen in the case of [98,99] in Table 4, [116,117] in Table 5, [151,160] in Table 8, and [86,103] in Table 9 to name a few.

The unbalance in the accuracy may be due to the selection of different feature extraction methods or data for testing and training purposes. However, this discrepancy in results has created confusion while selecting a suitable classifier for a particular problem. In addition to these problems, the available datasets need to rationalize by banishing the redundant, noisy, missing and unbalanced data. These datasets should be up to date with modern and sophisticated attacks, and missing values should be detected and removed from the required data.

The detection speed of particular cyber threat and prompt action are critical challenges for machine learning models. The time complexity of machine learning models matters in this case. Faster and robust models will detect the cyber threat beforehand and stop it from creating any problems for network and system. However, in order to have a better detection rate and to take a prompt action, the models with lower time complexity are suggested to use. Generally, the models with linear complexities such as O(n) and log-logarithmic are considered best. The choice of machine learning model will vary from case to case. The time complexity of each model is obtained after a rigorous literature review and web search. We have provided the time complexity of significant machine learning models used in cybersecurity in Table 2. It is essential to have efficient models to produce the best detection rate and quick response in some scenarios, such as the military. If the detection rate of models will be slower, then attackers will dodge the model to harm the system before any preventative measures are taken.

Deep learning can handle data without human interaction; however, still have several limitations. In comparison to machine learning techniques, deep learning techniques need an enormous amount of data and costly hardware components to produce better results. The substantial magnitude of time and power is required to process more massive datasets. In order to train the model, there is a need for high computing hardware such as GPUs and parallel processing to expedite the learning and classification processes. The growing rate of unlabeled, sparse, and missing data also affects the training process of the models. There is a need to have high computational efficiency where maximum throughput is trying to achieve with limited resources.

There is a significant need to have high-level of correctness instead of speed and response of prediction in some scenarios. Trustworthiness is essential when machine learning techniques are being applied in life-critical or mission-critical applications such as self-driving cars. Image classification is very critical to correctly read a traffic sign by self-driving cars. Prediction cannot be applied with blind faith where robots are doing the treatment and surgery of cancer patients.

We have also observed that, even in 2020, researchers are applying and testing the latest machine learning and deep learning techniques on outdated datasets. It is apparent from the year column of Tables 4–9 that latest machine learning techniques are being tested on DARPA and KDD Cup datasets which are more than 15 years old. There is a need to have the latest, benchmark, and real-time datasets to evaluate latest machine learning and deep learning models. There are benchmark datasets for intrusion detection like DARPA and KDD Cup. However, we have perceived that there is a deficiency of state-of-the-art datasets for spam and malware detection problems. Researchers are applying latest machine learning techniques on their customized datasets. They claim to have the better accuracy of their models without disclosing their datasets and code to regenerate the results. Customized datasets are often collected in a particular fashion that lack diversity and their proposed model(s) performed well on those datasets. However, when the same models are tested on other similar problem domain on different dataset, the models don't show the similar best results as claimed by authors on their customized datasets.

Furthermore, researchers are publishing the performance evaluation of their proposed models using different metrics. Some are publishing the recall while others are only focusing on accuracy. There should be standardized metrics to compare the performance of models. It can be observed from Tables 4–9 that most of the researchers have published the accuracy of their model, leaving other metrics. False-negative is a value that describes the case whereby an illegitimate user has been granted with access to a system and network. This would have a drastically worse effect on system performance rather than considering accuracy. There should be standard metrics in order to compare the models using different measurements. This would then be a milestone for future research to improve the performance of models.

Further, there is a need to have robust machine learning models to handle adversarial inputs. There should be an emphasis on training the model in adversarial settings to develop robust models against adversarial inputs. We have reviewed the six machine learning models based on several datasets to detect a cyber threat but encourage a beginner in this domain to delve into the extensive bibliography presented in this review paper. In future work, we will analyze more ML and DL techniques against several other cybersecurity threats. We will evaluate the ML models in other areas of cybersecurity, such as IoT, smart cities, methods based on API calls, cellular network, and smart grids.

Author Contributions: Conceptualization, K.S., S.L.; Data analysis and initial draft: K.S., S.L.; Methodology: K.S., S.L., and V.V.; First draft preparation, K.S., S.L., V.V., I.A.H., S.C.; Writing—Reviewing and editing, K.S., S.L., V.V., I.A.H., S.C., D.L., and J.L.; Supervision: S.L., V.V., I.A.H., S.C., D.L., and J.L.; Funding Acquisition, I.A.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. APC is paid by Norwegian University of Science and Technology, Norway.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. ICT Facts and Figures 2017. Available online: https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx (accessed on 9 October 2019).
2. Craigen, D.; Diakun-Thibault, N.; Purse, R. Defining cybersecurity. *Technol. Innov. Manag. Rev.* **2014**, *4*, 13–21. [CrossRef]
3. Farahmand, F.; Navathe, S.B.; Enslow, P.H.; Sharp, G.P. Managing vulnerabilities of information systems to security incidents. In Proceedings of the 5th International Conference on Electronic Commerce, Pittsburgh, PA, USA, 30 September–3 October 2003; pp. 348–354.
4. Szor, P. *The Art of Computer Virus Research and Defense: ART COMP VIRUS RES DEFENSE _p1*; Pearson Education: London, UK, 2005.
5. Firdausi, I.; Erwin, A.; Nugroho, A.S. Analysis of machine learning techniques used in behavior-based malware detection. In Proceedings of the 2010 Second International Conference on Advances in Computing, Control, and Telecommunication Technologies, Jakarta, Indonesia, 2–3 December 2010; pp. 201–203.
6. Michie, D.; Spiegelhalter, D.J.; Taylor, C. Machine learning. *Neural Stat. Classif.* **1994**, *13*, 1–298.
7. Shaukat, K.; Nawaz, I.; Zaheer, S. *Students Performance: A Data Mining Perspective*; LAP Lambert Academic Publishing: Saarbrücken, Germany, 2017.
8. Shaukat, K.; Nawaz, I.; Aslam, S.; Zaheer, S.; Shaukat, U. Student's performance in the context of data mining. In Proceedings of the 2016 19th International Multi-Topic Conference (INMIC), Islamabad, Pakistan, 5–6 December 2016; pp. 1–8.
9. Shaukat, K.; Masood, N.; Mehreen, S.; Azmeen, U. Dengue fever prediction: A data mining problem. *J. Data Min. Genom. Proteom.* **2015**, *2015*. [CrossRef]
10. Jusas, V.; Samuvel, S.G. Classification of motor imagery using combination of feature extraction and reduction methods for brain-computer interface. *Inf. Technol. Control* **2019**, *48*, 225–234. [CrossRef]
11. Uktveris, T.; Jusas, V. Comparison of feature extraction methods for EEG BCI classification. In Proceedings of the International Conference on Information and Software Technologies, Vilnius, Lithuania, 10–12 October 2015; pp. 81–92.

12. Shaukat, K.; Rubab, A.; Shehzadi, I.; Iqbal, R. A Socio-Technological analysis of Cyber Crime and Cyber Security in Pakistan. *Transylv. Rev.* **2017**, *1*, 84.
13. Canhoto, A.I.; Clear, F. Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. *Bus. Horiz.* **2019**, *63*, 183–193. [CrossRef]
14. Maqsood, H.; Mehmood, I.; Maqsood, M.; Yasir, M.; Afzal, S.; Aadil, F.; Selim, M.M.; Muhammad, K. A local and global event sentiment based efficient stock exchange forecasting using deep learning. *Int. J. Inf. Manag.* **2020**, *50*, 432–451. [CrossRef]
15. Dey, S.; Ye, Q.; Sampalli, S. A machine learning based intrusion detection scheme for data fusion in mobile clouds involving heterogeneous client networks. *Inf. Fusion* **2019**, *49*, 205–215. [CrossRef]
16. Geluvaraj, B.; Satwik, P.; Kumar, T.A. The future of cybersecurity: Major role of artificial intelligence, machine learning, and deep learning in cyberspace. In *International Conference on Computer Networks and Communication Technologies*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 739–747.
17. Jain, P. Machine Learning Versus Deep Learning for Malware Detection. Master's Thesis, San Jose State University, San Jose, CA, USA, 2019.
18. Rao, R.S.; Pais, A.R. Detection of phishing websites using an efficient feature-based machine learning framework. *Neural Comput. Appl.* **2019**, *31*, 3851–3873. [CrossRef]
19. Alauthman, M.; Almomani, A.; Alweshah, M.; Omoushd, W.; Alieyane, K. Machine Learning for phishing Detection and Mitigation. *Mach. Learn. Comput. Cyber Secur. Princ. Algorithmsand Pract.* **2019**, *26*, 48–74.
20. Alurkar, A.A.; Ranade, S.B.; Joshi, S.V.; Ranade, S.S.; Shinde, G.R.; Sonewar, P.A.; Mahalle, P.N. A Comparative Analysis and Discussion of Email Spam Classification Methods Using Machine Learning Techniques. In *Applied Machine Learning for Smart Data Analysis*; CRC Press: Boca Raton, FL, USA, 2019; p. 185.
21. Dada, E.G.; Bassi, J.S.; Chiroma, H.; Adetunmbi, A.O.; Ajibuwa, O.E. Machine learning for email spam filtering: Review, approaches and open research problems. *Heliyon* **2019**, *5*, e01802. [CrossRef] [PubMed]
22. Shukur, H.A.; Kurnaz, S. Credit Card Fraud Detection Using Machine Learning Methodology. *Int. J. Comput. Sci. Mob. Comput.* **2019**, *8*, 257–260.
23. Afek, Y.; Bremler-Barr, A.; Feibish, S.L. Zero-day signature extraction for high-volume attacks. *IEEE/ACM Trans. Netw.* **2019**, *27*, 691–706. [CrossRef]
24. Saad, S.; Briguglio, W.; Elmiligi, H. The Curious Case of Machine Learning In Malware Detection. *arXiv* **2019**, arXiv:1905.07573.
25. Ambalavanan, V. Cyber Threats Detection and Mitigation Using Machine Learning. In *Handbook of Research on Machine and Deep Learning Applications for Cyber Security*; IGI Global: Hershey, PA, USA, 2020; pp. 132–149.
26. Shah, N.F.; Kumar, P. A comparative analysis of various spam classifications. In *Progress in Intelligent Computing Techniques: Theory, Practice, and Applications*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 265–271.
27. Chandrasekar, C.; Priyatharsini, P. Classification techniques using spam filtering email. *Int. J. Adv. Res. Comput. Sci.* **2018**, *9*, 402. [CrossRef]
28. Shafi'I, M.A.; Latiff, M.S.A.; Chiroma, H.; Osho, O.; Abdul-Salaam, G.; Abubakar, A.I.; Herawan, T. A review on mobile SMS spam filtering techniques. *IEEE Access* **2017**, *5*, 15650–15666.
29. Chen, C.; Zhang, J.; Xie, Y.; Xiang, Y.; Zhou, W.; Hassan, M.M.; AlElaiwi, A.; Alrubaian, M. A performance evaluation of machine learning-based streaming spam tweets detection. *IEEE Trans. Comput. Soc. Syst.* **2015**, *2*, 65–76. [CrossRef]
30. Biggio, B.; Fumera, G.; Pillai, I.; Roli, F. A survey and experimental evaluation of image spam filtering techniques. *Pattern Recognit. Lett.* **2011**, *32*, 1436–1446. [CrossRef]
31. Kumar, A.D.; KP, S. DeepImageSpam: Deep Learning based Image Spam Detection. *arXiv* **2018**, arXiv:1810.03977.
32. Jusas, V.; Japertas, S.; Baksys, T.; Bhandari, S. Logical filter approach for early stage cyber-attack detection. *Comput. Sci. Inf. Syst.* **2019**, *16*, 491–514. [CrossRef]
33. Xin, Y.; Kong, L.; Liu, Z.; Chen, Y.; Li, Y.; Zhu, H.; Gao, M.; Hou, H.; Wang, C. Machine learning and deep learning methods for cybersecurity. *IEEE Access* **2018**, *6*, 35365–35381. [CrossRef]
34. Gandotra, E.; Bansal, D.; Sofat, S. Malware analysis and classification: A survey. *J. Inf. Secur.* **2014**, *5*, 56. [CrossRef]
35. Dharamkar, B.; Singh, R.R. A review of cyber attack classification technique based on data mining and neural network approach. *Int. J. Comput. Trends Technol.* **2014**, *7*, 100–105. [CrossRef]

36. Ford, V.; Siraj, A. Applications of machine learning in cyber security. In Proceedings of the 27th International Conference on Computer Applications in Industry and Engineering, San Diego, CA, USA, 12–14 October 2015.
37. Jiang, H.; Nagra, J.; Ahammad, P. Sok: Applying machine learning in security—A survey. *arXiv* **2016**, arXiv:1611.03186.
38. Hodo, E.; Bellekens, X.; Hamilton, A.; Tachtatzis, C.; Atkinson, R. Shallow and deep networks intrusion detection system: A taxonomy and survey. *arXiv* **2017**, arXiv:1701.02145.
39. Apruzzese, G.; Colajanni, M.; Ferretti, L.; Guido, A.; Marchetti, M. On the effectiveness of machine and deep learning for cyber security. In Proceedings of the 2018 10th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia, 29 May–1 June 2018; pp. 371–390.
40. Yin, X.C.; Liu, Z.G.; Nkenyereye, L.; Ndibanje, B. Toward an Applied Cyber Security Solution in IoT-Based Smart Grids: An Intrusion Detection System Approach. *Sensors* **2019**, *19*, 4952. [CrossRef]
41. Eder-Neuhauser, P.; Zseby, T.; Fabini, J. Malware propagation in smart grid networks: Metrics, simulation and comparison of three malware types. *J. Comput. Virol. Hacking Tech.* **2019**, *15*, 109–125. [CrossRef]
42. Ndibanje, B.; Kim, K.H.; Kang, Y.J.; Kim, H.H.; Kim, T.Y.; Lee, H.J. Cross-method-based analysis and classification of malicious behavior by api calls extraction. *Appl. Sci.* **2019**, *9*, 239. [CrossRef]
43. Torres, J.M.; Comesaña, C.I.; García-Nieto, P.J. Machine learning techniques applied to cybersecurity. *Int. J. Mach. Learn. Cybern.* **2019**, *10*, 2823–2836. [CrossRef]
44. Ucci, D.; Aniello, L.; Baldoni, R. Survey on the usage of machine learning techniques for malware analysis. *arXiv* **2017**, arXiv:1710.08189.
45. Buczak, A.L.; Guven, E. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Commun. Surv. Tutor.* **2015**, *18*, 1153–1176. [CrossRef]
46. Das, R.; Morris, T.H. Machine Learning and Cyber Security. In Proceedings of the 2017 International Conference on Computer, Electrical & Communication Engineering (ICCECE), Kolkata, India, 22–23 December 2017; pp. 1–7.
47. Li, J.-H. Cyber security meets artificial intelligence: A survey. *Front. Inf. Technol. Electron. Eng.* **2018**, *19*, 1462–1474. [CrossRef]
48. Ucci, D.; Aniello, L.; Baldoni, R. Survey of machine learning techniques for malware analysis. *Comput. Secur.* **2019**, *81*, 123–147. [CrossRef]
49. Veiga, A.P. Applications of artificial intelligence to network security. *arXiv* **2018**, arXiv:1803.09992.
50. Sagar, B.; Niranjan, S.; Kashyap, N.; Sachin, D. Providing Cyber Security using Artificial Intelligence–A survey. In Proceedings of the 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 27–29 March 2019; pp. 717–720.
51. Kwon, D.; Kim, H.; Kim, J.; Suh, S.C.; Kim, I.; Kim, K.J. A survey of deep learning-based network anomaly detection. *Clust. Comput.* **2019**, *22*, 949–961. [CrossRef]
52. Berman, D.S.; Buczak, A.L.; Chavis, J.S.; Corbett, C.L. A survey of deep learning methods for cyber security. *Information* **2019**, *10*, 122. [CrossRef]
53. Fischer, E.A. *Creating a National Framework for Cybersecurity: An Analysis of Issues and Options*; Nova Science Publishers: Hauppauge, NY, USA, 2009.
54. Shaukat Dar, K.; Javed, I.; Asad Ammar, A.; Konain Abbas, S.; Asghar, S.; Abu Bakar, M.; Shaukat, U. A survey-data privacy through different methods. *J. Netw. Commun. Emerg. Technol.* **2015**, *5*, 1–7.
55. Purkait, S. Phishing counter measures and their effectiveness–literature review. *Inf. Manag. Comput. Secur.* **2012**, *20*, 382–420. [CrossRef]
56. Shelly, G.B.; Vermaat, M.E. *Discovering Computers-Fundamentals 2011 Edition*; Course Technology Press: Boston, MA, USA, 2010.
57. Shelly, G.B.; Vermaat, M.E. *Discovering Computers*; Course Technology: Boston, MA, USA, 2012.
58. Lippmann, R.P.; Fried, D.J.; Graf, I.; Haines, J.W.; Kendall, K.R.; McClung, D.; Weber, D.; Webster, S.E.; Wyschogrod, D.; Cunningham, R.K. Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation. In Proceedings of the DARPA Information Survivability Conference and Exposition. DISCEX'00, Hilton Head, SC, USA, 25–27 January 2000; pp. 12–26.
59. Panigrahi, R.; Borah, S. A detailed analysis of CICIDS2017 dataset for designing Intrusion Detection Systems. *Int. J. Eng. Technol.* **2018**, *7*, 479–482.

60. Xie, M.; Hu, J.; Slay, J. Evaluating host-based anomaly detection systems: Application of the one-class SVM algorithm to ADFA-LD. In Proceedings of the 2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Xiamen, China, 19–21 August 2014; pp. 978–982.
61. Tavallaee, M.; Bagheri, E.; Lu, W.; Ghorbani, A.A. A detailed analysis of the KDD CUP 99 data set. In Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, Ottawa, ON, Canada, 8–10 July 2009; pp. 1–6.
62. Ferrag, M.A.; Maglaras, L.; Moschoyiannis, S.; Janicke, H. Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *J. Inf. Secur. Appl.* **2020**, *50*, 102419. [CrossRef]
63. Torrano-Gimenez, C.; Pérez-Villegas, A.; Álvarez, G.; Fernández-Medina, E.; Malek, M.; Hernando, J. An Anomaly-based Web Application Firewall. In Proceedings of the International Conference on Security and Cryptography—Volume 1: SECRYPT, Milan, Italy, 7–10 July 2009.
64. Spambase Dataset. Center for Machine Learning and Intelligent Systems at UC Irvine. Available online: https://archive.ics.uci.edu/ml/datasets/Spambase (accessed on 29 October 2019).
65. Mamun, M.S.I.; Rathore, M.A.; Lashkari, A.H.; Stakhanova, N.; Ghorbani, A.A. Detecting malicious urls using lexical analysis. In Proceedings of the International Conference on Network and System Security, Taipei, Taiwan, 28–30 September 2016; pp. 467–482.
66. Shiravi, A.; Shiravi, H.; Tavallaee, M.; Ghorbani, A.A. Toward developing a systematic approach to generate benchmark datasets for intrusion detection. *Comput. Secur.* **2012**, *31*, 357–374. [CrossRef]
67. Gonzalez, H.; Stakhanova, N.; Ghorbani, A.A. Droidkin: Lightweight detection of android apps similarity. In Proceedings of the International Conference on Security and Privacy in Communication Networks, Orlando, FL, USA, 23–25 October 2009; pp. 436–453.
68. Angra, S.; Ahuja, S. Machine learning and its applications: A review. In Proceedings of the 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), Chirala, India, 23–25 March 2017; pp. 57–60.
69. Le Callet, P.; Viard-Gaudin, C.; Barba, D. A convolutional neural network approach for objective video quality assessment. *IEEE Trans. Neural Netw.* **2006**, *17*, 1316–1327. [CrossRef]
70. Deng, L.; Yu, D. Deep learning: Methods and applications. *Found. Trends Signal Process.* **2014**, *7*, 197–387. [CrossRef]
71. Gelly, G.; Gauvain, J.-L. Optimization of RNN-based speech activity detection. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2017**, *26*, 646–656. [CrossRef]
72. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, *77*, 354–377. [CrossRef]
73. Fischer, A.; Igel, C. An introduction to restricted Boltzmann machines. In Proceedings of the Iberoamerican Congress on Pattern Recognition, Havana, Cuba, 28–31 October 2019; pp. 14–36.
74. Hinton, G.E. Deep belief networks. *Scholarpedia* **2009**, *4*, 5947. [CrossRef]
75. Vincent, P.; Larochelle, H.; Lajoie, I.; Bengio, Y.; Manzagol, P.-A. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *J. Mach. Learn. Res.* **2010**, *11*, 3371–3408.
76. Salakhutdinov, R.; Larochelle, H. Efficient learning of deep Boltzmann machines. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010; pp. 693–700.
77. Burges, C.J. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [CrossRef]
78. Frank, E.; Hall, M.A. *Data Mining: Practical Machine Learning tOols and Techniques*; Morgan Kaufmann: Burlington, MA, USA, 2011.
79. Agrawal, R.; Srikant, R. Mining sequential patterns. In Proceedings of the eleventh international conference on data engineering, Taipei, Taiwan, 6–10 March 1995; pp. 3–14.
80. Jain, A.K.; Mao, J.; Mohiuddin, K.M. Artificial neural networks: A tutorial. *Computer* **1996**, *29*, 31–44. [CrossRef]
81. Ross, Q.J. *C4. 5: Programs for Machine Learning. San Mateoca*; Morgan Kaufmann: Burlington, MA, USA, 1993.
82. Jain, A.K.; Dubes, R.C. *Algorithms for Clustering Data*; Prentice-Hall, Inc.: Upper Saddle River, NJ, USA, 1988.
83. Iyer, S.S.; Rajagopal, S. Applications of Machine Learning in Cyber Security Domain. In *Handbook of Research on Machine and Deep Learning Applications for Cyber Security*; IGI Global: Hershey, PA, USA, 2020; pp. 64–82.

84. Saxena, H.; Richariya, V. Intrusion detection in KDD99 dataset using SVM-PSO and feature reduction with information gain. *Int. J. Comput. Appl.* **2014**, *98*, 25–29. [CrossRef]
85. Tzortzis, G.; Likas, A. Deep belief networks for spam filtering. In Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007), Patras, Greece, 29–31 October 2007; pp. 306–309.
86. Awad, W.; ELseuofi, S. Machine learning methods for spam e-mail classification. *Int. J. Comput. Sci. Inf. Technol.* **2011**, *3*, 173–184. [CrossRef]
87. Pervez, M.S.; Farid, D.M. Feature selection and intrusion classification in NSL-KDD cup 99 dataset employing SVMs. In Proceedings of the 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014), Dhaka, Bangladesh, 18–20 December 2014; pp. 1–6.
88. Khan, Z.; Qamar, U. Text Mining Approach to Detect Spam in Emails. In Proceedings of the International Conference on Innovations in Intelligent Systems and Computing Technologies (ICIISCT2016), Las Piñas, Philippines, 24–26 February 2016; p. 45.
89. Najadat, H.; Abdulla, N.; Abooraig, R.; Nawasrah, S. Mobile sms spam filtering based on mixing classifiers. *Int. J. Adv. Comput. Res.* **2014**, *1*, 1–7.
90. Stein, G.; Chen, B.; Wu, A.S.; Hua, K.A. Decision tree classifier for network intrusion detection with GA-based feature selection. In *Proceedings of the 43rd Annual Southeast Regional Conference-Volume 2*; ACM: New York, NY, USA, 2005; pp. 136–141.
91. Feng, P.; Ma, J.; Sun, C.; Xu, X.; Ma, Y.J.I.A. A Novel Dynamic Android Malware Detection System With Ensemble Learning. *IEEE Access* **2018**, *6*, 30996–31011. [CrossRef]
92. Lee, J.; Kim, J.; Kim, I.; Han, K. Cyber Threat Detection Based on Artificial Neural Networks Using Event Profiles. *IEEE Access* **2019**, *7*, 165607–165626. [CrossRef]
93. Sharma, R.K.; Kalita, H.K.; Borah, P. Analysis of machine learning techniques based intrusion detection systems. In *Proceedings of the 3rd International Conference on Advanced Computing, Networking and Informatics*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 485–493.
94. Khan, L.; Awad, M.; Thuraisingham, B. A new intrusion detection system using support vector machines and hierarchical clustering. *VLDB J.* **2007**, *16*, 507–521. [CrossRef]
95. Kokila, R.; Selvi, S.T.; Govindarajan, K. DDoS detection and analysis in SDN-based environment using support vector machine classifier. In Proceedings of the 2014 Sixth International Conference on Advanced Computing (ICoAC), Chennai, India, 17–19 Decmber 2014; pp. 205–210.
96. Horng, S.-J.; Su, M.-Y.; Chen, Y.-H.; Kao, T.-W.; Chen, R.-J.; Lai, J.-L.; Perkasa, C.D. A novel intrusion detection system based on hierarchical clustering and support vector machines. *Expert Syst. Appl.* **2011**, *38*, 306–313. [CrossRef]
97. Masduki, B.W.; Ramli, K.; Saputra, F.A.; Sugiarto, D. Study on implementation of machine learning methods combination for improving attacks detection accuracy on Intrusion Detection System (IDS). In Proceedings of the 2015 International Conference on Quality in Research (QiR), Lombok, Indonesia, 10–13 August 2015; pp. 56–64.
98. Naz, S.; Singh, D.K. Review of Machine Learning Methods for Windows Malware Detection. In Proceedings of the 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 6–8 July 2019; pp. 1–6.
99. Zhu, H.-J.; Jiang, T.-H.; Ma, B.; You, Z.-H.; Shi, W.-L.; Cheng, L.J.N.C. HEMD: A highly efficient random forest-based malware detection framework for Android. *Neural Comput. Appl.* **2018**, *30*, 3353–3361. [CrossRef]
100. Cheng, Y.; Fan, W.; Huang, W.; An, J. A Shellcode Detection Method Based on Full Native API Sequence and Support Vector Machine. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Sanya, China, 12–15 November 2019; p. 012124.
101. Mohaisen, A.; Alrawi, O. Unveiling zeus: Automated classification of malware samples. In Proceedings of the 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil, 13–17 May 2013; pp. 829–832.
102. Shijo, P.; Salim, A.J.P.C.S. Integrated static and dynamic analysis for malware detection. *Procedia Comput. Sci.* **2015**, *46*, 804–811. [CrossRef]
103. Karthika, R.; Visalakshi, P.J.W.T.C. A hybrid ACO based feature selection method for email spam classification. *WSEAS Trans. Comput.* **2015**, *14*, 171–177.
104. Jain, G.; Sharma, M.; Agarwal, B. Spam detection on social media using semantic convolutional neural network. *Int. J. Knowl. Discov. Bioinform.* **2018**, *8*, 12–26. [CrossRef]

105. Sagar, R.; Jhaveri, R.; Borrego, C.J.E. Applications in Security and Evasions in Machine Learning: A Survey. *Electronics* **2020**, *9*, 97. [CrossRef]
106. Quinlan, J.R. *C4. 5: Programs for Machine Learning*; Elsevier: Amsterdam, The Netherlands, 2014.
107. Mishra, P.; Varadharajan, V.; Tupakula, U.; Pilli, E.S. Tutorials. A detailed investigation and analysis of using machine learning techniques for intrusion detection. *IEEE Commun. Surv. Tutor.* **2018**, *21*, 686–728. [CrossRef]
108. Kavzoglu, T.; Colkesen, I. The effects of training set size for performance of support vector machines and decision trees. In Proceedings of the 10th international symposium on spatial accuracy assessment in natural resources and environmental sciences, Florianópolis, Brazil, 10–13 July 2012; p. 1013.
109. Salehi, Z.; Sami, A.; Ghiasi, M.J.C.F. Using feature generation from API calls for malware detection. *Security* **2014**, *2014*, 9–18. [CrossRef]
110. Jamil, Q.; Shah, M.A. Analysis of machine learning solutions to detect malware in android. In Proceedings of the 2016 Sixth International Conference on Innovative Computing Technology (INTECH), Dublin, Ireland, 24–26 August 2016; pp. 226–232.
111. Kevric, J.; Jukic, S.; Subasi, A.J.N.C. An effective combining classifier approach using tree algorithms for network intrusion detection. *Applications* **2017**, *28*, 1051–1058. [CrossRef]
112. Gaikwad, D.; Thool, R.C. Intrusion detection system using ripple down rule learner and genetic algorithm. *Int. J. Comput. Sci. Inf. Technol.* **2014**, *5*, 6976–6980.
113. Ingre, B.; Yadav, A.; Soni, A.K. Decision tree based intrusion detection system for NSL-KDD dataset. In Proceedings of the International Conference on Information and Communication Technology for Intelligent Systems, Ahmedabad, India, 15–16 May 2020; pp. 207–218.
114. Ahmim, A.; Maglaras, L.; Ferrag, M.A.; Derdour, M.; Janicke, H. A novel hierarchical intrusion detection system based on decision tree and rules-based models. In Proceedings of the 2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS), Santorini Island, Greece, 29–31 May 2019; pp. 228–233.
115. Relan, N.G.; Patil, D.R. Implementation of network intrusion detection system using variant of decision tree algorithm. In Proceedings of the 2015 International Conference on Nascent Technologies in the Engineering Field (ICNTE), Navi Mumbai, India, 9–10 January 2015; pp. 1–5.
116. Goeschel, K. Reducing false positives in intrusion detection systems using data-mining techniques utilizing support vector machines, decision trees, and naive Bayes for off-line analysis. In Proceedings of the SoutheastCon 2016, Norfolk, VA, USA, 30 March–3 April 2016; pp. 1–6.
117. Malik, A.J.; Khan, F.A. A hybrid technique using binary particle swarm optimization and decision tree pruning for network intrusion detection. *Clust. Comput.* **2018**, *21*, 667–680. [CrossRef]
118. Moon, D.; Im, H.; Kim, I.; Park, J.H. DTB-IDS: An intrusion detection system based on decision tree using behavior analysis for preventing APT attacks. *J. Supercomput.* **2017**, *73*, 2881–2895. [CrossRef]
119. Santos, I.; Brezo, F.; Ugarte-Pedrero, X.; Bringas, P.G. Opcode sequences as representation of executables for data-mining-based unknown malware detection. *Inf. Sci.* **2013**, *231*, 64–82. [CrossRef]
120. Islam, R.; Tian, R.; Batten, L.M.; Versteeg, S. Classification of malware based on integrated static and dynamic features. *J. Netw. Comput. Appl.* **2013**, *36*, 646–656. [CrossRef]
121. Yan, P.; Yan, Z. A survey on dynamic mobile malware detection. *Softw. Qual. J.* **2018**, *26*, 891–919. [CrossRef]
122. Saab, S.A.; Mitri, N.; Awad, M. Ham or spam? A comparative study for some content-based classification algorithms for email filtering. In Proceedings of the MELECON 2014-2014 17th IEEE Mediterranean Electrotechnical Conference, Beirut, Lebanon, 13–16 April 2014; pp. 339–343.
123. Zhang, Y.; Wang, S.; Phillips, P.; Ji, G. Binary PSO with mutation operator for feature selection using decision tree applied to spam detection. *Knowl. -Based Syst.* **2014**, *64*, 22–31. [CrossRef]
124. Sharma, S.; Arora, A. Adaptive approach for spam detection. *Int. J. Comput. Sci. Issues* **2013**, *10*, 23.
125. Alom, M.Z.; Bontupalli, V.; Taha, T.M. Intrusion detection using deep belief networks. In Proceedings of the 2015 National Aerospace and Electronics Conference (NAECON), Dayton, OH, USA, 15–19 June 2015; pp. 339–344.
126. Tyagi, A. *Content Based Spam Classification-A Deep Learning Approach*; University of Calgary: Calgary, AB, Canada, 2016.
127. He, S.; Lee, G.M.; Han, S.; Whinston, A.B. How would information disclosure influence organizations' outbound spam volume? Evidence from a field experiment. *J. Cybersecur.* **2016**, *2*, 99–118. [CrossRef]

128. Zhang, Y.; Li, P.; Wang, X.J.I.A. Intrusion detection for IoT based on improved genetic algorithm and deep belief network. *IEEE Access* **2019**, *7*, 31711–31722. [CrossRef]
129. Ye, Y.; Wang, D.; Li, T.; Ye, D.; Jiang, Q. An intelligent PE-malware detection system based on association mining. *J. Comput. Virol.* **2008**, *4*, 323–334. [CrossRef]
130. Alkaht, I.J.; Al-Khatib, B. Filtering SPAM Using Several Stages Neural Networks. *Int. Rev. Comp. Softw.* **2016**, *11*, 2. [CrossRef]
131. Yuan, Z.; Lu, Y.; Xue, Y. Droiddetector: Android malware characterization and detection using deep learning. *Tsinghua Sci. Technol.* **2016**, *21*, 114–123. [CrossRef]
132. Jo, S.; Sung, H.; Ahn, B. A comparative study on the performance of intrusion detection using decision tree and artificial neural network models. *J. Korea Soc. Digit. Ind. Inf. Manag.* **2015**, *11*, 33–45.
133. Ammar, A. A decision tree classifier for intrusion detection priority tagging. *J. Comput. Commun.* **2015**, *3*, 52. [CrossRef]
134. Li, Y.; Ma, R.; Jiao, R. A hybrid malicious code detection method based on deep learning. *J. Secur. Appl.* **2015**, *9*, 205–216. [CrossRef]
135. Rizk, Y.; Hajj, N.; Mitri, N.; Awad, M. Deep belief networks and cortical algorithms: A comparative study for supervised classification. *Appl. Comput. Inform.* **2019**, *15*, 81–93. [CrossRef]
136. Phan, T.D.; Zincir-Heywood, N. User identification via neural network based language models. *Int. J. Netw. Manag.* **2019**, *29*, e2049. [CrossRef]
137. Shrivas, A.K.; Dewangan, A.K. An ensemble model for classification of attacks with feature selection based on KDD99 and NSL-KDD data set. *Int. J. Comput. Appl.* **2014**, *99*, 8–13.
138. Shabtai, A.; Moskovitch, R.; Feher, C.; Dolev, S.; Elovici, Y.J.S.I. Detecting unknown malicious code by applying classification techniques on opcode patterns. *Secur. Inform.* **2012**, *1*, 1. [CrossRef]
139. Ahmad, I.; Abdullah, A.B.; Alghamdi, A.S. Artificial neural network approaches to intrusion detection: A review. In Proceedings of the 8th Wseas International Conference on Telecommunications and Informatics, Istanbul, Turkey, 30 May–1 June 2009.
140. Soranamageswari, M.; Meena, C. A novel approach towards image spam classification. *Int. J. Comput. Theory Eng.* **2011**, *3*, 84. [CrossRef]
141. Chen, Y.; Narayanan, A.; Pang, S.; Tao, B. Multiple sequence alignment and artificial neural networks for malicious software detection. In Proceedings of the 2012 8th International Conference on Natural Computation, Chongqing, China, 29–31 May 2012; pp. 261–265.
142. Arram, A.; Mousa, H.; Zainal, A. Spam detection using hybrid Artificial Neural Network and Genetic algorithm. In Proceedings of the 2013 13th International Conference on Intellient Systems Design and Applications, Salangor, Malaysia, 8–10 December 2013; pp. 336–340.
143. Qureshi, A.-U.-H.; Larijani, H.; Mtetwa, N.; Javed, A.; Ahmad, J.J.C. RNN-ABC: A New Swarm Optimization Based Technique for Anomaly Detection. *Computers* **2019**, *8*, 59. [CrossRef]
144. Sheikhan, M.; Jadidi, Z.; Farrokhi, A. Intrusion detection using reduced-size RNN based on feature grouping. *Neural Comput. Appl.* **2012**, *21*, 1185–1190. [CrossRef]
145. Liangboonprakong, C.; Sornil, O. Classification of malware families based on n-grams sequential pattern features. In Proceedings of the 2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA), Melbourne, Australia, 19–21 June 2013; pp. 777–782.
146. Hardy, W.; Chen, L.; Hou, S.; Ye, Y.; Li, X. DL4MD: A deep learning framework for intelligent malware detection. In Proceedings of the International Conference on Data Mining (DMIN), Las Vegas, NV, USA, 12–15 July 2010; p. 61.
147. Foqaha, M.A.M. Email spam classification using hybrid approach of RBF neural network and particle swarm optimization. *Int. J. Netw. Secur. Appl.* **2016**, *8*, 17–28.
148. Bassiouni, M.; Ali, M.; El-Dahshan, E.A. Ham and Spam E-Mails Classification Using Machine Learning Techniques. *J. Appl. Secur. Res.* **2018**, *13*, 315–331. [CrossRef]
149. Gao, Y.; Wu, H.; Song, B.; Jin, Y.; Luo, X.; Zeng, X.J.I.A. A Distributed Network Intrusion Detection System for Distributed Denial of Service Attacks in Vehicular Ad Hoc Network. *IEEE Access* **2019**, *7*, 154560–154571. [CrossRef]
150. Siddiqui, M.; Wang, M.C.; Lee, J. Detecting internet worms using data mining techniques. *J. Syst. Cybern. Inform.* **2009**, *6*, 48–53.

151. Rathi, M.; Pareek, V. Spam mail detection through data mining-A comparative performance analysis. *Int. J. Mod. Educ. Comput. Sci.* **2013**, *5*, 31. [CrossRef]
152. Zhou, Y.-Y.; Cheng, G. An Efficient Network Intrusion Detection System Based on Feature Selection and Ensemble Classifier. *arXiv* **2019**, arXiv:1904.01352.
153. Xu, H.; Sun, W.; Javaid, A. Efficient spam detection across online social networks. In Proceedings of the 2016 IEEE International Conference on Big Data Analysis (ICBDA), Hangzhou, China, 12–14 March 2016; pp. 1–6.
154. Gupta, G.P.; Kulariya, M. A framework for fast and efficient cyber security network intrusion detection using apache spark. *Procedia Comput. Sci.* **2016**, *93*, 824–831. [CrossRef]
155. Vinayakumar, R.; Alazab, M.; Soman, K.; Poornachandran, P.; Al-Nemrat, A.; Venkatraman, S.J.I.A. Deep Learning Approach for Intelligent Intrusion Detection System. *IEEE Access* **2019**, *7*, 41525–41550. [CrossRef]
156. Prakash Chandra, P.U.K.; Lilhore, P.N.A. Network intrusion detection system based on modified Random forest classifiers for kdd cup-99 and nsl-kdd Dataset. *Int. Res. J. Eng. Technol.* **2017**, *4*, 786–791.
157. Vivek Nandan Tiwari, P.S.R. Enhanced Method for Intrusion Detection over KDD Cup 99 Dataset. *Int. J. Curr. Trends Eng. Technol.* **2016**, *2*, 218–224.
158. Galal, H.S.; Mahdy, Y.B.; Atiea, M.A. Behavior-based features model for malware detection. *J. Comput. Virol. Hacking Tech.* **2016**, *12*, 59–67. [CrossRef]
159. Mosli, R.; Li, R.; Yuan, B.; Pan, Y. A behavior-based approach for malware detection. In Proceedings of the IFIP International Conference on Digital Forensics, Orlando, FL, USA, 30 January–1 February 2017.
160. Lee, S.M.; Kim, D.S.; Kim, J.H.; Park, J.S. Spam detection using feature selection and parameters optimization. In Proceedings of the 2010 International Conference on Complex, Intelligent and Software Intensive Systems, Krakow, Poland, 15–18 February 2010; pp. 883–888.
161. Mccord, M.; Chuah, M. Spam detection on twitter using traditional classifiers. In Proceedings of the International Conference on Autonomic and Trusted Computing, Banff, AB, Canada, 2–4 September 2011; pp. 175–186.
162. Jiang, L.; Zhang, H.; Cai, Z. A novel Bayes model: Hidden naive Bayes. *IEEE Trans. Knowl. Data Eng.* **2008**, *21*, 1361–1371. [CrossRef]
163. Panda, M.; Patra, M.R. Network intrusion detection using naive bayes. *Int. J. Comput. Sci. Netw. Secur.* **2007**, *7*, 258–263.
164. Fan, C.-I.; Hsiao, H.-W.; Chou, C.-H.; Tseng, Y.-F. Malware detection systems based on API log data mining. In Proceedings of the 2015 IEEE 39th Annual Computer Software and Applications Conference, Taichung, Taiwan, 1–5 July 2015; pp. 255–260.
165. Sharma, S.K.; Pandey, P.; Tiwari, S.K.; Sisodia, M.S. An improved network intrusion detection technique based on k-means clustering via Naïve bayes classification. In Proceedings of the IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM-2012), Nagapattinam, India, 30–31 March 2012; pp. 417–422.
166. Jackson, T.R.; Levine, J.G.; Grizzard, J.B.; Owen, H.L. An investigation of a compromised host on a honeynet being used to increase the security of a large enterprise network. In Proceedings of the Fifth Annual IEEE SMC Information Assurance Workshop, West Point, NY, USA, 10–11 June 2004; pp. 9–14.
167. Khammas, B.M.; Monemi, A.; Bassi, J.S.; Ismail, I.; Nor, S.M.; Marsono, M.N. Feature selection and machine learning classification for malware detection. *J. Teknol.* **2015**, *77*, 234–250. [CrossRef]
168. Bhat, A.H.; Patra, S.; Jena, D. Machine learning approach for intrusion detection on cloud virtual machines. *Int. J. Appl. Innov. Eng. Manag.* **2013**, *2*, 56–66.
169. Gharibian, F.; Ghorbani, A.A. Comparative study of supervised machine learning techniques for intrusion detection. In Proceedings of the Fifth Annual Conference on Communication Networks and Services Research (CNSR'07), Frederlcton, NB, Canada, 14–17 May 2007; pp. 350–358.
170. Renuka, D.K.; Visalakshi, P.; Sankar, T.J.I.J.C.A. Improving E-mail spam classification using ant colony optimization algorithm. *Int. J. Comput. Appl.* **2015**, *2*, 22–26.