Journal of Computer Science

Original Research Paper

Cyber Security Threats and Countermeasures using Machine and Deep Learning Approaches: A Survey

Manjula M, Venkatesh and Venugopal K. R.

Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bengaluru, India

Article history
Received: 16-09-2022
Revised: 17-11-2022
Accepted: 29-11-2022

Corresponding Author:
Manjula M
Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bengaluru, India
Email: manjula.m82@gmail.com

Abstract: Recent advancements in e-business, e-healthcare, e-governance, and online digital transactions have brought valuable benefits. Unfortunately, it raises severe cyber-attacks. Cyberattacks disrupt normal operations, try to retrieve confidential information and defense secrets, and subvert the nation's defense systems and Internet-connected devices. Cyber security solutions are required to detect, analyze, defend against threats and protect sensitive data from unauthorized access. This study gives a detailed survey of different cybersecurity attacks, like Denial-of-service attacks, Botnet Evasion Attacks, Malware invasions, Spam and phishing invasion, Spoofing, Domain Generation algorithms, Probing attacks, R2L, and U2R attacks. This research review emphasizes Machine Learning and Deep Learning-based approaches to Cybersecurity problems. This study's key highlights are the research challenges, cybersecurity issues, cyber security domains, and tools for the Intrusion detection system. Data sets play a vital role in cybersecurity research; hence, Private and Publically available datasets are reviewed in this study. Various performance matrices are discussed in this survey which can be used to evaluate the effectiveness of cybersecurity solutions.

Keywords: Cyber Security Threats, Cyber Security, Machine Learning (ML), Deep Feature Learning (DL), Botnet Attacks, Malware Attacks, Evasion Attack, Adversarial Machine Learning Algorithms

Introduction

The recent advancements in cloud computing, wireless communication, big data, social network, the Internet of Things, and the availability of high-speed internet has enabled rapid growth in cyberspace. Cyberspace is a global domain combining computer communication systems with technology (Starodubtsev *et al*., 2020). Although advancements in technologies bring valuable benefits, unfortunately, it raises stringent cyber threats as well. Cyber security can be coined as "Information Technology Security". With the massive growth in Internet-connected devices, securing these devices is a high need. The process of protecting hosts in networks, applications, computing devices, and data from adversaries' attacks is referred to as *'Cyber security'*. This involves the collection of policies, processes, techniques, and technologies to prevent vulnerabilities/attacks (Berman *et al*., 2019). Generally, cyber security systems protect user data and devices via encryption, user authentication, anti-virus software, Intrusion Detection Systems (IDS), and firewalls. Under different contexts, cyber security is termed network security, disaster recovery, end-user training, operational security, and information security.

A cyberattack is a conscious endeavor by an individual or organization to gain the confidential data of individuals or organizations. In a cyber-attack, the attacker introduces an attack on the network, disabling applications, malfunctioning devices to interrupt routine services, and stealing confidential information. The attacker could make the Internet-connected devices malfunction. Cyber-attacks exploit vulnerabilities in software and hardware design through malware. Denial of Distributed Service attacks is introduced to overwhelm the target websites. Through hacking, attackers pierce the defense of potential computer systems and integrity with their functionality. The attacker tries to retrieve confidential information from the organization; the retrieved information can be shared with an unauthorized person for financial or business benefits. Cyber-attacks interrupt the normal operations of software applications. Inexperienced or untrained personnel, improper systems configuration, and

insufficient procedures increase computer network systems' vulnerabilities. The COVID-19 pandemic also has raised many people to adopt technology, thus exposing people to cyber vulnerabilities. During the COVID-19 pandemic, many cyber security firms identified a drastic increase of 35% in cyber-attacks (Andrade *et al*., 2020).

There is a need for novel and efficient methods in cyber security as there is an emergence of new smart network technologies (Duić *et al*., 2017). These systems should be protected from digital attack, damage, or unauthorized access. Efficient Cyber security solutions are required in various domains such as business applications, online transactions, cloud computing, mobile computing, software solutions, etc. Cyber Security solution is very much required as it encloses protecting sensitive data from unauthorized access. There is a requirement to employ threat intelligence and Machine Learning approaches to identify, analyze and defend against cyber security risks in real time. In recent years, emerging technologies that can be used in the detection of cyber-attacks are Dynamic Networks, Predictive Semantics, Quantum Computing, Behavioral Identity, Cloud Computing, etc., (Geluvaraj *et al*., 2019). This study is comprehensive about the current research work in the field of Cybersecurity using the ML and DL approaches.

Research Challenges and Issues in Cyber Security

Cybersecurity is a dynamic area where challenges will always proliferate and professionals or individuals must be ready to face these challenges. Figure 1 gives the major sectors affected by Cyber-attacks in the year 2021 (Gmcdouga, 2022). E-business, online transactions are more vulnerable to cyber threats. It results in the loss of confidential information, reputation damage and even being liable for legal issues. Cyber threats can be of any one form based on the motives of the attacker, such as *Cybercrime, Cyberwarfare, Cyberterrorism*, and *Cyber Espionage* (Rohith and Batth, 2019). Cybercrime leads to various criminal activities causing significant financial losses to businesses and individuals. Because of Cyber Espionage, a massive amount of data, sensitive information, and intellectual property are extracted from government and private sector websites for economic benefit or political reasons. Statistics revealed that 11% of cyber-attacks are because of espionage.

The Aerospace and defense sectors face cyber threats with intentions of stealing intellectual property and defense secrets. In Cyber warfare, Cyber attackers monitor, infiltrate and subvert other nations' defense to disrupt their critical infrastructure (Duić *et al*., 2017). Cyberattacks on defense have cascading effects and breach the national security system. Cyber-attacks are launched covertly to weaken or strike at an adversary to achieve political objectives. The enemy is unseen and the victim is unsure how and where to react. The attacker does not leave any proof of their involvement in these attacks. The attackers are called *non-state attackers* (Goel and Nussbaum, 2021). To name a few, non-state actors include criminal organizations, script kiddies, hacktivists, scammers, and black hat hackers. The future war will be *'Cyberterrorism'* or no contact war wherein there is no 'physical' or 'kinetic' action across the borders, which is constantly increasing. In Cyberterrorism, cyberspace is deliberately used for devising terrorist attacks. Recently terrorists are using cyberspace for their communication, to have command and control, to brainwash innocent people, and for training and funding goals. Providing Cyber security in the defense system is a complex issue that requires multidimensional, multilayered initiatives and responses. According to Professor Dorothy E. Denning, the definition of cyber terrorism is: Unlawful invasion and imminence of attack against computing nodes, networks, and critical data when done to bogart or compel a government or public servants or elected people emphasizing socio-political goals (Luiijf *et al*., 2013). The other terms for cyber terrorism are cyber jihad, electronic jihad, e-jihad, and Internet jihad (Malin *et al*., 2017).

Satellite communication systems, navigation systems, and Earth observation systems often pose threats from cyber-attacks (Caprolu *et al*., 2020). Cyber attackers can use software mechanisms, amplifiers, transmitters, and steerable antennas to interfere with or generate satellite signals. The vulnerabilities in the satellite communication systems are mission-critical as they can disturb launch systems, telemetry, tracking, and command and communications. Continuous monitoring and protection measures have to be taken to protect these space-based systems.

A myriad of cyber threats plays the health sector. Data privacy in healthcare is more concern in many countries. Cyber threats to the health sector may arise from Malware that compromises the virtue of the system or from DDOS attacks by losing patients' privacy or disrupting the facilities available to patients. Cyber threats in the health sector have ramifications beyond financial loss and breach of privacy (Thamer and Alubady, 2021). For instance, 'Ransomware' malware for hospitals steals patient data and puts patient lives at risk. It is reported that more than 18 million patient data are afflicted by ransomware attacks. This sector is more prone to cyber-attacks since the patient's personal information and medical data were compromised by Cyber attackers. The illness data can be used to blackmail the confidential information of patients such as results of diagnosis, severity, types of treatment, and diseases shared with the marketing firms for advertising and recommending their products. In the Electronic Healthcare Record (EHR) system, the patient records are maintained and the medical

devices such as infusion pumps, remote e-patient observation equipment, and Ventilation and Air Conditioning (HVAC) systems are connected to EHR systems.

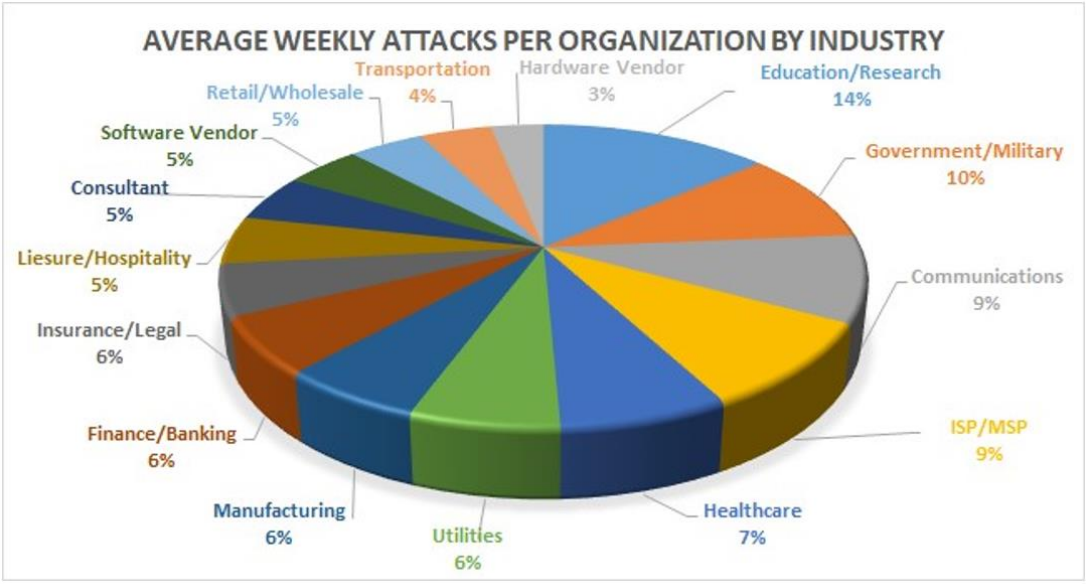

Fig. 1: Average weekly attacks per organization by Industry (2021)

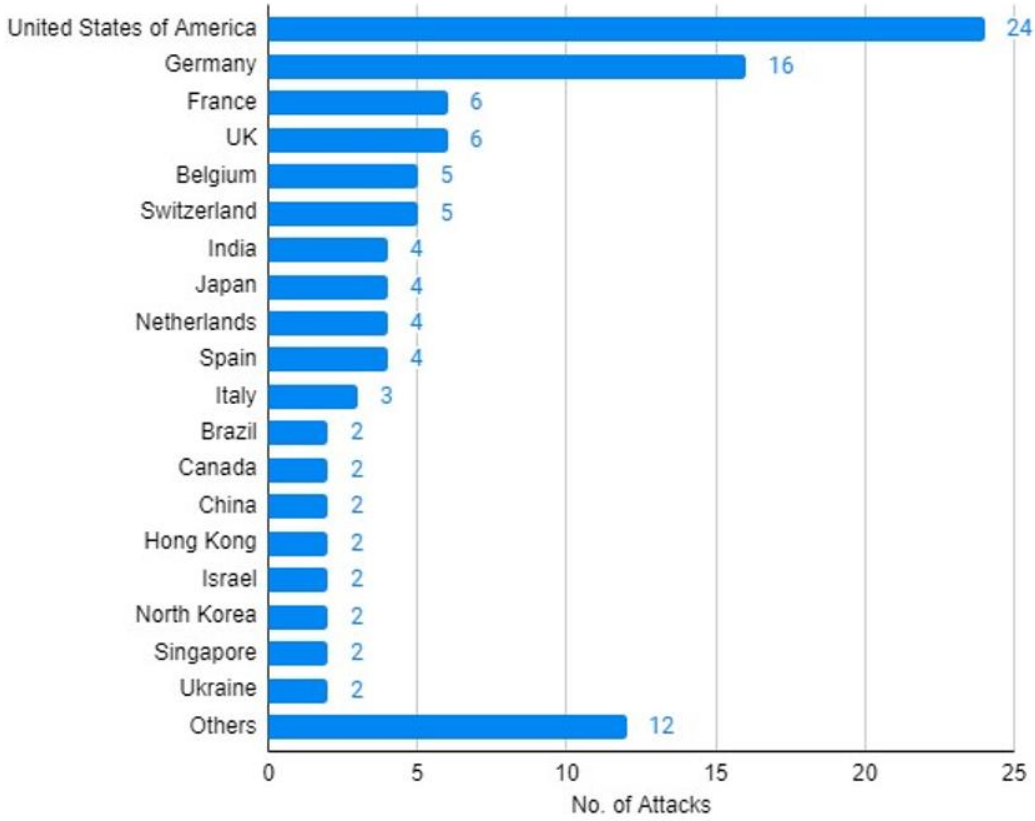

Fig. 2: Countries affected by major cyber-attacks in January 2022

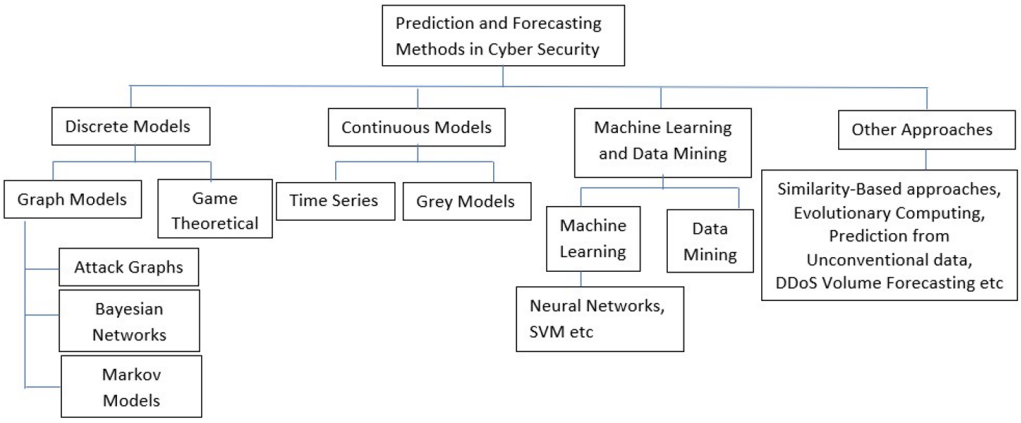

Fig. 3: Different approaches for cyber security

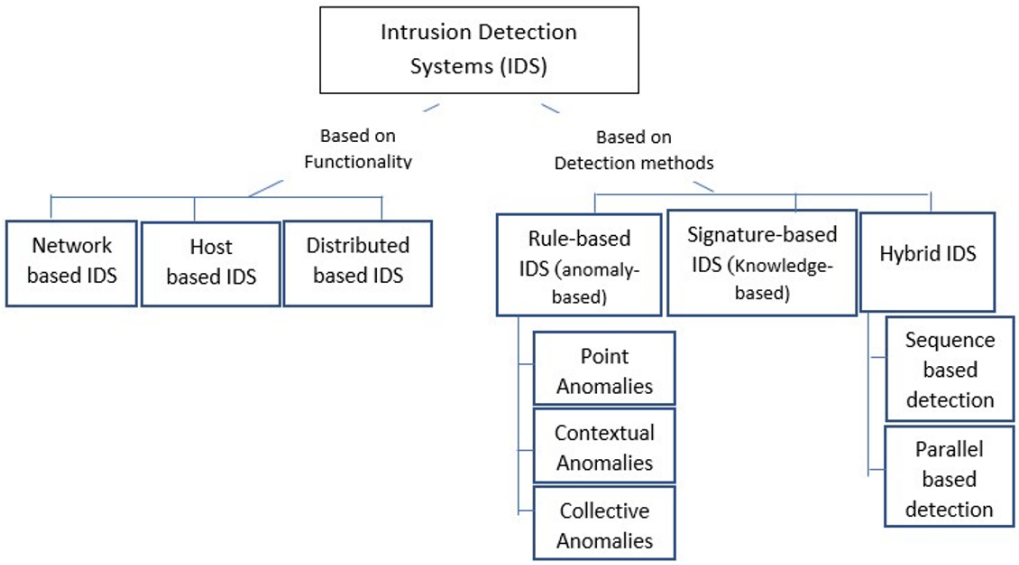

Fig. 4: Intrusion Detection System (IDS)

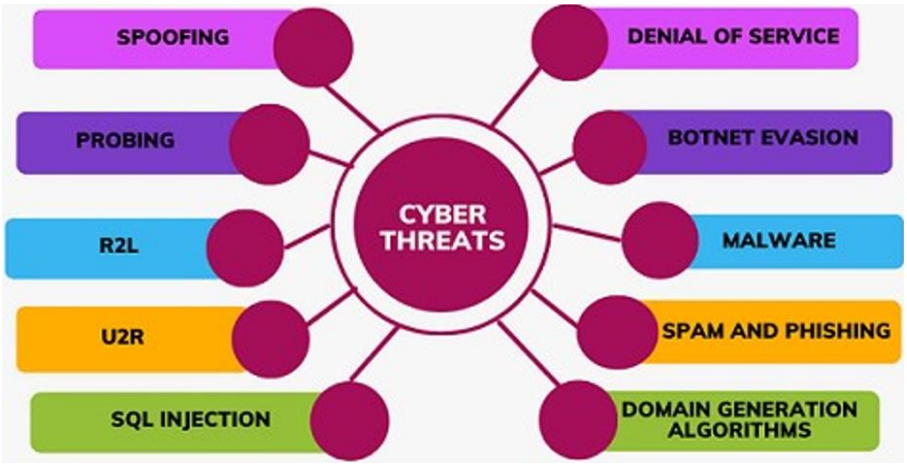

Fig. 5: Types of cyber security threats

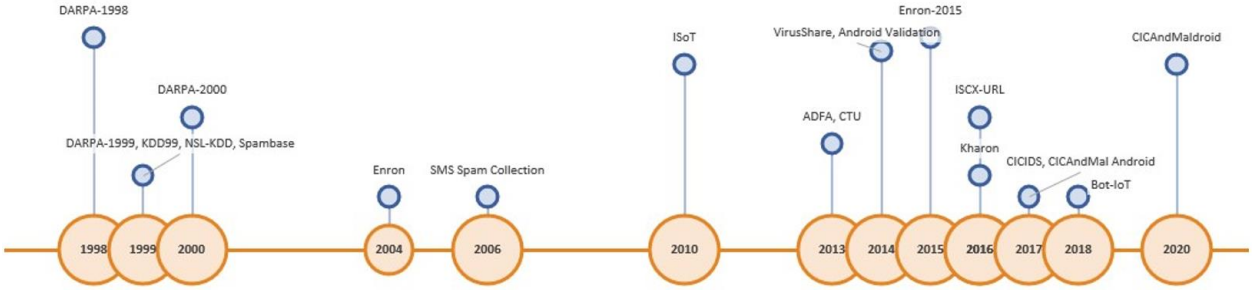

Fig. 6: Evolution of frequently used security datasets

Table 1: NSL-KDD dataset KDD99/ KDD CUP 99 dataset features

No.	Features	Type	Feature category
1	Duration	Integer	Basic features
2	Protocol_type	Nominal	Basic features
3	Service	Nominal	Basic features
4	Src_bytes	Integer	Basic features
5	Dst_bytes	Integer	Basic features
6	Flag	Nominal	Basic features
7	Land	Binary	Basic features
8	Wrong_fragment	Integer	Basic features
9	Urgent	Integer	Basic features
10	Hot	Integer	Content features
11	Num_failed_logins	Integer	Content features
12	Logged_in	Binary	Content features
13	Num_compromised	Integer	Content features
14	Root_shell	Binary	Content features
15	Su_attempted	Binary	Content features
16	Num_root	Integer	Content features
17	Num_file_creations	Integer	Content features
18	Num_shells	Integer	Content features
19	Num_access_files	Integer	Content features
20	Num_outbound_cmds	Integer	Content features
21	Is_hot_login	Binary	Content features
22	Is_guest_login	Binary	Content features
23	Count	Integer	Time traffic features
24	Serror_rate	Real	Time traffic features
25	Rerror_rate	Real	Time traffic features
26	Same_srv_rate	Real	Time traffic features
27	Diff_srv_rate	Real	Time traffic features
28	Srv_count	Integer	Time traffic features
29	Srv_serror_rate	Real	Time traffic features
30	Srv_rerror_rate	Real	Time traffic features
31	Srv_diff_host_rate	Real	Time traffic features
32	Dst_host_count	Integer	Machine traffic features
33	Dst_host_srv_count	Integer	Machine traffic features
34	Dst_host_same_srv_rate	Real	Machine traffic features
35	Dst_host_diff_srv_rate	Real	Machine traffic features
36	Dst_host_same_src_port_rate	Real	Machine traffic features
37	Dst_host_srv_diff_host_rate	Real	Machine traffic features
38	Dst_host_serror_rate	Real	Machine traffic features
39	Dst_host_srv_serror_rate	Real	Machine traffic features
40	Dst_host_rerror_rate r	Real	Machine traffic features
41	Dst_host_srv_error_rate	Real	Machine traffic features

Table 2: Different attack class with relevant features of KDD CUP99 data set

Attack class	Attack type	The best set of relevant features
Dos	neptune, teardrop, back, land, pod, smurf	2, 3, 9, 26, 41,4, 26,27,41
U2R	loadmodule, rootkit, buffer_overflow, perl	6, 11, 29, 30, 3, 10, 14
R2L	imap, ftp_write, multihop, guess_passwd, phf, warezclient, warezmaster, spy	1,2,7,33,3,40,34,30,21
Probe	ipsw eep, nmap, portsweep, satan	2,3, 9, 30, 34, 38, 40

Table 3: CICIDS 2017 dataset

Day	Date	Description	Attack type	Attack sub-type	Number of instances
Monday	July 3, 2017	Normal Activity	Benign (Normal human activities)	NA	2359087
Tuesday	July 4, 2017	Attacks + Normal Activity	Benign, Brute Force	FTP-Patator	7938
				SSH-Patator	5897
Wednesday	July 5, 2017	Attacks + Normal Activity	Benign,	DoS slowloris	5796
			DoS / DdoS	DoS Slowhttptest	5499
				DoS Hulk	231072
				DoS Golden Eye	10293
				Heartbleed	11
Thursday	July 6, 2017	Attacks + Normal Activity	Benign,	Web Attack - Brute Force	1507
			Web Attack	Web Attack – XSS	652
				Web Attack - Sql Injection	21
			Infiltration	Dropbox	36
				Cool disk - MAC	
Friday	July 7, 2017	Attacks + Normal Activity	Benign,	Botnet ARES	1966
			Botnet ARES	DDoS LOIT	41835
			DDoS LOIT Port Scan	Port Scan	158930

Table 4: ISCX URL-2016 dataset

Sl No.	URL type	Sources	Number of samples
1	Benign	Alexa top websites	35,300
2	Spam	Publicly available WEBSPAM-UK2007 dataset	12,000
3	Phishing	Open-Phish	10,000
4	Malware	DNS-BH	11,500
5	Defacement	Alexa ranked trusted websites	45,450

Table 5: Four well-known Malware families and their top 5 features of DREBIN dataset

SI No.	Malware family	Top 5 features
1	Fake installer	Send SMS
		android.hardware.telephony
		SEND_SMS
		READ_PHONE_STATE
		Send Text Message
2	Droid kungfu	getSubscriberId
		SIG_STR
		READ_PHONE_STATE
		BATTERY_CHANGED_ACTION
		system/bin/su
3	Gold dream	lebar.gicp.net
		sendSMS
		DELETE_PACKAGES
		getSubscriberId
		android.provider.Telephony.SMS_RECEIVED
4	Ginger master	READ_PHONE_STATE
		getSubscriberId
		USER_PRESENT
		system/bin/su
		Http Post

Table 6: The CTU-13 botnet dataset scenario

Scenario Id	Botnet name	Number of infected nodes	Number of packets	Botnet type
1	Neris	1	71,971,482	IRC botnet
2	Neris	1	71,851,300	IRC botnet
3	Rbot	1	167,730,395	IRC botnet
4	Rbot	1	62,089,135	IRC botnet
5	Virut	1	4,481,167	HTTP botnets
6	Menti	1	38,764,357	IRC botnet
7	Sogou	1	7,467,139	HTTP botnets
8	Murlo	1	155,207,799	IRC botnet
9	Neris	10	115,415,321	IRC botnet
10	Rbot	10	90,389,782	IRC botnet
11	Rbot	3	6,337,202	IRC botnet
12	NSIS.ay	3	13,212,268	P2P botnets
13	Virut	1	50,888,256	HTTP botnets

Table 7: A bird view on state-of-art machine learning techniques for cyber security

Reference	Year	Attack	Scenario	ML model	Dataset	Performance results
Kilincer *et al.*	2021	DoS, DDoS, Brute Force, Exploit, SQL Injection	Intrusion Detection Systems	SVM, KNN, DT	NSL-KDD, Kyoto, DARPA, KDD 99, CIDDS-001, ISCX-2012, AWID, UNSW-NB15, CSE-CIC-IDS (2017, 2018)	-
Taheri *et al.*	2020	KNN-based, LR, Trivial, Distribution, Antcolony based	Android malware detection	AML Random Forest, Bagging, and SVM.	Drebin dataset, Genome dataset, Contagio dataset	F-score-90-95% Accuracy 90.55%
Zhida Li *et al.*	2019	Network Anomalies and Intrusions	Classifying Network Anomalies and Intrusions	Broad Learning-System (BLS), LSTM, GRU, RNNs	BGP Datasets, NSL-KDD	-
Baptista *et al.*	2019	Malware attack	Malware Detection System	Self-Organizing Incremental Neural Networks (SOINN)	Virus Share	Accuracy 74%
Guo *et al.*	2021	Black-Box Attack	Adversarial Example	KNN, MLP, CNN, and ResNet	KDD99 dataset, CSE-CIC-IDS2018	Recall 91.8% Accuracy 98%
Li *et al.*	2018	Malware 5494 benign-apps,	Malware detection system	SVM	Randomly Generated from Google play -store data	Accuracy 96.47% Detection rate 94.62%
Xin *et al.*	2018	Intrusion Detection	NIDS for IoT Security	KNN, SVM, and DT	DARPA Intrusion Detection Data Sets; KDD CUP 99, NSL-KDD, ADFA.	Accuracy 99.41%
Chaabouni *et al.*	2019	Intrusion Detection	NIDS for IoT Security	MLP, SVM, ANN	KDDCUP99, UNSW-NB15, DEFCON, ADFA IDS, PREDICT, KYOTO, CAIDA, ISCX 2012 and ICS	Accuracy 99%
Li *et al.*	2018	Malware	DGA-Based Malware Detection	DT-J48, ANN, SVM, LR, NB, GBT, RF	CryptoLocker, Tovar, Dyre,Nymaim,Locky	Accuracy 95.89%
Zhang *et al.*	2020	GAN attack	A Brute-Force Black-Box Method	LR, DT, MLP, NB and RF	NSL-KDD, DREBIN	-
Pu *et al.*	2021	DoS, Probe, U2R, and R2L	Hybrid Unsupervised Clustering-Based	SSC, OCSVM	NSL-KDD	-
Seth *et al.*	2021	Bot, Brute Force, DDoS, DoS, Infiltration, Web Attacks	Intelligent Intrusion Detection System	KNN, DT, RF, Extra Trees, XGB, HBGB, Light GBM	CSE-CIC-IDS2018	Accuracy 96.97%, recall rate of 97.4%,
Alam *et al.*	2020	Phishing Attacks	Phishing Attacks on Websites	PCA, RF, DT	Phishing Attacks from Kaggle	Accuracy 97%
Giovanni *et al.*	2018	Botnets, DGA, Malware, Spam, Phishing	Intrusion (Spam, Phishing) Detection, Malware Analysis,	RF, Feedforward Fully Connected Deep Neural Network.	DREBIN	-
Feng *et al.*	2018	DoS, DDoS	Distributed Cyber Attack using C&C communication	PCA, RF, SVM	CCC datasets consisting C10, C08, C09 and C13 datasets	-
Gupta *et al.*	2021	Phishing Attacks	Lexical based ML in real-time environment	RF, KNN, SVM, LR	ISCXURL-2016	Accuracy 99.57%
Hegde *et al.*	2020	Botnet	Botnet Activity in IoT Network Traffic	RF, DT, two class NNs, Multiclass DT, and Multi class NNs.	Stratosphere Lab IoT-23 data	Accuracy 99.9%
Lakshmanarao *et al.*	2021	Phishing Attacks	Phishing Attacks on Websites	LR, SVM, KNN, DT, RF, AdaBoost, Gradient Boosting	UCI	Accuracy 97%

Table 7: Continue

Singhal *et al*.	2020	Phishing attacks	Malicious website Detection	Gradient boosted trees, RF and Feedforward Neural Networks	Phish tank, Malware Domain List (MDL)	Accuracy 96.4%
Xiujuan *et al*.	2019	Phishing attacks	Spear-phishing Emails Detection Based on Authentication	SVM, KNN, RF	Enron Email Dataset	Accuracy 95.05%
Chaudhary *et al*.	2020	Probe, DoS, U2R, and R2L	Intrusion Detection Malware Analysis	RBF-SVM, Seq2Seq	KDD Cup 99, ADFA	Accuracy 99.90%
Begli *et al*.	2019	DoS, U2R	IDS for Critical Infrastructure	SVM	NSL-KDD	Detection rate 95.01%
Hagos *et al*.	2017	U2R, R2L, DoS, Probe	Network Intrusion Detection	SVM	NSL-KDD	Accuracy 97%
Krishnaveni *et al*.	2021	DoS, Probe, U2R and R2L	Network Intrusion Detection on Cloud Computing	SVM, NB, LR, and DT	NSL-KDD, Honeypot Kyoto 2006 +	Accuracy 96.06%
Wang *et al*.	2018	U2R, R2L, DoS, Probe	Intrusion Detection	KNN, K-Means Method	KDD	Precision 87.99%

Table 8: Three major categories of DL techniques

Model type	Features	Examples
Discriminative (supervised)	Labeled/self-annotations	Recurrent Neural Networks (RNN), Convolutional Neural Network (CNN), Multiple-Layer Perceptrons (MLP), Deep Stacking Network (DSN)
Generative (unsupervised)	Unlabeled data	Generative Adversarial Network (GAN), Auto Encoder (SAE, DAE,SDAE), Boltzmann Machine (DBM, RBM, DBN), Sum-Product Network (SPN)
Hybrid	Combines both Generative and Discriminative architectures	Deep Neural Network (DNN), Deep Transfer Learning (DTL), Deep Reinforcement Learning (DRL)

Table 9: A bird view on state-of-art deep learning techniques for cyber security

Reference	Year	Threat name	Scenario	DL model	Dataset	Performance results %
Lee *et al*.	2017	Intrusion detection	Event Profiling method for applying Artificial Intelligence Techniques	FCNN, CNN and LSTM	NSLKDD, CICIDS2017	Accuracy 98.00
Ferrag *et al*.	2020	Brute Force, DoS, DDOS, Botnet attack, SQL Injection	Approaches for cyber security Intrusion Detection	RBM, DNN, CNN, RNN, DBM, DBN, DA	Bot-IoT; CSE-CIC-IDS2018	Accuracy 98.39
Andresini *et al*.	2020	Zero Day Attack	Intrusion Detection	MINDFUL, NN, ANN, CNN and ACNN	KDDCUP991, UNSW-NB15, CICIDS 2017 AWID KDD 99	Accuracy 97.90 F-score- 94.93
Kim *et al*.	2017	Impersonation Detection	Intrusion Detection	SAE	AWID	-
Otoum *et al*.	2019	Intrusion detecting	Critical networks applications based on WSN	RBM, Restricted Boltzmann-based Clustered IDS (RBC-IDS)	KDD 99	Accuracy 99.91
Alom *et al*.	2015	Intrusion detection	Intrusion detection in network traffic	DBN	NSL-KDD	Accuracy 97.50
Sohn *et al*.	2020	Intrusion detection	IDS models based on DBN	DBN, RBM	UNSW-NB15, ADFA, NSL-KDD and KDD Cup 99,	-
Malaiya *et al*. (2018)	2019	Malicious activities	Anomaly Detection in the Network	VAE, FCN, LSTM, Seq2Seq	NSL-KDD, MAWILab traces, UNSWNB15, IDS2017, Kyoto-Honeypot	Accuracy 99.00
Maimo *et al*.	2018	Intrusion/Anomaly Detection system	Anomaly Detection in 5G Networks	DBN, SAE, SVM, LSTM	CTU	Accuracy 99.90 Recall 99.340
Alabadi *et al*.	2020	Anomaly Detection	Anomaly Detection	CNN	UNSW-NB15, IDS2017, NSL-KDD, KDD99, CICIDS2017	-
Hwang *et al*.	2020	DDoS (UDP, ACK, SYN, and HTTP floods)	Early Detection of Anomaly in Network Traffic	CNN	Mirai-CCU, USTC-TFC2016, Mirai-RGU	Accuracy 99.77 Precision 99.93 Recall 99.17
Apruzzese *et al*.	2020	Botnet	Botnet Evasion Attacks	DRL (RF, WnD)	CTU, BOTNET	Precision 99.80 Recall 99.80
Yu *et al*.	2019	Malwares with DGA	Domain Generation Algorithms (DGAs)	RNNs, LSTMs, CNNs and hybrid CNN/RNN	AlexaBamb	Accuracy 98.95
Shahzad *et al*.	2021	Malwares with DGA	DGA Domain Detection	RNN (LSTMs, BiLSTMs, GRUs)	Alexa Top 1M domains, Cisco umbrella popularity list, OSINT,	Accuracy 87.00 TPR of 81.00 Netlab 360.00
Vinayakumar *et al*.	2019	Malware	Metamorphism and polymorphism as	CNN, LSTM, DNN	Malimg, Privately collected samples from	Accuracy 96.30

Table 9: Continue

			obfuscation techniques for detecting signature -based attacks		Virus sign and Virus share	
Xu *et al.*	2021	Anomaly detection system	Network Anomaly Detection	Autoencoder	NSL-KDD	Accuracy 90.61 F1-score 92.26
Hindy *et al.*	2020	zero-day attacks	Intrusion Detection Systems (IDS)	Autoencoder, SVM	CICIDS2017 NSL-KDD	Accuracy 99.67
Djellali *et al.*	2019	Intrusion detection	Pattern Recognition	MLP	NSL-KDD	Accuracy 98.17
Mahdavifar *et al.*	2020	Malware detection	Semi-Supervised Deep Learning	Feed-Forward Neural Network (FFNN)	CICMalDroid2020	Accuracy 96.70 F1-Score 97.84 Precision 99.16
Li *et al.*	2020	Malware detection	Malware classification using Fuzzy decision trees and SVM	Gradient based adversarial methods	Virus Share	Accuracy 99.80
Li *et al.*	2020	Malware detection	Gradient free attacks, Transfer attacks, gradient based attacks, obfuscation, mixture of attacks	Adversarial deep ensemble, DNN	Drebin, and Androzoo	Accuracy 99.14

The cyber attacker can compromise the EHR system, and the devices connected to the EHR system and can introduce cyber-attacks (de la Torre *et al*., 2017).

The cyberattacks in IoT (Internet-of-Things), wherein devices connected to the network, are more susceptible when adversaries try to capture the IP address, application port, DNS server, and server IP address. IoT devices are sensitive to cyber-attacks as most IoT nodes are constantly attached and share data over the Internet (Roopak *et al*., 2019). The current development of portable and IoT devices has further amplified the consequence of malware attacks (He *et al*., 2021). As a result, the risks are exponentially more significant for IoT devices. Safeguarding the IoT device is complicated by the scale and scope of data being generated and collected. Software piracy and malware attacks put organizations and specific operational capabilities at risk. These attacks on the IoT with the growth in ubiquitous devices, the number of security threats increased (Ullah *et al*., 2019; Le Jeune *et al*., 2021). Cyber security algorithms assist in defending hosts, can defend against these applications and data, and recover from failure in a controlled, measurable way.

Although there are many cyber-attack detection mechanisms available, the rapid enhancement in hacking skills and the increase in the number of cyberattacks demands new cyber-attack detection systems (Duić *et al*., 2017). In recent days, all digital transactions such as online e-business transactions, online banking transactions, online stock market transactions, and online patient record maintenance. All these online transactions are prone to cyberattacks wherein the attacker interprets and eavesdrops on vital information related to the business transaction. Similarly, in online bank transactions, adversaries capture user login credentials, loan credentials, etc., to know the financial status of customers. It reuses login credentials, and financial status to introduce security threats.

Cyber security attacks on the power grid of a country are considered as vulnerable because vital information on nuclear reactor design and operations of reactors could be shared with other countries. The cyberattack on the power grid could result in the interruption of power and abruptly terminate the reactor functionality (Farooq *et al*., 2021). The primary threat to the Nuclear Power Plant is Cyber Sabotage. Cyber Sabotage can physically disrupt nuclear equipment, introduce viruses/malware into the power plant and lead to a nuclear explosion. Some examples of the cyberattack on nuclear power plants are the Stuxnet computer worm attack on Iran's Nuclear Power Plants, the cyberattack on India's Kudankulam Nuclear power plant, and the Ukraine power grid hack (Kumar and Gupta, 2021).

The new generations of Cyber-Physical Systems (CPS) consisting of software and physical parts (Akazaki *et al*., 2018) are more vulnerable to threats and easily breach the integrity of these systems. Moreover, the sensors of these systems can be hacked by hackers and false data can be infiltrated into the system so that the controller works on the malicious data. The attacker can even compromise the actuators so they won't function properly (Humayed *et al*., 2017).

In tech support scams, cyber attackers use fear tactics to convince people to pay for overpriced "help services" to diagnose their computer hardware and software-related technical issues. It's been noticed that cyberattacks impersonalize pension disbursing officials and collect the information of senior citizens to seek benefits it. Online gaming has become a way of entertainment for youngsters but this raises the opportunity for adversaries to introduce cyber-attacks. The most common cyber security threats in online gaming include disclosing personal information, location, IP addresses of devices, flooding, hacking, server maintenance problems, etc., (Shabut *et al*., 2016).

The safety and data confidentiality of citizens are the primary concerns in a smart city. The authors (Hamid *et al*., 2019) have discussed major cybersecurity issues and challenges in smart cities. The cyber security issues are categorized from three perspectives: Governance, technological and socio-

economic. The governance process uses Information Communication Technology (ICT) tools and the Internet to deliver information and public services. Smart cities have to assure citizens' privacy, providing confidentiality and security benchmarks to ensure security and privacy.

Recent years have witnessed the boom of social networks becoming more popular. Billions of people are using social media such as Instagram, Facebook, Twitter, YouTube, LinkedIn, etc., to socialize and interact with each other. However, targeted spam, phishing, defamation, impersonation, cyberbullying, and fake accounts are some of the threats most common in social cybersecurity (Thakur *et al*., 2019). Moreover, cyber attackers are sending out phony copyright complaint notices to Twitter, and Facebook users which contain harmful links, and clicking such links can damage the devices or corrupt the software on the device.

Nowadays, many search engines rank web pages to give relevant search results when the user queries through Search Engine Optimization (SEO) (Dramilio *et al*., 2020). Though many organizations use special techniques to place their pages in search results, cybercriminals can use SEO poisoning to design malicious websites and use search engine optimization tactics to make them appear predominantly in search results. This type of attack method is also referred to as search poisoning. Blockchain and cryptocurrency are proliferating and attracting more interest than ever (Tsochev *et al*., 2021). Crypto transactions are digital and business entities must apply relevant cybersecurity measures to protect against security breaches, identity theft, and other potential threats. Secure key storage management and inviolable computing are critical needs for Blockchain devices (Urien, 2021).

As Cloud Computing relies on the Internet, the usage of the cloud is increasing tremendously and has become a competitive need; securing could architecture is a significant concern (Krishnaveni *et al*., 2021). Some major threats to the could architecture are DoS Attacks, Insider Risks, Account hijacking, Data breaches, Misconfiguration, and Reduced Infrastructure Visibility.

An insider attack is one more security threat to any organization (Suresh and Madhavu, 2021). Here the attackers may be current or former employees, business partners, officeholders, consultants, or the Board of Directors. The disgruntled employees may decide to bring harm to the organization purposefully. Employees with malicious intentions may disclose the organization's secrets to outsiders as they may know network design, susceptibilities, and access codes. Though reckless users may not intend to cause any harm, they have access to the organization's information and proprietary data that they accidentally expose. Cyber security is the biggest challenge with insider attacks, where insider fraud has to be detected and rectified (Pantelidis *et al*., 2021). Insider attacks pose severe threats to CPS like Smart Cities and their components (Hossain *et al*., 2020).

5G is emerging quickly, providing high speed and responsiveness for wireless communication technology (Cabaj *et al*., 2018). But the new technologies come with unknown risks for which cyber security professionals have to find solutions for potential threats. 5G networks play an important role in smart cities, identity authentication, online banking, etc. Cybersecurity is essential to secure transactions, mitigate identity theft, protect user data and identities, and additional Intelligent Access Control (IAC) mechanisms (Sedik *et al*., 2021).

It is observed that Cyber threat is a global issue and many countries are getting affected by them. Figure 2 gives a glimpse of countries affected by major Cyber-attacks (Statistics, 2021). There is an increased need to address Cyber-attacks and defense mechanisms worldwide.

Domains In Cybersecurity

The domains of cyber security are discussed in this section. There is no rigid boundary for these domains as they keep evolving. The major domains identified are the social domain, information domain, Physical domain, and cognitive domain. The physical domain includes protecting the system/desktop machine and the peripheral hardware components from cyber thefts. The information domain focuses on Confidentiality, Integrity, and availability of data. The Information domain proposes strategies for shielding programs, data, computers, and networks from unauthorized access or attacks. The information security model is for an organization's policies that keep the data safe. The perception of the data, the analysis, and the way data is used in decision-making explain the Cognitive domain. Social domain deals with the norms, ethics, and policies of the organization and the broad social landscape (Collier *et al*., 2013). The technical implementation of different forms of security involves application security, information security, vulnerability management, network security, cloud security, cryptography, critical infrastructure security, etc.

Different forecast or prediction methods are used to find the consequence of cyber-attacks (Husák *et al*., 2018). The taxonomy of attack prediction and forecasting methods in cybersecurity is explained in more detail in Fig. 3. In the attack projection and Intention recognition method, the intentions of the attacker and their next move are predicted. The upcoming cyber-attacks are predicted in Intrusion prediction. The prediction of cyber-attacks on the whole network is done in network security situation forecasting. There are various approaches for formulating cybersecurity threats as models like Bayesian networks, Markov models, attack graphs, or continuous models like grey models, time series, etc. The cyber security issues

can also be tackled by machine learning, deep multilayer representational learning, and knowledge discovery approaches.

Intrusion Detection Systems (IDS)

Modern networked enterprises require highly sophisticated technology to safeguard the organizations. Intrusion discovery Systems are used as security tools to detect possible intrusions in a network or a host (Berman *et al*., 2019; Sharma *et al*., 2016; Alom *et al*., 2015; Uğurlu and Doğru, 2019; Kim and Aminanto 2017). An intruder within or outside the organization may initiate anomalous activity to disturb network operations. IDSs protect the system by providing user authentication and ensuring protective access from unauthorized users to gain more system privileges or misuse their privileges. It also guarantees to prevent the loss of data privacy (Alom *et al*., 2015; Gümüşbaş *et al*., 2020). IDSs can distinguish between malicious and benign actions (Ferrag *et al*., 2020). Based on the functionality, IDSs can be Network-based, Host-based, and distributed based (Alom *et al*., 2015; Karatas *et al*., 2018). Similarly, based on detection methods, intrusion discovery systems can be operating as (i) rule-based (also called anomaly-based), (ii) signature-based (also called misuse-based) while analyzing and detecting attacks, and (iii) hybrid (Gümüşbaş *et al*., 2020; Karatas *et al*., 2018; Macas and Wu, 2020; Chaudhary *et al*., 2020; Lakshminarayana *et al*., 2019).

In a rule-based, the normal behavior states of the system are stored in the database. The program behavior is continuously monitored if any deviation apart from these specified rules is indicated with alarms. A malware detector has a data collector that collects information about the program interpreter and data matcher. Program interpreter converts the data to a useful representation, data matchers compare the interpreted data with the program behavior (Al-Janabi and Altamimi, 2020). Most of the literature classifies anomaly detection as shown below (Alabadi and Celik, 2020):

1) *Point Anomalies*-It is the data point that is treated as abnormal when it is compared with the remaining data
2) *Contextual Anomalies*-The abnormality is based on a particular context
3) *Collective Anomalies*-It is the collection of data points as a dataset which is treated as anomalous

The anomaly-based IDS performance is better for unknown and complex attacks than the Signature-based attacks (Hindy *et al*., 2020). It is good at detecting unknown attack types (also called zero-day exploits) (Rashid *et al*., 2020). The major challenge with anomaly detection is to segregate normal and aberrant behavior. The anomaly detection solutions are not standard across the applications. It is difficult to predict since malicious activities keep evolving continuously. And also, all unseen behaviors are treated as anomalies, raising false alarm rates.

In signature-based (also known as Knowledge-based), the attack patterns usually are stored in the data repository. These attack patterns are compared in the network by the IDS. In signature-based attacks, the attack types which are already known can be detected with high accuracy and they don't generate false alarms. Usually, Signature based IDS achieves higher detection performance compared to anomaly-based for the known attack types (Kim and Aminanto, 2017). The drawback of Signature based is that it can identify attacks mentioned only in the database. The administrator must update the database rules and signatures very frequently (Gümüşbaş *et al*., 2020; Kilincer *et al*., 2021; Mahdavifar and Ghorbani, 2019; Hwang *et al*., 2020). Extracting the different signatures requires a lot of time and effort. It doesn't provide accurate results for zero-day attacks and viruses which have polymorphic behavior (Al-Janabi and Altamimi, 2020; Pu *et al*., 2020). "*Zero-day*" refers to newly realized security susceptibilities that attackers can exploit to attack systems. In other words, the vendor or the developer has "zero days" to fix it (He *et al*., 2021). The vendor or developer would have just learned about the flaw.

Another distinct way of detecting the intrusion is known as the Hybrid method. This method integrates the advantages of both anomaly and misuse detection. It increases the intrusion identification rate and minimizes the false positive rate for unknown attack types. Most of the ML/DL techniques are hybrid intrusion detection (Mahdavifar and Ghorbani, 2019). In hybrid attacks, unknown attack types are identified by anomaly detection and known attacks are detected by misuse detection. Hybrid detection is divided into 2 categories (Rashid *et al*., 2020) sequence-based detection, (ii) parallel based detection. In the former, either misuse or anomaly detection is used first. In the latter approach, multiple detectors are applied in parallel to obtain multiple outputs for decision. The complete classification of IDS is viewed in Fig. 4.

An Intrusion Prevention System (IPS) is a protection mechanism for interconnected devices that continuously observes malicious activity in the network (Krishna *et al*., 2020; Chandre *et al*., 2018). It takes suitable action to prevent those activities by blocking, dropping, or reporting them. An Intrusion Prevention System (IPS) shall be a signature-based or statistical anomaly-based system (Krishna *et al*., 2020). With Intrusion prevention systems, one can control access to an IT network and protect it from misuse and attack.

Some opensource IDS/IPS systems are OSSEC (Open-Source Security), SNORT, Suricata, Zeek, Samhain, Fail2ban, Security Onion, Bro-IDS, Kismet, OpenWIPS-ng, Sagan, etc., (Sokolov *et al*., 2019; Chaabouni *et al*., 2019).

Types of Cyber Threats

The Cyber Security world is not static. Cyber threats are changing at a rapid speed. Also, Cyber Security defense tactics and attack methods are changing and enhancing every day. Some of the major cyber threats addressed in this study are given in Fig. 5.

Cyber threats are broadly categorized as follows.

A. Denial of Service (DoS)

Generally, the banking sector, government organizations, commercial applications, media companies, etc. are vulnerable to DoS attacks (Hamid *et al*., 2019). In a denial-of-service episode, the hacker floods the systems, networks, or servers with undesirable requests. It makes the server resources and bandwidth exhausted with attackers' traffic. As a result of the denial of service, the system cannot fulfill legitimate requests from legitimate users.

Suppose the attackers use several compromised devices to venture DoS attacks; the attack is well-known as a Distributed Denial-of-Service (DDoS). According to several researchers in the literature, it is proved that DDoS has several repercussions. The motivations behind introducing DDoS attacks are to disturb the traffic of a targeted service or service for financial benefit, economic growth, to take revenge, ideological belief, intellectual challenge, Cyber Warfare, etc., (Zargar *et al*., 2013). DDoS attacks are categorized into various types of divisions based on the objectives of the attack (Chen, 2020) they are.

1. DDoS Flooding Invasion at Network/Transport-Level

The attacker makes Network's bandwidth resources unavailable. This flooding attack is further classified into different types as follows (Chaabouni *et al*., 2019).

a) Flooding Attacks

In flooding attacks, the hacker overwhelms the target's network bandwidth by sending false requests, mainly with ICMP or UDP packets, ultimately disrupting the legitimate user.

These kinds of attacks can be initiated with botnets.

b) Protocol Exploitation Flooding Attacks

Protocol attacks utilize the processing capability of network infrastructure resources like firewalls, servers, and load balancers. They target Layer 3 and Layer 4 protocols with malignant requests for connection.

c) Reflection-Based Flooding Attacks

Here the attacker spoofs the target's IP address and transmits the request to the devices that provide service. The server responds and replies to the target's IP address. To do this the attacker primarily uses UDP or TCP in some cases, thus having the same protocol as *'Reflection'* in both directions.

d) Amplification-Based Flooding Attacks

In Amplification attacks, attackers send the "trigger packet" to reflector devices by setting the source IP address as their target's IP address. It in turn overwhelms the victims' machine with the trigger packets. The attacker can send millions of these requests to vulnerable services, thereby generating considerably enormous responses than the original request and significantly boosting the size and bandwidth allocated to the target.

2. Attack on System's Resources (SYN Flooding Attack)

This attack uses the TCP handshake process required to launch a TCP connection. In this, the invader floods the SYN messages to the server, for which it responds with a confirmation. As the requests are fake, the server waits for the client to complete the handshake mechanism and retransmits SYN + ACK continuously until timeout. Ultimately the server is called on to keep open many half-open connections that eventually overwhelm resources such as CPU time, memory, and other device resources, often to the point where the server crashes.

3. Application-Level DDoS Flooding Attacks

These sophisticated DDoS attacks exploit weaknesses at the application layer. It opens connections, initiates processes, and performs transactions that would deplete finite resources like disk space and available memory. Application-level DDoS attacks are categorized into (Vanitha *et al*., 2017; Zargar *et al*., 2013):

a) Flooding attacks with Reflection/Amplification: It is similar to network/transport level attack
b) HTTP flooding attacks: Four varieties of this type of attacks are as follows

i) Session Flooding Attacks

It exhausts the server resources by sending a high rate of session connection requests to the server e.g.: HTTP get/post flooding attack.

ii) Request Flooding Attack

In a request flooding attack, attackers send sessions containing more requests than usual which results in a

DDoS flood denying the service to the client e.g.: A session HTTP get flooding/HTTP post flooding.

(iii) Asymmetric Attacks

In asymmetric attack, attackers transmit sessions that include high workload requests of several HTTP requests embedded in a single packet e.g.: Multiple HTTP get/post flood.

iv) Slow Request/Response Attacks

In a slow request/response attack, the attacker sends partial HTTP requests that consistently and rapidly grow, gradually update, and will not terminate. The episode persists until these requests take up all available sockets and the web server becomes inaccessible. e.g.: Slow Loris attack, HTTP fragmentation attack, Slow post attack, slow reading attack.

B. Botnet Evasion Attacks

The botnet attack is a multi-stage, predominant cyber-attack that begins with scanning network devices. It infects the devices with malicious software, like viruses (Hussain *et al*., 2021). To increase the magnitude of their attacks, attackers can gain control of a botnet without the device owner's understanding. Further, a botnet overwhelms systems in networks in a DDoS attack. Even though the actual target for botnets is computers, in recent years' adversaries are targeting Internet of Things (IoT) devices more often (Yamaguchi, 2020). In 2016, the Mirai botnet targeted half a million IoT devices with open telnet ports and used default usernames and passwords to log in to those devices and turn them into zombies (Kambourakis *et al*., 2017). The intention of launching a botnet attack is to initiate malicious activities such as spam generation, key logging, copyright violation, etc. Habitually bots use various invasive approaches to gain the maximum benefit (Karim *et al*., 2014). The originator of botnets is commonly known as Bot Masters, typically a person or an association of people who have the intention of launching malicious activities.

Botnet communications are classified into (Dhayal and Kumar, 2018):

1) Centralized botnet (i.e., the client-server model)
2) Decentralized Botnet (i.e., peer to peer communication model), and
3) Hybrid model

For instance, *Mirai, Muhstik, Toraii, Hakai, Trojan, Gagfyt, Okiru, Kenjiro, Hajime, IRCBot, Hide and seek botnets* are the most common botnet attacks (Hegde *et al*., 2020).

C. Malware Attack

'Malware' is a term derived from the words 'malicious' and 'software'. Malware is broadly used to refer to worms, ransomware, viruses, spyware, adware, Trojans, and other types of harmful software. In a Malware attack, adversary's trespasses network vulnerable links when a person connects a suspected link or opens an attachment of an email. It leads to the installation of unsecured or untrustworthy software on the system. Recently a large number of new malwares are generated by using metamorphic, polymorphic, and different evasive techniques (Vinayakumar *et al*., 2019). Initially, the malware is in an incubation period during which it will be propagating silently in the network by infecting the hosts. In the incubation period malware does not harm any system in the network and the attack is launched only when it's guaranteed enough systems are infected. During the expansion period, it propagates the entire network by launching/infecting bots. The malware detection probability and the incubation period are the key factors that determine the extent of malware attack severity (Xu *et al*., 2020).

Malware got different names based on its behavior and its purpose. The most common types of malwares include Malvertising, Cryptojacking, Spyware, Adware, Ransomware, Trojan horse, Worms, Rootkits, Man-In-TheMiddle (MitM), Backdoors, Viruses, Bot, Scareware, Man-InThe-Mobile (MitMo), etc., (Al-Janabi and Altamimi, 2020). According to the author (TM, 2020), the recent malware attacks were Shlayer, ZeuS, Agent Tesla, NanoCore, CoinMiner, Delf, Gh0st, Jupyter, Arechclient2, Mirai.

Generally, the analysis and detection techniques for malware attacks are classified into (Al-Janabi and Altamimi, 2020; Top 10 Malware, 2020; Albasir *et al*., 2018; Lin *et al*., 2020; Baptista *et al*., 2019):

1) Dynamic
2) Static and
3) Hybrid

Static analysis is faster as they can analyze the code without running and they deal with false-positive. Techniques based on static analysis are computationally effective and safer. The static analysis does not predict malware more accurately since it shows up only for some of the patterns. It can detect the most common types of malwares. However inefficient for advanced malwares which utilizes advanced evasion detection techniques such as polymorphism and obfuscation (where evidence of malicious activities is hidden) (Mahdavifar *et al*., 2020).

Dynamic analysis works with executed code and is effective against obfuscation. The dynamic analysis uses the characteristics of the Malware and the malware functionalities to determine the severity of the Malware. Moreover, the behavior of malware functionalities is determined after executing executable code in a sandbox environment. Dynamic analysis can detect any unseen samples as the file is analyzed in virtual environment

systems to improve performance. As the file will be executed, dynamic analysis achieves better accuracy and determines all matching patterns. Hybrid techniques leverage the advantages of both static as well as dynamic methods.

D. Spam and Phishing Attack

Phishing is sending fraudulent communications that seem to come from a reputed origin, usually by email. Through phishing attacks, the intruders pose as trustworthy contacts and gain sensitive information from the user (Sajal *et al*., 2019; Singh, 2020). To capture vital sensitive information like credit card numbers, PINs, and login information of users, phishing attacks are launched by adversaries. In phishing either login credentials or malware, the software is installed on the victim's machine. Phishing is a common cyber threat in social media such as Twitter, Facebook, etc. Phishing emails convince users with faultless words and original logos. Phishing links are linked to websites that are malware infected. Phishing attacks are exploiting human vulnerabilities more than system vulnerabilities. It makes the user enter his/her details into a fake website that resembles a legitimate website (Al-Janabi and Altamimi, 2020; Patil *et al*., 2017; Gupta *et al*., 2021).

Singh (2020); Tang and Mahmoud (2021) the phishing attacks are classified into two major types: (i) Social engineering (i.e., deceptive phishing) (ii) Malware-based phishing. Social engineering attacks usually with the psychological manipulation of users to make some mistakes or share their confidential information. In Malware based phishing, malicious softwares are executed on the user's machine to fetch users' confidential information. Malware-based phishing attacks are DNS phishing, Session hijacking, content injection phishing, key loggers, phone phishing, link manipulation, system reconfiguration etc. It is possible with Phishing attacks to install malwares in the victim's machine which can change the victim's machine into a Botnet and botnets can now be able to launch DDoS or any other kind of attack.

Author in (Alam *et al*., 2020) explained the classification of Phishing attacks. Different phishing attacks are as follows.

1) Algorithm-Based Phishing

It was first identified in the year 1996, wherein the phisher developed an algorithm to generate random credit card numbers to match the original credit card numbers of America Online (AOL) Accounts (Tang and Mahmoud, 2021; Khonji *et al*., 2013).

2) Deceptive Phishing

In a deceptive attack, the attacker uses emails or SMSs to send fraudulent links and trick people to click the links. The websites behind the links snatch and store the personal information of the victim.

3) URL Phishing

In this attack, the attackers use the phishing page's URL to infect the target. The hidden link is to the hacker's website. When the victim clicks on the URL, it is directed to the hacker's website snatching the victim's information.

4) Hosts File Poisoning

The host file in the operating systems is poisoned so that when the user requests the desired website, either it is rerouted to another website or it returns a "Page Not Found" error. When it is redirected to a fake website, the user data is stolen. By poisoning the host file, the way the OS resolves a DNS name is altered.

5) Content-Injection Phishing

Content Injection phishing is a common web security vulnerability. The vulnerable web applications make the actual content on the web page be spoofed or modified. Content injection phishing occurs when the application is not properly handling the user-supplied data and the attacker can supply the content to the web applications.

6) Clone Phishing

In Clone Phishing, an email that is sent before containing any link is used to create an identical copy of the email but with a malicious link. This new email is just a replica of the original but with fake links or attachments. This duplicated email is sent to all contacts from the target’s inbox. The person receiving the cloned mail clicks on the fake links, assuming it to be a legitimate email. This attack is hazardous as the recipients will never suspect the email.

7) Whaling

Whaling attack always targets high-profile executives like CEOs, CTOs, and Directors (Sajal *et al*., 2019). The attackers usually make the victims act such as fund transfer. It is difficult to find these attacks as they often don’t use malicious URLs or weaponized attachments.

8) Spear Phishing

Phishing usually targets a large number of recipients but spear phishing emails are carefully designed to get data in the form of a response from a particular person. Though the risk rate is high, spear phishing is having a high success rate and has become one of the major aspects affecting network security (Xiujuan *et al*., 2019).

Email phishing and URL phishing are difficult to identify as attackers frequently change their strategies (Alam *et al*., 2020). Some of the protection approaches against phishing attacks include Client-side tools, Authentication, Server-side filters and classifiers, network-level protection and also educating the users (Singh, 2020).

E. Domain Generation Algorithm (DGA)

It's a type of attack in which adversaries design a software program that generates an extensive number of pseudo-random domain names (Shahzad *et al*., 2021). With this DGA, the malware will generate hundreds to thousands of domain names randomly in a short period. The generated domain names are explicitly assigned to sites. The domain names assigned for the sites will receive control from the malware and give their instructions. DGAs are common in Malware, which endeavors to install command and control communication with the botmaster and the infected machine (Chen *et al*., 2021). This is referred to as "command and control" or C2 (Yu *et al*., 2019). Since domain names are short-lived, it is a challenge for the defenders or the analysts to detect them (Li *et al*., 2019). Using DGAs, the attackers can manage the infection-spreading websites and deploy the command and control (C&C). The DGA attack constitutes the following phases: Infection, C&C, Lateral Spreading, and Data exfiltration. DGA attacks can be broadly classified into Binary and Script based, depending on how they are deployed (Sood and Zeadally, 2016).

F. Spoofing

Spoofing is also called an impersonation attack. In spoofing, the attacker steals the user authentication credentials to gain unauthorized access to the services. The user credential can be obtained by eavesdropping on the network or can be stolen from the device using a phishing attack. The attacker links their MAC address to the IP address of an unprotected network. It becomes easy for the attacker to perform theft or delete the data in this vulnerable network. Spoofing can be commonly categorized into ARP (Address Resolution Protocol) Spoofing, IP Spoofing, and DNS Spoofing (Hamid *et al*., 2019; Chaabouni *et al*., 2019). In ARP spoofing, the attacker sends the spoofed ARP message into the LAN. The Media Access Control (MAC) address of the attacker is then attached to any one of the legitimate users in the LAN. With this, the attacker will be able to modify, steal or even stop network traffic. IP address spoofing is done by modifying the source IP address with which the sender's identity is modified.

DNS spoofing occurs by modifying the entries of a DNS server (which maps domain names to IP addresses). The attacker can now be able to reroute the particular domain name to a malicious or infected system.

G. Probing

The attacker uses probing to get to know the weak points in the system and attain entry to it. The hackers send the scan packets into the system and efficiently collect the information and data. Examples of attacks include Nmap, Satan, port sweep, IP sweep, mscan, etc (Gümüşbaş *et al*., 2020; Dixit and Silakari, 2021).

H. Remote-to-Local (R2L)

The invader, in an R2L attack, identifies the device's vulnerability by sending packets over the network. Then the attacker acquires unauthorized access to the victims' machine (Elsayed *et al*., 2020). The attacks are usually caused by buffer overflow (as in *imap, named, sendmail*), misconfigured security policies (as in *ftpwrite*), or Trojans (*xsnoop*). R2L attack may be challenging to detect as it involves both network-level and host-level features (Rodriguez *et al*., 2021).

I. User-to-Root (U2R)

In a U2R attack, the user gets legal entry to the account (target machine), with which the attacker illegally attempts to acquire superuser permissions of the root by using the susceptibilities of the system (Sapre *et al*., 2021; Begli *et al*., 2019). Examples of this attack are Load Module, Eject, Buffer_overflow, and Perl attacks.

J. SQL Injection

SQL injection attacks mostly attack web applications. The loopholes in the websites' databases are used to compromise the database. With this, hackers can access confidential user information on the website (Kilincer *et al*., 2021). The hackers can even modify, delete or update the user information on the website. These attacks allow attackers to spoof identity, cause repudiation problems, destroy the data or make data inaccessible and even change the administrator's setting of the website server. Attackers can access the backend data based on the methods, such as Out-of-band SQLi, In-band SQLi (Classic), and Inferential SQLi (Blind).

Data Sets used in ML and DL Techniques

Data and datasets play a crucial role in cyber security research to conduct research and evaluate the research activities in the field of Cyber security. It is essential to identify and use the relevant dataset to conduct the research experiments to estimate the significance and performance of suggested Cyber Security solutions. The effectiveness and implementation of the ML and DL models rely on the size of the datasets that are used in training ML and DL methods (Xin *et al*., 2018). To construct efficient IDS, relevant heterogeneous and massive datasets are required in training proposed models and evaluating the performance of proposed IDS (Sohn, 2021). Some of the vital dataset over some time is depicted in Fig. 6.

Ferrag *et al*. (2020) have classified available public datasets into 7 categories. The classified datasets are based on network traffic, internet traffic, electrical network, virtual private network android applications, IoT traffic, and internet-connected devices.

Buczak and Guven (2015) have illustrated the significance of ML approaches in intrusion detection systems. The authors used packet-level, Netflow, and public data to evaluate the ML algorithm.

The MIT Lincoln Labs have developed and maintained several public datasets. The datasets are commonly referred to as DARPA datasets and their datasets are available for public use. (Husák *et al*., 2018; MITLL, 2022; Lippmann *et al*., 2000). The publicly available datasets are useful for researchers to conduct experiments in the area of Cyber security. The distinct datasets are DARPA 1998, DARPA 1999, and DARPA 2000.

Generally, the performance of ML and DL based Intrusion Detection Systems are evaluated using available datasets such as NSL-KDD, KDD CUP 99, UNB ISCX 2012, ADFA, UNSW-NB15, CSE-CIC-IDS 2018, CIC IDS 2017 (Gümüşbaş *et al*., 2020; Sohn, 2021; Uğurlu and Doğru, 2019; Karatas *et al*., 2018). DEFCON, CAIDAs, LBNL, CDX, KYOTO, TWENTE, UMASS, and ADFA2013 are some more IDSs dataset evaluation frameworks (Ferrag *et al*., 2019).

A. KDD99/KDD CUP 99 Dataset

KDD99 datasets are created in a Competition, namely, 3rd International Knowledge Discovery and Data Mining Tools, that was held in association with KDD-99, The 5th International Conference on Knowledge Discovery and Data-Mining (KDD). These datasets are based on the DARPA 1998 PCAP files (Gümüşbaş *et al*., 2020). These are widely used in differentiating intrusion and normal traffic (Andresini *et al*., 2020). There are more than 5 million different data available in the dataset. In total, there are 41 traffic features mainly to give the information about source IP address and source port. These 41 features can be categorized into 3 classes such as content features, traffic features, and basic features (Ferrag *et al*., 2020). The Basic features (also referred to as Intrinsic attributes) are extracted from the network packet's header. The Content attributes are extracted from the contents area of the network packets. Traffic attributes are calculated based on the previous connections. Traffic attributes are grouped into (1) Time-based traffic features and (2) Host-based (machine) traffic features (Chahira, 2020). The summary of the features is given in Table 1.

The KDD99 dataset has a collection of 20 different types of attacks (Rashid *et al*., 2020). The attack traffic can be classified as (Berman *et al*., 2019; Alom *et al*., 2015; Uğurlu and Doğru, 2019; Ferrag *et al*., 2020; Karatas *et al*., 2018; Li *et al*., 2019; Kadam *et al*., 2020):

1) Normal (nonattack type data)
2) Remote to Local (R2L)
3) Probing
4) User to Root (U2R)
5) Denial of Service Attack (DoS)

The different attack class with relevant features of the KDD CUP99 data set is summarized in Table 2.

KDD99/KDD CUP 99 Dataset induces many drawbacks such as duplicated samples, unbalanced classes, training, and test data having different probability distributions. Moreover, these datasets do not include the newest attack types (Gümüşbaş *et al*., 2020; Hindy *et al*., 2020). These data sets are available in two files such as 'KDDTrain^{+}.csv' and 'KDDTest^{+}.csv'.

B. NSL-KDD Dataset

It is the next version of the KDD-CUP-99 Dataset. These datasets are reformed by removing duplicative instances in KDD-CUP-99 and reconstructing the structure of the datasets (Sohn, 2021). Restructured data enables the classifier not to give any biased results. Compared to KDD CUP 99, NSL-KDD Dataset gives a lower reduction ratio since there is no repetitive data. There are 41 attributes in the dataset giving the different features of the traffic. The features are classified into numerical (38 features) and categorical (3 features such as "flag", "service", and "protocol_type") (Li *et al*., 2019). The 4 major attack categories are listed in (Karatas *et al*., 2018):

1) Remote to Local (R2L)
2) DoS
3) Probe
4) User to Root (U2R)

The NSL-KDD dataset is divided into KDD Train^{+} and KDD Test^{-21}. KDD Train^{+} is used to train an IDS model, and KDD Test^{-21} is used for testing the datasets (Sohn, 2021; Li *et al*., 2019).

C. UNB ISCX 2012 Dataset

UNB ISCX 2012 Dataset was created at the University of New Brunswick (UNB) in 2012. UNB dataset includes traffic with normal data and attacks data for the DoS, infiltration, DDoS, and SSH attacks with brute force method (Gümüşbaş *et al*., 2020). The UNB ISCX 2012 dataset for Intrusion Detection Systems is from the realistic network and traffic covering diverse intrusion scenarios. This dataset has statistical features such as protocol, source_bytes, direction, time_stamp, source_packets, dst_bytes, dst_packets, source_ip, Tag, and dst_ip. The real network traffic of POP3, IMAP, SMTP, HTTP, FTP, and SSH protocols are analyzed to determine the expected behavior of computers (Ghurab *et al*., 2021). It consists of traces of labeled network data that include the payload of a full packet in *pcap* (Packet Capture) format. Datasets are made open for public use by researchers (Shiravi *et al*., 2012).

D. UNSW-NB15 Dataset

The datasets are generated at the Cyber Range Laboratory by the Australian Center for Cyber Security

(ACCS), a cyber-security research team employing IXIA Perfect Storm, Argus, Tcpdump, and Bro-IDS tools. The tools are specifically designed to create DoS, generic, shellcode, reconnaissance, worms, and exploits (Ferrag *et al*., 2020). Datasets have 49 features with two million and 540,044 vectors. The features are grouped as basic, time, content, flow, connection, labeled, and general (Gümüşbaş *et al*., 2020; Ferrag *et al*., 2020; Sohn, 2021). The dataset is publicly available and consists of 45 distinct IP addresses.

E. CICIDS 2017 Dataset

Sharafaldin *et al*. (2018a); Ring *et al*. (2017) of the Canadian Institute for Cybersecurity Intrusion Detection System (CICIDS) created this dataset in 2018. The dataset comprises the normal data and attack data are gathered for five days. Table 3 summarizes the data collection (Pantelidis *et al*., 2021; Sharafaldin *et al*., 2018b). The database consists of 2830540 samples containing 83 features. The important features extracted are flow duration, destination port, total backward packets, total forward packets, etc., (Ho *et al*., 2021). It covers a variety of insider and outsider attacks. The common attacks covered in the Dataset are DoS, Web Attacks, Botnet, Brute Force SSH, DDoS, Brute Force FTP, Heartbleed, Infiltration, etc., (Rashid *et al*., 2020).

F. CSE-CIC-IDS2018 Dataset

The CSE-CIC-IDS2018 dataset is formed by the Canadian Institute for Cyber-security (CIC) and the Communications Security Establishment (CSE). It addresses seven distinct types of attacks, such as heartleech, bruteforce, botnet, DoS, DDoS, infiltration, and web attack (Rashid *et al*., 2020; Shibahara *et al*., 2016). To collect the data, the victim organization Constituted five departments with 420 personal computers with 30 servers and the attacking infrastructures consisted of 50 machines. The CSE-CIC-IDS2018 dataset consisted of network flow data, event files from each victim's machine and 80 different network traffic attributes from CICFlowMeter-V3 (Rashid *et al*., 2020).

G. BGP Datasets

Li *et al*. (2019) discussed the Border Gateway Protocol (BGP) datasets. Datasets include the routing logs from Reseaux IP Europeens (RIPE) and BCNET. The dataset includes the data of the day of the attack (anomalous data points) and the data collected two days before and two days after the invasion (regular data points). The record has 37 features that are extracted from the dataset. This dataset possesses information regarding the attacks such as Code Red I, Nimda, and Slammer. BCNET contains the regular data.

H. ISCX URL-2016 Dataset

The dataset contains samples of different types of URLs (Mamun *et al*., 2016). The collected URLs are classified into five different types of URLs:

1) Benign URLs-these are legitimate URLs that lead to any malicious websites
2) Spam URLs-it contain out-of-context links to websites, discussion forms, etc., to promote spammer sites
3) Phishing URLs-the URLs make the user visit a fake website and thus steal the personal information of the victim
4) Malware URLs-the URLs take the user to a malicious website that installs some malware on the victim's device
5) Defacement URLs-here the Hacktivists try to deface a website for their benefit which technically means penetrating a website

Table 4 shows the various URL types, their sources, and several samples.

I. CICMalDroid2020 Dataset

Mahdavifar *et al*. (2020) a new dataset named CICMalDroid2020 is proposed. CICMalDroid2020 is a collection of samples taken from five distinct classes of Android applications such as *SMS*, *Adware*, *Riskware*, *Banking*, and *Benign*. This includes 17,341 samples capturing the static and dynamic features from the publicly available datasets.

J. CIDDS-001/CIDDS-002 Dataset

Coburg Intrusion Detection Dataset (CIDDS) was created by Ring *et al*. (2017). CIDDS-001/CIDDS-002 datasets are used for evaluating anomaly-based intrusion detection. 32 million of data are collected from Open stack servers and External servers.

K. Drebin

Drebin is a famous, widely accepted malware dataset used in the Android operating system and was part of the Mobile Sandbox project (Mishra *et al*., 2021). The Drebin dataset comprises 5,615 malicious Android packages and 123,453 benign examples of SHA256 values and covers 129,013 mobile applications. The records are accumulated from August 2010 to October 2012. The Drebin dataset consists of 545,356 features, which results in a feature vector with high sparsity. It can be represented as a matrix of size 129,013 × 545,356 (Arp *et al*., 2014; Salah *et al*., 2020). Four well-known malware families and their top 5 features are listed in Table 5 (Arp *et al*., 2014).

L. Kyoto University Honeypot Dataset

Kyoto 2006+ is created from the traffic data over three years from Nov. 2006 to Aug. 2009, using various

variants of honeypots (Song *et al*., 2011). The dataset includes 24 statistical features, of which 14 features are conventional features extracted from the KDD Cup 99 data set, a prevalent and widely used habituated evaluation data in intrusion detection (Krishnaveni *et al*., 2021). The dataset has ten additional features such as Source_Port_Number, Ashula_detection, Source_IP_Address, IDS_detection, Malware_detection, Label, Destination_Port_Number, Destination_IP_Address, Protocol and Start_Time (Ghurab *et al*., 2021).

Various honeypots such as Windows machines, Linux/Unix machines, and dedicated honeypots introduced in network printers, home appliances, etc. were used to collect the actual data. The honeypots were deployed on five different networks, such as 1 class A and 4 class B networks inside and outside Kyoto University. Research on the Kyoto dataset mainly concentrated on detecting anomalies, notably feature analysis, ensemble classifier, and dimensionality reduction (Salo *et al*., 2019).

M. The CTU-13 Dataset

CTU-13 Dataset consists of labeled data with Botnet, Normal and backdrop traffic grabbed in the Czech Technical University (CTU), Czech Republic, in the year 2011 (Garcia *et al*., 2014). It consists of thirteen different captures and is referred to as scenarios of distinct botnet samples. Every botnet system is for specific Malware with a precise infection of the virtual appliances, captured through different pcap files. The Dataset includes various types of botnets with HTTP, IRC, and P2Pbased communication techniques with invasions like Click Fraud, Port Scan, DDos, Spam, and FastFlux (Kim *et al*., 2021). The CTU13 Botnet dataset scenario is shown in Table 6 (Huang *et al*., 2021; Garcia and Uhlir, 2014).

N. ADFA Dataset

Created at the Australian Defense Academy (ADFA) by Creech and Hu (2013), this public dataset was devised to overcome the constraints of the KDD Cup 99 dataset. The two variants of the ADFA data set are ADFA-LD (ADFA Linux Dataset) and ADFA-WD (ADFA Windows Dataset), containing the data from each operating system. ADFA includes 833 normal training datasets and 4373 normal validation datasets. Constructed by evaluating the system-call-based HIDS, it represents the recent attacks' format and process. The dataset contains the attack types such as Websell, Adduser, HydraSSH, Hydra-FTP, and Jara-Meterpenter.

O. AWID Dataset

Aegean WiFi Intrusion Dataset (AWID) is a public set of data consisting of real data of regular and malicious traffic of the 802.11 networks (Kolias *et al*., 2015). This dataset is exclusively for detecting intrusion in Wireless Networks. The data is collected from a devoted WEP-protected 802.11 network with actual network utilization. A physical lab was set up emulating a typical Small Office/Home Office (SOHO) infrastructure to collect the AWID data. With the labeling method, the dataset is accessed in two sets: AWID-CLS (for Classes) and AWID-ATK (for Attacks). Each of these sets has a complete subset (AWIDCLS-F and AWID-ATK-F) and a lessened subset (AWID-CLSR and AWID-ATK-R). Each subset has two versions, one for training and the other one for testing. AWID-CLS-R-Trn and AWID-ATK-R-Trn consist of 1,795,575 records, of which 1,633,190 is regular traffic and 162,385 records are intrusive. AWID-ATK-F-Trn and AWID-CLS-F-Trn have 37,817,835 records, of which 1,085,372 have some kind of attack. The AWID dataset is used for other wireless technologies such as WiMax, UMTS, LTE, or different 802.11 settings (e.g., vehicular networks, mesh mode).

P. Other Datasets

Many researchers used other datasets besides the well-known datasets for Cyber Security. Researchers (Xiujuan *et al*., 2019) used Email Dataset named Enron from CALO (A Cognitive Assistant that Learns and Organizes) project. The data is a collection of 150 users' emails, mainly the Enron Corp's senior manager. The Federal Energy Management Committee examines and publishes this email dataset to the network. In the research work (Lorenzen *et al*., 2018), data was collected from "Cybersecurity Environment for Detection, Analysis, and Reporting" (CEDAR).

These datasets are used to analyze the deep learning algorithms segregating normal and benign network activities. Li *et al*. (2019) used DGA-based domain data such as Nymaim, Tovar, CryptoLocker, Locky, and Nymaim to evaluate their proposed model for Malware detection. Authors (Singhal *et al*., 2020) used public blocklists such as PhishTank and Malware Domain List (MDL) to collect malicious URLs. OpenDNS operate PhishTank to distribute and verify phishing websites. MDL holds an archive of malware-infected websites. Yu *et al*. (2019) used AlexaBamb training data constituting domain names of Alexa, which is benign, and Bambenek, which is non-benign. Apart from these, researchers have also used the datasets from other sources such as the Cisco umbrella popularity list, Alexa Top 1M domains, OSINT DGA feed from Bambenek, and Netlab 360 for the most famous domain names for DGA Domain Detection (Shahzad *et al*., 2021).

Machine Learning-Based Approaches for Cyber Security Problems

Machine Learning algorithms build behavior models using mathematical techniques across massive datasets and make imminent predictions with the new set of input

data. Machine learning methods are adequate for intrusion discovery mechanisms. Machine Learning (ML) lets the computer learn without explicitly programming them. The frontier person of ML, Arthur Samuel, explained ML as, *a branch of computer science that emphasizes how to make the computer think (i.e., artificial intelligence) without giving explicit instructions to the machines* (Gordon, 1995). ML performs categorization and regression established on previously learned features from the set of training instances. The strategy consists of two phases: Training and testing (Buczak and Guven, 2015).

Machine Learning approaches are commonly classified as Unsupervised, Supervised, and Reinforcement techniques. The algorithm/system is trained with a set of labeled input and output data in a supervised learning algorithm. The training is done with the feature set of input and correct output that makes the model learn over time. That is, the training dataset is having the target vector. Whereas in Unsupervised Learning, algorithms learn from the training data but without any target vector available (Sharma *et al*., 2016; Martínez *et al*., 2019; Apruzzese *et al.*, 2018; Hu and Tan, 2017; Yavanoglu and Aydos, 2017; Djellali *et al*., 2019). Different algorithms and computation approaches are used in supervised techniques. The most commonly used supervised learning methods are classification and regression established on the target labels, which can be either discrete or numeric (Liang *et al*., 2019). Unsupervised learning includes Dimensionality reduction, Density estimation, and Clustering (Liang *et al*., 2019). As Unsupervised Learning doesn't require labeled training data to detect malicious activity, they are best suited for cyber security compared with supervised learning which needs labeled training data (Geluvaraj *et al*., 2019). In reinforcement learning, the machine learned by trial and error in an interactive setting with the experience and predicted output is evaluated based on positive or negative reward (Alabadi and Celik, 2020). The major reinforcement methods are Value function approximation and Policy search (Liang *et al*., 2019).

ML algorithms effectively determine zero-day attacks and unusual system characteristics (Chaudhary *et al*., 2020; Vinayakumar *et al*., 2019). Machine learning approaches that are used in threat detection systems are Selforganization maps, Support Vector Machines (SVM), Naive Bayes (NB), Bayesian classifiers, Decision trees (DT), Neural network classifiers (Sharma *et al*., 2016; Kilincer *et al*., 2021; Singh *et al*., 2014; Sohn, 2021; Iyer *et al*., 2021).

The research experiment of (Apruzzese *et al*., 2018). uses Feedforward Fully Connected Deep Multi-Layer Neural Network and Random Forest algorithms. The ML algorithms are applied in (i) Intrusion identification, (ii) Analysis of Malware, and (iii) Detection of spam. DGA Detection, Network Intrusion Detection, and Botnets are focused on Intrusion Detection. The other machine learning algorithms that are used in cybersecurity are the Bayesian approach- Bayesian classifiers, and Markov models. K Nearest Neighbor (KNN), Naive Bayesian classification, SVM, and Neural Networks are the machine learning techniques that are used in spam filtering (Patil *et al*., 2017).

Pu *et al*. (2020) proposed a blended unsupervised method for anomaly detection process combining cluster-based techniques such as One-Class SVM(OCSVM) and Subspace Clustering (SSC). SSC is an extension of the traditional clustering approaches. SVM is a supervised approach that investigates data and identifies patterns. OCSVM is an extension of the SVM model and is specifically appropriate for unlabeled data. The proposed method is evaluated utilizing the notable NSL-KDD dataset (Sohn, 2021).

The attackers use malicious websites to acquire control of the system and inject Malware to collect user details or harm the system. Generally, the attackers keep changing the URL of the malicious websites. Singhal *et al*. (2020) suggested a method to categorize website URLs as malicious or benign. The authors used the Machine Learning classifiers like Gradient Boosted Decision Trees, Random Forests, and Deep Neural Networks for the classification. For these classifiers, they used Content-Based, Host-Based, and Lexical features from the URLs. The author highlighted drift in websites to address the vibrant nature of malicious websites. Web drifts are observed by changing the association between the input data and the target variable.

In malware analysis, the ML approaches are utilized for Malware detection and classify the Malware into different categories (Li *et al*., 2020). In malware detection, algorithms classify software as malicious or benign. The major challenge with Malware is that they incorporate metamorphic, polymorphic, and other evasive techniques which can modify their behaviors and create a new type of malwares (Vinayakumar *et al*., 2019). These obfuscation techniques are used by hackers against traditional signature-based techniques. Baptista *et al*. 2019) present methods for malware detection. The proposed method is established with Self-Organizing Incremental Neural Networks (SOINN) and binary visualization. Binary data of any file is converted into an image and malicious traffic is analyzed and detected using SOINN. The converted images are preprocessed and extracted features are given to SOINN for clustering and classification. A similar process happens during the testing phase. The algorithm achieves 74% of the overall detection rate with false positives at 12% and false negatives at 14%.

To effectively detect Malware, authors (Li *et al*., 2018) designed the Significant Permission IDentification (SigPID) framework, which adopts an

SVM classifier. SigPID framework pulls effective permissions from the applications and effectively utilizes the extracted data to detect Malware employing supervised learning algorithms. To extract significant permissions, the authors proposed a Multilevel Data Pruning (MLDP) approach with Support-based Permission Ranking (SPR), Permission Mining with Association Rules (PMAR), and Permission Ranking with Negative Rate (PRNR). The authors then used an SVM classifier to categorize Malware and benign applications. The proposed framework achieves better accuracy, precision, and recall in Malware detection, which is the main objective of the framework.

The SVM classifier is one more efficient method for the detection of Malware. The authors (Hegde *et al*., 2020) proved the effectiveness of SVM classifiers in detecting botnet activities for a home IoT environment. The performance metrics used are false alarm rate, detection rate, and testing accuracy. The classifiers used for detecting botnet activities are Random Forest, Decision Trees, Two class Neural Networks, Multiclass Decision Trees, and Multiclass Neural Networks. The author concluded that the performance of the classifiers increased with the dataset size and amount and diversity of the malicious activities.

Intrusion detection systems are used to monitor malicious activities in the system. The ML-based IDS approach involves three categories such as data classification, anomaly-based method, and data clustering (Bahl and Sharma 2015). Data classification is a supervised machine learning strategy where the dataset is classified into different types of attacks. The deviations from the expected behavior are identified by an anomaly-based method, a semi-supervised machine learning technique. In data clustering, the data is clustered based on patterns.

Adaptive Bayesian Algorithm (ABA), Artificial Neural Networks (ANN), KNN, DT, and SVM are machine learning techniques that research scholars in literature extensively used to detect intrusion. The machine learning model, Radial Based Function SVM (RBF-SVM), resulted in the most increased accuracy (Chaudhary *et al*., 2020). Otoum *et al*. (2018) the author proposes an Adaptively Intrusion Detection System (Adaptive-IDS), called the Adaptively Supervised and Clustered Hybrid Intrusion Detection System (ASCH-IDS) to classify the aggregated data. This model uses machine learning techniques namely random forest-based classifier as misuse detection subsystem to detect known attacks and enhanced-DBSCAN classifier as anomaly detection subsystem to detect unknown attacks.

Begli *et al*. (2019); Hagos *et al*. (2017) used SVM in designing an intrusion detection system to prevent possible attacks like U2R, DoS, etc. The proposed methodology uses the SVM to classify the malicious traffic pattern from the typical traffic pattern, which happens to be non-linear.

Though intrusion detection employs many machine learning algorithms, each has its benefits and de-benefits. Each algorithm performs differently on different attacks. Ensemble in machine learning is a technique in which several base models of machine learning models combine to have an optimal predictive model. These ensemble models proved to be efficient in detecting cyber-attacks. Feng *et al*. (2018) (designed a unique Intelligent Intrusion Detection System framework to address multi-attack classification based on the CIC-IDS 2018 dataset. Their ensemble technique uses a blended mode of feature-selecting approach employing Random Forest (RF) and Principal Component Analysis (PCA). The other Machine Learning algorithms utilized in the suggested work are KNN, DT, Extra Trees, Light GBM, Gradient Boosting based on Histogram (HBGB), and Extreme Gradient Boosting (XGB). The framework is tested and compared with other approaches as well. Krishnaveni *et al*. (2021) proposed an ensemble method for efficient feature selection and classification of network intrusion detection for the current threats in cloud computing. This proposed approach relies on the univariate ensemble feature selection technique, with reduced feature sets selected from intrusion datasets such as Honeypot real-time dataset, Kyoto, and NSLKDD.

Wang *et al*. (2018) used the K-NN technique for supervised learning and the K-Means method in KNN classifier for unsupervised learning to enhance the performance of the intrusion classifier for U2R attacks. The authors introduce feature weighting and unsupervised learning methods in the KNN process to achieve this. Observed results reveal that the suggested approach can efficiently classify network attacks and significantly enhance the classification of U2R attacks.

Buczak and Guven (2015) proposed the ML and DL-based approaches for detecting cyber intrusion and misuse attacks that are applied in wired and wireless networks. The author focused on Misuse Detection, Anomaly Detection, and Hybrid Detection for the various models of ML and DL such as (i) Bayesian Networks, (ii) Evolutionary Computation, (iii) Artificial Neural Networks, (iv) Clustering, (v) Decision Trees, (vi) Association Rules and Fuzzy Association Rules, (vii) Sequential Pattern Mining, (viii) Inductive Learning, (ix) Support Vector Machine, x) Hidden Markov Models and (xi) Naive Bayes. These models' performances are compared with the parameters such as time to train a model, classify unidentified examples with a trained ML model, Comprehend the conclusive results (classification), and Accuracy. This research work highlights the requirement of retraining data and labeled data.

In the research work by Xin *et al*. (2018), the authors detailed the ML and Deep Multilayered Representative

Learning strategies that are employed in detecting network intrusion. They considered SVM, KNN, Decision Trees, Deep Belief Networks (DBN), Recurrent Neural Networks (RNN), and finally Convolution Neural Networks (CNN) in their study. They highlighted some problems such as the unavailability of benchmark datasets, irregular evaluation metrics, and insufficient measurement of the efficiency of the algorithms.

Feng *et al*. (2018) in their research use ML approaches to detect Distributed Cyber Attacks. The work focuses on identifying C&C (Command and Control) communication between the C&C server and the bots that are compromised. The C&C contact occurs in the preparation stage of distributed attacks. The authors used 55 features to select C&C traffic to detect the DDoS attacks early. They used mainly PCA and SVM for feature selection. SVM and RF methods are used to build the classifier. The experiment focused on decreasing the number of features used and finding the critical features necessary for the early detection of C&C communication. The study concluded that though more features are used in the detection, as the count reaches around 40, the detection performance will not very much.

The literature proved that Machine Learning algorithms are best suited for phishing attacks since they have most of the common characteristics in common (Lakshmanarao *et al*., 2021). Many ML algorithm-based results have been presented in publications to thwart phishing attacks. However, the existing ML-based solutions have higher response times, and high false-positive rates and involve third parties' (unauthenticated) information. Gupta *et al.* (2021), proposed a solution for phishing attacks that detects URL phishing attacks in a real-time environment. The authors have used well-known algorithms such as Random Forest, Spearman correlation, and K best for identifying phishing attacks. The proposed work used nine lexical-based features to achieve high accuracy with Random forests with a very low response time. The authors have done a detailed study on the response time that includes the time for feature extraction, dataset preparation, loading of modules, and predicting the results as valid or phishing attacks. Authors have concluded that the Random Forest algorithm has the highest response time and SVM has the minimum response time.

The effectiveness of other classified algorithms is verified by Iyer *et al.* (2021). The classification algorithms used are DT, K-NN, SVM, Logistic Regression (LR), RF, and Ensemble learning. Authors (Iyer *et al*., 2021) applied fusion classifiers based on priority-based algorithms such as Priority Algorithm 1 (PA1) and Priority Algorithm 2 (PA2). A final fusion is then applied based on the priorities obtained in PA1 and PA2 to achieve an accuracy of 97%.

Phishing prediction can be done using different machine learning methods such as SVM, Decision Tree, Random Forest, Naive Bayes, Bayesian Classification, K-Nearest Neighbor, and Artificial Neural Networks. The feature selection is classified as Source code features, URL features, and Image features, and these are based on rules (Singh, 2020; Tang and Mahmoud; 2021; Alam *et al*., 2020) Random forests and Decision Trees are used for detecting phishing attacks. The datasets are collected from Kaggle and feature selection is made by Principal Component Analysis (PCA). It identifies and classifies the dataset components. Decision Trees are used to categorize the website and for classification, Random Forest is used. High accuracy was achieved through Random Forest. Research works (Xiujuan *et al*., 2019) propose spear-phishing email detection based on Authentication (SPBA) which uses personality features, stylometric features, and gender features extracted from the emails of the same sender with which the identity portrait model of the sender is created. For authentication, KNN, SVM, and Random Forest are used as classifiers. The real portrait of the sender is then compared with the portrait of the uncertain email. If it is found identical then the email is treated as normal otherwise the email is classified as spear phishing from a disguised sender. This study outperforms the PHILFER and FSSPD concerning detection rate and accuracy.

Apruzzese *et al*. (2018) showed that ML algorithms are used in many problems, whereas DL algorithms are mainly used for Malware investigation, less in Intrusion detection. Unsupervised DL algorithms are used in spam detection. The results provided strong evidence that ML techniques are having shortfalls in their effectiveness for Cyber Security. A lack of human surveillance can allow professional attackers to penetrate, loot the data and even vandalize an enterprise. The authors concluded that the ML methods are prone to adversarial attacks, the algorithms need continuous re-training and the parameters need to be carefully tuned (Apruzzese *et al*., 2018). The attacker can perform adversarial attacks on the machine learning algorithms during the training or testing (inferring) period (Chaudhary *et al*., 2020).

Adversarial Machine Learning is the ML method that makes the machine malfunction by providing wrong input to the model while training the machine. This forces the machine to make false predictions. The attack by the attacker can be a targeted attack where a specific part of the training sample is targeted or it can be a random attack where any part of the training sample is targeted. In both methods, the ultimate goal is to misclassify the output result. The adversarial effect can be an integrity violation, availability violation, or privacy violation based on the adversary's goals (Dixit and Silakari, 2021). The targeted attack on the neural network which leads to misclassification is referred to as an Integrity violation. If the targeted system is unavailable to users for a certain period, it is called an availability violation. Privacy

violation occurs if the adversary is successful in compromising confidential information. However, adversarial examples can be leveraged to enhance ML models' performance or robustness.

Research showed that ML techniques in IDS attain a heightened detection rate but a less false positive rate. But it is also observed that the ML algorithms can misclassify the network data due to poison learning (Sharma *et al*., 2016; Xin *et al*., 2018).

The process of making Machine Learning algorithms perform undesirable activity/function is referred to as an Adversarial Machine attack. The adversarial machine attacks are categorized as (Liang *et al*., 2019):

1) Poisoning (also known as a causative attack)
2) Evasion attack and
3) Exploratory attack

A poisoning attack is a kind of adversarial invasion in which the adversary in the poison attack manipulates the training dataset of a machine learning model. In a poisoning attack, the adversary gives carefully designed training data and these are induced into the system while at the training stage. The contaminated/poisoned datasets result in incorrect behavior of the model and thus resulting in a performance decrease. This definitely will affect the accuracy of the system. Poisoning attacks can be of two types: Poisoning with changing features (labels) and poisoning without changing the features (Chaudhary *et al*., 2020). In an Exploratory episode, the adversary learns the model algorithm and can manipulate the parameters of the system so that they can reach their goals (Yu and Deng 2010).

Another commonly known attack is the evasion attack. In an Evasion attack, the malicious samples are evaded/misclassified as valid during test time. Evasion attack is on the learned models during the testing phase producing adversary-selected outputs. Through an Evasion attack, the adversary can pass through the test process by altering the test samples, and the model results in incorrect output (Liang *et al*., 2019). The evasion attacks can be classified into three types. A black box attack is the most frequently used attack type where the attacker will have zero knowledge about the ML/DL models. In a white box attack, the hacker has a permit to access the parameters of the prototype, whereas, in the grey box model, the attacker has moderate knowledge about the model (Dixit and Silakari, 2021; Taheri *et al*., 2020). The testing phase attacks are Deep Fool, Fast Gradient Sign Method (FGSM), Optimization-based method, Jacobian-based Saliency Map Approach (JSMA), etc., (Chaudhary *et al*., 2020). The thwarting techniques for attacks on the ML models are categorized into four types: Security Assessment mechanisms, counteractant in the training and testing stage, Data Security, and Privateness. Some examples of defensive techniques are Adversarial Training, Ensemble Method, Data Deduplication, Secure Data Deduplication, Data Sensitization, Reject on Negative Impact (RONI), Identity Based Encryption, Defense Distillation, Differential Privacy, Blockchain Based Solution, Homomorphic Encryption, etc., (Chaudhary *et al*., 2020).

Guo *et al*., (2021) also propose a black box attack method for models which detect anomaly network flow using machine learning algorithms. The proposed Black Box adversarial example generation method uses the White box attack on the substitute model. The target model and substitute model are trained identically on the KDD99 dataset and the CSE-CICIDS2018 dataset. The attacker can launch an attack on the substitute model with the white box method. These crafted adversarial examples are then used in the target model to check whether these adversarial examples can misclassify the target model. Experiment results showed that the authors effectively generated adversarial examples based on network flow, which can mislead the detection models that are machine learning-based.

In general, adversaries use Adversarial Machine Learning Algorithms (AML), so that the machine learning algorithms misclassify the benign sample. The main reason is to make a machine learning model malfunction. For this adversary use poison data. This data may be to exploit particular vulnerabilities and compromise the outcomes. Some of the AML models are Droid API Miner, Mystique, Pin droid, and Droid Chameleon which reduce the detection rate of classification of Machine Learning models (Taheri *et al*., 2020). The adversarial classification can be False negative or False positive. In False positive, the attacker wrongly calculates a negative instance to classify it as positive. In contrast, in a false negative, the benign data is added with Malware so it can bypass the detection (Taheri *et al*., 2020).

Taheri *et al*. (2020) propose Ant Colony Optimization (ACO) algorithm to produce poison malware samples. In this approach, a linear regression algorithm is applied first to choose the malware instances almost identical to the benign examples in the training dataset. Next, the ACO Function is utilized to find the adversary sample data. The ACO pheromone value used is the number of features changed. The algorithm starts with one feature and the new samples are produced by modifying the malware samples without attributes present in legitimate applications. It is repeated by utilizing more additional features. The distance between the recently generated sample and the discriminator is estimated. If it is within the specified Malware and the discriminator range, the newly generated sample is added to recently developed samples; otherwise, discard this sample. The feature values are changed and the distance is recalculated. This is repeated until the maximum iteration, or the classifier misclassifies malware samples.

The domain names generated by DGA are generally detected by extracting the features of DNS traffic and statistical characteristics of the domain name language. Later, the ML algorithms analyze the extracted features to identify and classify the DGA domain names (Chen *et al*., 2021). The authors (Chen *et al*., 2021) used a deep neural LSTM network to propose a DGA domain name detection model. Li *et al*. (2019) presented a model to handle DGA threats as conventional malware control approaches (like blacklisting) cannot handle them. The paper focuses on the machine learning framework which can identify and detect DGA attacks. It also proposes the Deep Learning technique (DNN) to organize those large numbers of domain names. This study presents the machine learning framework with a two-level model and prediction model. In the first level of classification, the paper identifies Decision Tree-J48 as the best classifier among NB, ANN, LR, SVM, RF, and Gradient Boosting Tree (GBT), to classify DGA domains. The DT-J48 classification algorithm worked with high accuracy and minimum classification time. The framework uses the DBSCAN algorithm for second-level clustering, which is a density-based clustering. As the HMM model performs well with a quick run time and elevated match accurateness, it is used to analyze the clustering results. When compared with the DT-J48 classification algorithm, the DNN model works better to classify large datasets. The research work extrapolates that deep learning algorithms perform better when compared with machine learning algorithms to classify large data sets.

Nowadays, ML approaches are susceptible to adversarial instances through Generative Adversarial Networks (GANs). It is an unsupervised ML technique that combines a generator and a discriminator (Gümüşbaş *et al*., 2020; Rao *et al*., 2020). GAN poses severe problems for Cybersecurity applications that are security-critical. More work is required to study the effect of adversarial examples in Cybersecurity. The generator produces data from the random distribution which could easily be mistaken for real data and a segregator (discriminator) separates real data from the false data. They learn the data distribution through unsupervised methods (Gümüşbaş *et al*., 2020). The generator is a convolutional neural network and the discriminator is a DE convolutional neural network. The data produced by the generator is matching to the probability distribution of training data. Whereas the discriminator distinguishes the training data from the generated data (Rao *et al*., 2020). The generated samples can increase the detection performance. GANs can be used in addressing Missing Data Problems (Ren and Xu, 2019), to generate negative samples to satisfy the negative samples which are needed to train deep networks.

Zhang *et al*. (2020) have suggested a Brute-Force Black-Box method to launch an invasion of systems that work with Machine Learning. The proposed method detects Network Intrusion Detection (NIDS) since ML techniques are vulnerable to adversarial examples. The Brute Force Attack Method (BFAM) framework evaluates the resilience of the ML classifiers in detecting cyber security. It uses the confidence scores from the target classifiers to develop the adversarial examples so that BFAM can be used for other adversarial invasions in cyber security. To utilize the excellent performance of Wasserstein GAN (WGAN), authors used this in their GAN model. Other GAN models such as MalGAN by Kim *et al*. (2018), and IDSGAN by Lin *et al*. (2020) are capable of generating adversarial malware, which misleads malware detection systems based on ML.

Table 7 summarizes the various ML algorithms discussed in this section. ML techniques can be used efficiently for defending against Cyberattacks, moreover, ML-based systems used offensively against all types of attacks. Kamoun *et al*. (2020) studied various AI/ML models for cyber security defense. The authors also list the misuse of AI/ML itself for Cyber security threats. Generally, AI/ML models, frameworks, and tools are available as open source, the hackers can easily adapt these models for their benefit. The adversarial AI/ML-based attack models are featured with speed, automation, scale, and sophistication. Based on the activities/actions, (Kamoun *et al*. 2020) categorize the AI/ML-powered cyberattacks into Probing, Scanning, Spoofing, Flooding, Misdirection, Execution of malicious processes, and Bypassing.

Nguyen and Armitage (2008) systematically explain the performance of ML algorithms differently for different applications of cyber security. The author concluded that it is better to use a combination of classification models (Nguyen and Armitage, 2008).

Deep Learning Solutions to Cyber Security

Deep Learning (DL) is considered a sub-category of ML that establishes a layered neural network to stimulate human intelligence for coherent thinking (Martínez *et al*., 2019; Hu and Tan, 2017). Deep learning algorithms have proved that they can overcome the constraints of machine learning algorithms. Deep learning algorithms benefit from traditional machine learning algorithms (Aslan and Yilmaz, 2021) where the high-level features are generated from existing features automatically.

DL algorithms lower the requirement for feature engineering and feature space. It can perform well on supervised, unsupervised, and semi-supervised learning efficiently. DL algorithms process enormous datasets and they can handle unstructured data efficiently. DL algorithms play a vital role in solving problems in various research domains: Image processing, Bioinformatics, Game playing, Speech recognition, Object detection, Segmentation, Classification, Pattern recognition, and matching, Customer Relationship Management automation,

Vehicle automation system, etc., (Karatas *et al*., 2018; Mahdavifar and Ghorbani, 2019; Rodriguez *et al*., 2021).

The deep learning techniques' robustness, rapidness, accuracy, and ability to handle extensive data have drawn researchers' concentration in recent years.

Deep Learning (DL) algorithms efficiently detect advanced cyber security threats. It is evident that DL techniques can be used for cybersecurity problems. Deep Learning algorithms can identify known and unknown attacks, it can manage incomplete, inconsistent, and composite data (Geluvaraj *et al*., 2019). The authors (Lakshminarayana *et al*., 2019; Kim and Aminanto, 2017) studied various DL algorithms and then classified DL algorithms into Generative (Unsupervised), Discriminative (Supervised), and Hybrid. Table 8 describes some of the DL techniques under these categories (Sarker, 2022).

Aslan and Yilmaz (2021) suggested a framework for DL models and explored the utilization of deeply layered learning models for detecting several cybersecurity problems such as Intrusion, Malware, Spam Phishing, and Website Defacement. The authors used generative deep learning models over discriminative or hybrid approaches. Authors have highlighted the advantages of semi-supervised learning for unlabeled data.

Deep Boltzmann Machine (DBM), DBN, CNN, Restricted Boltzmann Machine (RBM), Deep Neural Network (DNN), Deep Reinforcement Learning (DRL), Generative Adversarial Network (GAN,) Stacked Autoencoder (SAE), LSTM (Long Short-Term Memory), RNN, Deep Auto Encoder(DAE), Deep feedforward neural network and combination learners are some of the Deep learning mechanisms which are useful for the cyber-attack detection (Gümüşbaş *et al*., 2020; Ferrag *et al*., 2020; Mahdavifar and Ghorbani 2019; Dixit and Silakari, 2021; Sohn, 2021).

Compared with classic ML techniques, deep networks can acquire the features automatically from data, reducing the effort of pre-processing the input data and not relying on human-engineered features. This makes Deep learning algorithms fit for much real-time processing. But DL algorithm's performance declines if the algorithms are not provided with sufficient numbers of appropriate training data (Mahdavifar and Ghorbani, 2019; Yu *et al*., 2019). The existing machine learning techniques do not scale over a huge volume of data and detecting cyberattacks in large loosely coupled devices is a great challenge. It is observed that ML techniques are inefficient in detecting intrinsic attacks or unidentified malware and are very poor in preserving users' privacy (Sapre *et al*., 2021).

The DL methods can overcome the drawbacks of ML models for existing cyber security solutions. DL has the potential at handling complex patterns and builds robust and reliable models. The DL techniques are faster and more accurate in processing since it has self-learning capabilities that improve the processing speed as well as the accuracy of the applications (Imamverdiyev and Abdullayeva, 2020). DL methods are suitable for Malware Detection, Network Intrusion Detection, DDoS attacks, Phishing/Spam Detection, Behavior Anomaly Detection, Botnet Detection, and Website Defacement Detection (Chen, 2020).

Lee *et al*. (2019) used various Artificial Neural Network methods, such as CNN, LSTM, and FCNN to develop an Artificial Intelligence-Security Information and Event Management (AI-SIEM). The proposed model can discriminate between true positive and false positive notifications. This model enables Cyber security analysts to identify cyber threats and defend against them quickly. The author inferred that AI-SIEM has relevance in learning-based network intrusion detection models. They also concluded that multiple deep learning approaches could be efficiently used to enhance the threat predictions to avoid cyber-attacks.

The research by Xin *et al*. (2018); Karatas *et al*. (2018) highlights the differences between DL and ML techniques used for cybersecurity. DL algorithms perform well when large data volume is available and it requires high-performance machines with GPUs which are not applicable to ML algorithms. In Machine Learning, feature extraction is done by an expert wherein in Deep Learning, the algorithm tries to automatically extract the features. The ML algorithm's performance is gauged on the accuracy of the extracted features which is not the case in DL algorithms. With respect to problem-solving methods, ML divides the problem into sub problems and then solves those sub problems whereas DL algorithms do end-to-end problem-solving. The training period is more in DL models but the testing time is very less compared to ML algorithms. This is reversed in the case of ML models. Machine Learning algorithms can work on any normal CPU, but to run the Deep Learning algorithms high performance machines are required. Manual feature extractions are done in the ML approaches, whereas deep learning algorithms automatically extract abstract and flexible features by generalization in classification (Mahdavifar *et al*., 2020). Hossain *et al*. (2020) in their work demonstrated that the LSTM Deep Learning approach outperforms the Machine Learning classifiers like J48, RF, KNN, NB, DT, and algorithms to detect FTP and SSH brute-force invasions effectively.

Ferrag *et al*. (2020) performed an exhaustive investigation on intrusion discovery systems, datasets, and also a comparative analysis of various DL models. The authors used Deep Learning strategies like DNN, RNN, RBM, CNN, DBN, DBM, and DA for detecting Intrusions such as Brute Force, DoS, DDOS, SQL Injection, and Botnet attack and compared with different machine learning approaches like RF, NB, SVM, ANN concerning global detection rate. DBM, RNN, and CNN

are the DL models incorporated for detecting network-based intrusion. Karatas *et al*. (2018) listed out components involved in IDSs to enhance network security. The components of IDS are data collection, feature selection, and decision engine. The third component is the critical one where the collected data is classified as benign or malicious based on previous knowledge.

It isn't easy to find anomalous features in the extensive network traffic samples. Feedforward neural network autoencoders are best suited for network anomaly detection since it is simple to train the input and reconstruct the output (Xu *et al*., 2021). Autoencoders are an unsupervised neural network learning approach. Autoencoders reduce dimensionality by compressing input data and rebuilding output data from their representation. It can discover structure within data to develop a compressed input representation. Xu *et al*. (2021). presented a novel Autoencoder-based method consisting of five layers for detecting anomalous traffic in the network. The approach transforms the input dataset into balanced datasets concerning data size and data types by removing outliers and avoiding bias in anomaly detection. In the 5-layer architecture, the hidden layer has the optimized count of neurons and the latent space layer provides the best performance compared to other architectures.

Autoencoders can also be used for feature learning and feature extraction. Authors (Andresini *et al*., 2020) used deep feature learning with multi-channel to detect intrusion in the system. The MINDFUL (MultI-chanNel Deep FeatUre Learning) framework uses Autoencoders. The Autoencoders are implemented by Hindy *et al*. (2020) for detecting zero-day attacks. This study tries to overcome the drawbacks of outlier-based zero-day detection, which has high false-negative rates. The authors built an IDS model to reduce the false negative rate (i.e., miss rate) with high recall (i.e., true-positive rate). The authors used the CICIDS2017 and NSL-KDD datasets. They remarked on an excellent accuracy rate compared to the One-Class Support Vector Machine (SVM).

An unsupervised Stacked Auto Encoder (SAE) is combined with weighted feature selections (Kim and Aminanto, 2017) to improve the feature learning process for IDS. The authors described that SAE is efficient and valuable for Feature Extraction, Clustering, and Classification mechanisms. The authors used SAE for classification and clustering. The results are validated using the Aegean Wi-Fi Intrusion Dataset (AWID) consisting of benign, injection, impersonation, and flooding classes. The authors concluded that IDS that used SAE as a classifier resulted in a low impersonation detection rate. Thus, SAE could be used as a classifier rather than a feature extractor.

An exhaustive investigation of deep learning-based intrusion detection is proposed by (Otoum *et al*., 2019). The proposed work uses the Adaptively Supervised and Clustered Hybrid (ASCH-IDS) methodology (Otoum *et al*., 2018). This intrusion detection model, Restricted Boltzmann-based Clustered IDS (RBC-IDS), is for Wireless Sensor Networks-based critical applications. The results showed that ML-based IDS is desirable when it resembles DL-based IDS concerning the accuracy, training, and testing time for WSN-based critical infrastructure monitoring. The research work in (Alom *et al*., 2015) performed a series of experiments on Intrusion detection using DBN. With these experiments, (Alom *et al*., 2015) could identify unknown attacks and, after 50 iterations, achieved 97.5% of accuracy.

Djellali *et al*. (2019), designed two deep learning techniques as Batch Gradient Descent and Stochastic Gradient Descent which are compared and tested on a resampling method for cybersecurity. Batch Gradient Descent is an iterative technique that uses complete input training patterns in order to optimize a cost function. In Stochastic Gradient Descent, the input training patterns are randomly selected to update the weights. The author concluded that Stochastic Gradient Descent provides an efficient optimization algorithm for cybersecurity with a good performance and less computational costs.

Sohn (2021) proposed a survey paper that describes the basics of the DBN-based intrusion detection model. The author compares the fundamental algorithms, the different training methods, and the data sets and interprets the results of various research works starting from 2016. Intrusion detection based on DBN used ADFA, NSL-KDD, UNSW-NB15, and KDD Cup 99 dataset. The DBN-IDS-based framework consists of components such as a data preprocessor training, Classifier, Optimizer, and fine-tuning algorithm (Sohn, 2021).

The malicious activities are detected from the network traffic using anomaly detection. Many deep-learning techniques have been proposed for anomaly detection. To detect anomalies, (Kim *et al*., 2018) designed a cluster of approaches established on Variational Autoencoder (VAE), Fully Connected Network (FCN), and LSTM Seq2Seq structures and concluded that Deep learning methods are a proper selection for convincing network anomaly detection. The authors examined the proposed architectures with various public traffic datasets, including IDS2017, UNSW-NB15, Kyoto-Honeypot, and NSL-KDD. In data preprocessing, numerical features are normalized using a z-score, and categorical features are turned to numerical by one-hot encoding-the preprocessed data fed into a connected network for training. The authors considered ReLU as the activation function in hidden layers. The Softmax layer produces the final output with a cross-entropy cost function which can be either normal or attack. Next, the two variants of VAE models such as VAE-Pure and VAEFCN models are tested. The original data and the detected data are compared to calculate the loss. The LSTM-Seq2Seq

model is based on RNN which yields a target sequence and conditional probability through an encoder and decoder. LSTM Seq2Seq structure showed a promising result of 99% of binary classification accuracy on both the NSL-KDD dataset and Kyoto University Honeypot data ("Kyoto-Honeypot"). Results of SVM and RF show less accuracy when classified with the NSL KDD dataset and high accuracy with the UNSW-NB15 dataset.

Maimó *et al*. (2018) proposed a two-level DL model, which acts as a robust system for detecting anomalies and defending against cyber-attacks in a 5G architecture for the mobile network. The supervised or semi-supervised learning method is used in the first level to implement a DBN or an SAE operating on every RAN. supervised LSTM Recurrent Network is used in the second level to confine the cyberattacks.

For anomaly detection with multi-dimensional input, a little investigation has been done by employing Convolutional Neural Networks (CNN) (Alabadi and Celik, 2020). Though Deep learning mechanisms are best suited for anomaly detection, the challenges faced are to identify the threats faster and the traffic profile should be auto-profiled. The traffic profile includes flow statistics such as transmission rate, packet count, flow size, etc. In CNN these kinds of features are automatically extracted from the traffic profile. Hwang *et al*. (2020), the traffic patterns are built by investigating the starting bytes of the first few traffic packets. As it uses only the first few packets for anomaly detection, the speed of threat detection is increased. The proposed system automatically uses the CNN module to know the source data's features. The model achieves 99.77% accuracy in detecting malicious activities and less than 1% FNR and FPR. The dataset comprises four DDoS attack classes: HTTP flood, ACK flood, UDP flood, and SYN flood (Hwang *et al*., 2020).

A Machine Learning approach that combines deep learning approaches with Reinforcement Learning (RL) is named Deep Reinforcement Learning (DRL). Apruzzese *et al*. (2020) used DRL mechanisms in their work to propose a design approach that protects botnet detectors from adversarial attacks. The novel strategy leverages DRL to improve the robustness of detectors. Botnet detectors use the classifiers Wide and Deep (WnD) and Random Forest (RF). The agents in the proposed model are based on deep reinforcement learning approaches such as Double Deep Q-Network(2DQN) and Deep State-action-reward-state-action (Sarsa), which use off-policy and on-policy methods, respectively. In the next phase, this trained DRL agent produces the adversarial attack. These samples can evade a botnet detector. With adversarial training, the model utilizes the samples for hardening the botnet detectors.

Deep Neural Networks are most suited for Domain Generation Algorithms as they can efficiently classify domain names as malicious and benign (Yu *et al*., 2019). For DGA detection, the authors examine the advantages of labeled data to train DL classifiers. For this, the authors used RNNs, LSTMs, CNN, and hybrid CNN/RNN models. Shahzad *et al*. (2021) used RNN architectures like Bidirectional LSTM (Bi-LSTM), Long Short-Term Memory networks (LSTMs), and Gated Recurrent Units (GRU) to calculate the performance of a DGA classifier. The suggested DGA classifier takes the domain names from the DNS queries and does not demand manual feature creation. Without any contextual information, the model performs multiclass classification to determine the domain family to which it belongs.

When compared to ML algorithms, DL algorithms are most suited for malware detection (Apruzzese *et al*., 2018; Li *et al*., 2020). The reason is the diminishing output of ML algorithms when the data size increases. DL algorithms enhance the performance though the input size is more. As Malwares are multiplying with the technology, malware detection should cope with the scalability issues. Vinayakumar *et al*. (2019) proposed a hybrid scalable deep learning framework named as Scale Mal Net, which handles large samples of malware. The model collects the malware samples and applies them to pre-process in a distributed way. The executable files are classified into benign or Malware samples using static and dynamic examination in the first phase. This is followed by a second stage where the malware executable files are separated into their families. However, the robustness of the DL techniques is not focused on the work; the authors conclude that the deep learning architectures outperform the classical machine learning models.

To classify Malware, authors (Aslan and Yilmaz, 2021) proposed a framework with a hybrid deep neural network. This hybrid approach combines several pre-trained network models and the test results proved that the suggested framework could segregate Malware with increased precision, recall, accuracy, and F score.

A framework for Malware category classifications for Android is performed (Mahdavifar *et al*., 2020). This framework uses dynamic malware category classification and also applies semi-supervised deep neural networks. The experiment results show that the F1 score is better and has a false positive rate of 2.76% outperforming the typical machine learning algorithms. The input layer consisted of 470 neurons and the output layer consisted of 5 neurons. The sigmoid function is used for activation and for optimization they use mini–Batch Gradient Descent.

Like Machine Learning algorithms, deep learning methods also get affected by adversarial attacks. Deep learning models are fragile under adversarial attacks (Li *et al*., 2020). The adversarial attacks can be gray-box, white-box, and black-box attacks. Many attack algorithms are proposed for adversarial sample generation for these threat models. Some of the attack algorithms are the Deep Fool, Fast Gradient Sign Method (FGSM), Optimization-

based method, Jacobian-based Saliency Map Approach (JSMA), Limited-memory Broyden-Fletcher Goldfarb-Shanno (L-BFGS) algorithm, the Basic Iterative Method (BIM)/Projected Gradient Descent (PGD), Carlini and Wagner (C&W) attacks and Distribution Ally Adversarial attack (Ren *et al*., 2020; Li *et al*., 2021).

Intrusion Detection Systems based Deep Learning Neural Networks are susceptible to attacks on white-box and backdoor adversarial scenarios (Alrawashdeh and Goldsmith, 2020). Much research work has been undergone in this field. One such work is investigating the adversarial examples affecting the interpretation of Intrusion Detection Systems using Deep Neural Networks (DNN) (Yang *et al*., 2018). The author illustrates that the adversary can generate adversarial examples to mislead the DNN model even though the models' internal information is isolated from the adversary. These adversarial examples are generated and evaluated in the black-box model. Though the internal details of the model are not accessed here, the adversary can still mislead the classifier to misclassify the attack input as normal input.

Shi and Sagduyu (2017) proposed a Machine learning classifier for generating and defending against evasion and causative attacks, combining the DL-based exploratory attack. Initially, the adversary creates a classifier using an exploratory attack established on Deep Learning (DL), similar to the original classifier. From the built classifier, the samples are collected and given to the original classifier. To achieve an evasion attack in the trained classifier, the adversary tries to deceive the machine learning algorithm by providing incorrect input data, which results in the wrong label, thus misclassifying the samples. For the causative attack, the adversary provides the target classifier with false class information, thus reducing the precision of the trained classifier. This study by the authors demonstrated that the evasion attack increased the error in the test phase and the causative attack increased the same during the training phase. They concluded the work by providing an aggressive defense mechanism with small perturbations showing that the error under attack is identical to the error when there is no attack.

Li and Li (2020) propose a mixture of attacks to produce adversarial malware examples. For this author uses multiple generative procedures and manipulation sets. To validate the malware detectors' robustness, the author uses 26 evasion attacks. These evasion attacks are categorized into gradient-based, gradient-free, transfer attack, obfuscation, and a mixture-of-attack approaches. Table 9 summarizes the various DL algorithms discussed in this section.

Performance Metrics

The essential metrics estimating the performance of DL and ML techniques are Confusion matrix, precision, Detection Rate (DR) (also called recall or true positive rate), false negative rate, false positive rate, true negative rate, F1-Score, accuracy (Chaudhary *et al*., 2020). The Receiver Operating Characteristics (ROC) curve and Area under the ROC (AuC) are also used to estimate the classification performance (Sohn, 2021).

In a dataset of random size, the component can belong to either binary or n-ary classification. In binary classification, the element can be considered as an attack or benign. The invasion is represented as positive and the benign category is denoted as negative (Vinayakumar *et al*., 2019). A True Positive (*TP*) is a component from the positive category that the algorithm treats as positive. Similarly, a True Negative (*TN*) is an element from the negative class that is treated correctly as negative by the algorithm. But in a False Positive (*FP*), the element is identified as an attack when in actuality, it isn't. Similarly, in False Negatives (*FN*), the algorithm fails to identify the attack.

Accuracy is measured as the fraction of elements that are correctly predicted:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision (Positive Predictive Value) is the fraction of elements that are predicted correctly to the overall predicted attacks. This identifies the number of attacks classified as positive:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

The Detection Rate (DR) reveals the count of attacks that are identified (Vinayakumar *et al*., 2019; Jayakumar *et al*., 2015):

$$DR = \frac{TP}{TP + FN} \quad (3)$$

The False Positive Rate (FPR) indicates the number of invasions that are not recognized:

$$FPR = \frac{FP}{TN + FP} \quad (4)$$

The recall is the calculated percentage of rightly classified attack data to the total count of attack data in a provided dataset-the more the recall rate, the better the machine learning model's performance:

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

F1-Score/F1-Measure is calculated as the harmonic mean of Precision and Recall. The increased rate of F1-Score illustrates that the machine learning algorithm is accomplished excellently:

$$F1-Score = 2X\frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

The high value of the false negative rate may demonstrate that the NIDS failed to identify known or anonymous attacks. In contrast, the increased false positive rate indicates the false alarms generated when there is no attack in the network (Kilincer *et al*., 2021)

Some of the metrics used in Generating AEs are the Total Time Cost (TTC), Adversarial Detection Rate (ADR), and Original Detection Rate (ODR) (Zhang *et al*. 2020). TTC is the total time required to build a set of AEs. The ODR determines the detection performance of the target classifiers contrary to the actual attack examples. The ADR implies the detection performance of the target classifiers contrary to the adversarial attacks:

$$ODR = \frac{No.of\ right\ indentified\ orignal\ attack\ examples}{No.of\ all\ the\ orignal\ attack\ examples} \quad (7)$$

$$ADR = \frac{No.of\ right\ indentified\ adversarial\ attack\ examples}{No.of\ all\ the\ asversarial\ attack\ examples} \quad (8)$$

Conclusion

Nowadays, cyber security attacks are increasing tremendously. The prevailing cyber security attacks are DOS attacks, Phishing, Malware attack, Botnet Evasion Attacks, Spoofing, R2L, Probing attacks, and U2R attacks. This survey paper details the different cyber security attacks and tools for detecting intrusion detection mechanisms. The paper also identifies cyber security domains and significant research challenges. Many traditional approaches are inefficient in detecting, analyzing, and defending against cyber-attacks. In current years, it has been evident that ML and Deep Feature Learning approaches efficiently solve cyber security attacks. This study reviewed several efficient algorithms of Machine and Deep feature learning to solve many cyber security problems. The article also addresses the adversarial attacks on Machine Learning Algorithms and Deep Learning Algorithms and the defense mechanisms against those adversarial attacks. The survey gives insights into private and publicly available datasets that are significant in analyzing the effectiveness of the proposed algorithms to defend against cyber security threats. The paper concludes with the various performance matrices utilized to estimate the efficiency of the suggested algorithms.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

Manjula M: Conceived ideas and designed the outline of the manuscript, collected the relevant data, papers from different source. Wrote review on each paper emphasizing on the concepts, proposed algorithms, limitations in the proposed approaches. Drafted the manuscript and designed the figures. Explored the research issues and challenges in cyber security. Examined the Machine Learning and Deep Learning approaches for the cyber security issues. Tabulated the different research work.

Venkatesh: Developed the framework for the paper, identified key research issues and challenges in cyber security, theory. Performed comprehensive review of papers on cyber security, approaches, design techniques. A complete review of manuscript, verified the different methods. Encouraged research scholar to investigated Machine Learning, Deep Learning based approaches for cyber security solutions and supervised the findings of this study.

Venugopal K. R.: Reviewed the manuscript, provided critical feedback and helped to shape the research, analysis of the manuscript. Verified the analytical methods. Motivated research scholar to explore Machine Learning, Deep Learning algorithms used for cyber security and supervised the findings of this study. Provided valuable feedback to enhance the quality of the work.

Ethics

I undersigned that this article has not been published elsewhere. The authors declare no conflict of interest.

References

Akazaki, T., Liu, S., Yamagata, Y., Duan, Y., & Hao, J. (2018, July). Falsification of cyber-physical systems using deep reinforcement learning. *In International Symposium on Formal Methods* (pp. 456-465). Springer, Cham https://doi.org/10.1109/TSE.2020.2969178

Alabadi, M., & Celik, Y. (2020, June). Anomaly detection for cyber-security based on convolution neural network: A survey. *In 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications* (HORA) (pp. 1-14). IEEE. https://doi.org/10.1109/HORA49412.2020.9152899

Alam, M. N., Sarma, D., Lima, F. F., Saha, I., & Hossain, S. (2020, August). Phishing attacks detection using *international conference on smart systems and inventive technology* (ICSSIT) (pp. 1173-1179). IEEE. https://doi.org/10.1109/ICSSIT48917.2020.9214225

Albasir, A., James, R. S. R., Naik, K., & Nayak, A. (2018, April). Using deep learning to classify power consumption signals of wireless devices: An application to cybersecurity. *In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP) (pp. 2032-2036). IEEE. https://doi.org/10.1109/ICASSP.2018.8461304

Al-Janabi, M., & Altamimi, A. M. (2020, November). A comparative analysis of machine learning techniques for classification and detection of malware. *In 2020 21st International Arab Conference on Information Technology* (ACIT) (pp. 1-9). IEEE. https://doi.org/10.1109/ACIT50332.2020.9300081

Alom, M. Z., Bontupalli, V., & Taha, T. M. (2015, June). Intrusion detection using deep belief networks. *In 2015 National Aerospace and Electronics Conference* (NAECON) (pp. 339-344). IEEE. https://doi.org/10.1109/NAECON.2015.7443094

Alrawashdeh, K., & Goldsmith, S. (2020, October). Optimizing Deep Learning Based Intrusion Detection Systems Defense Against White-Box and Backdoor Adversarial Attacks Through a Genetic Algorithm. *In 2020 IEEE Applied Imagery Pattern Recognition Workshop* (AIPR) (pp. 1-8). IEEE. https://doi.org/ 0.1109/AIPR50011.2020.9425293

Andrade, R. O., Ortiz-Garcés, I., & Cazares, M. (2020, July). Cybersecurity attacks on Smart Home during Covid-19 pandemic. *In 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability* (WorldS4) (pp. 398-404). IEEE. https://doi.org/10.1109/WorldS450073.2020.9210363

Andresini, G., Appice, A., Di Mauro, N., Loglisci, C., & Malerba, D. (2020). Multi-channel deep feature learning for intrusion detection. *IEEE Access*, 8, 53346-53359. https://doi.org/10.1109/ACCESS.2020.2980937

Apruzzese, G. Andreolini, M., Marchetti, M., Venturi, A., & Colajanni, M. (2020). Deep reinforcement adversarial learning against botnet evasion attacks. *IEEE Transactions on Network and Service Management*, 17(4), 1975-1987. https://doi.org/10.1109/TNSM.2020.3031843

Apruzzese, G., Colajanni, M., Ferretti, L., Guido, A., & Marchetti, M. (2018, May). On the effectiveness of machine and deep learning for cyber security. *In 2018 10th international conference on cyber Conflict* (CyCon) (pp. 371-390). IEEE. https://doi.org/10.23919/CYCON.2018.8405026

Arp, D., Spreitzenbarth, M., Hubner, M., Gascon, H., Rieck, K., & Siemens, C. E. R. T. (2014, February). Drebin: *Effective and explainable detection of android malware in your pocket.* In Ndss (Vol. 14, pp. 23-26). https://doi.org/10.14722/ndss.2014.23247

Aslan, Ö., & Yilmaz, A. A. (2021). A new malware classification framework based on deep learning algorithms. *Ieee Access*, 9, 87936-87951. https://doi.org/10.1109/ACCESS.2021.3089586

Bahl, S., & Sharma, S. K. (2015, May). Detection rate analysis for user to root attack class using correlation feature selection. *In International Conference on Computing, Communication and Automation* (pp. 66-71). IEEE. https://doi.org/10.1109/CCAA.2015.7148345

Baptista, I., Shiaeles, S., & Kolokotronis, N. (2019, May). A novel malware detection system based on machine learning and binary visualization. *In 2019 IEEE International Conference on Communications Workshops* (ICC Workshops) (pp. 1-6). IEEE. https://doi.org/10.1109/ICCW.2019.8757060

Begli, M., Derakhshan, F., & Karimipour, H. (2019, August). A layered intrusion detection system for critical infrastructure using machine learning. *In 2019 IEEE 7th International Conference on Smart Energy Grid Engineering (SEGE)* (pp. 120-124). IEEE. https://doi.org/10.1109/SEGE.2019.8859950

Berman, D. S., Buczak, A. L., Chavis, J. S., & Corbett, C. L. (2019). A survey of deep learning methods for cyber security. *Information*, 10(4), 122. https://doi.org/10.3390/info10040122

Buczak, A. L., & Guven, E. (2015). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications *surveys & tutorials*, 18(2), 1153-1176. https://doi.org/10.1109/COMST.2015.2494502

Caprolu, M., Di Pietro, R., Raponi, S., Sciancalepore, S., & Tedeschi, P. (2020). Vessels cybersecurity: Issues, challenges, and the road ahead. *IEEE Communications Magazine*, 58(6), 90-96. https://doi.org/10.1109/MCOM.001.1900632

Chaabouni, N., Mosbah, M., Zemmari, A., Sauvignac, C., & Faruki, P. (2019). Network intrusion detection for IoT security based on learning techniques. *IEEE Communications Surveys & Tutorials*, 21(3), 2671-2701. https://doi.org/10.1109/COMST.2019.2896380

Chahira, J. M. (2020). Model for intrusion detection based on hybrid feature selection techniques. *International Journal of Computer Applications Technology and Research*, 9, 115-124. https://doi.org/10.7753/IJCATR0903.1005

Chandre, P. R., Mahalle, P. N., & Shinde, G. R. (2018, November). Machine learning based novel approach for intrusion detection and prevention system: A tool based verification. *In 2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN)* (pp. 135-140). IEEE. https://doi.org/10.1109/GCWCN.2018.8668618

Chaudhary, H., Detroja, A., Prajapati, P., & Shah, P. (2020, December). A review of various challenges in cybersecurity using artificial intelligence. *In 2020 3rd International Conference on Intelligent Sustainable Systems* (ICISS) (pp. 829-836). IEEE. https://doi.org/10.1109/ICISS49785.2020.9316003

Chen, Y., Pang, B., Shao, G., Wen, G., & Chen, X. (2021). DGA-based botnet detection toward imbalanced multiclass learning. *Tsinghua Science and Technology*, 26(4), 387-402. https://doi.org/10.26599/TST.2020.9010021

Chen, Z. (2020, August). Deep Learning for Cybersecurity: A Review. *In 2020 International Conference on Computing and Data Science (CDS)* (pp. 7-18). IEEE. https://doi.org/10.1109/CDS49703.2020.00009

Collier, Z. A., Linkov, I., & Lambert, J. H. (2013). Four domains of cybersecurity: A risk-based systems approach to cyber decisions. *Environment Systems and Decisions*, 33(4), 469-470. https://doi.org/10.1007/s10669-013-9484-z

Creech, G., & Hu, J. (2013, April). Generation of a new IDS test dataset: Time to retire the KDD collection. *In 2013 IEEE Wireless Communications and Networking Conference* (WCNC) (pp. 4487-4492). IEEE. https://doi.org/10.1109/WCNC.2013.6555301

de la Torre, I., García-Zapirain, B., & López-Coronado, M. (2017). Analysis of Security in Big Data Related to Healthcare. *Journal of Digital Forensics, Security* and Law, 12(3), 5. https://doi.org/10.15394/jdfsl.2017.1448

Dhayal, H., & Kumar, J. (2018, April). Botnet and p2p botnet detection strategies: A review. *In 2018 International Conference on Communication and Signal Processing (ICCSP)* (pp. 1077-1082). IEEE. https://doi.org/10.1109/ICCSP.2018.8524529

Dixit, P., & Silakari, S. (2021). Deep learning algorithms for cybersecurity applications: A technological and status review. *Computer Science Review*, 39, 100317. https://doi.org/10.1016/j.cosrev.2020.100317

Djellali, C., Adda, M., & Moutacalli, M. T. (2019, August). A comparative study to deep learning for pattern recognition, by using online and batch learning; taking cybersecurity as a case. *In 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 907-912). IEEE. https://doi.org/10.1145/3341161.3343533

Dramilio, A., Faustine, C., Sanjaya, S., & Soewito, B. (2020, August). The effect and technique in search engine optimization. *In 2020 International Conference on Information Management and Technology (ICIMTech)* (pp. 348-353). IEEE. https://doi.org/10.1109/ICIMTech50083.2020.9211171

Duić, I., Cvrtila, V., & Ivanjko, T. (2017, May). International cyber security challenges. *In 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics* (MIPRO) (pp. 1309-1313). IEEE. https://doi.org/10.23919/MIPRO.2017.7973625

Elsayed, M. S., Le-Khac, N. A., Dev, S., & Jurcut, A. D. (2020, October). Detecting abnormal traffic in large-scale networks. *In 2020 International Symposium on Networks, Computers and Communications (ISNCC)* (pp. 1-7). IEEE. https://doi.org/10.1109/ISNCC49221.2020.9297358

Farooq, Q., Shan, L., & Yuan, H. Y. (2021, April). New Approach Towards Cyber Security for Nuclear Power Control System. *In 2021 IEEE Kansas Power and Energy Conference* (KPEC) (pp. 1-6). IEEE. https://doi.org/10.1109/KPEC51835.2021.9446235

Feng, Y., Akiyama, H., Lu, L., & Sakurai, K. (2018, August). Feature selection for machine learning-based early detection of distributed cyber attacks. *In 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress* (DASC/PiCom/DataCom/CyberSciTech) (pp. 173-180). IEEE. https://doi.org/10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00040

Ferrag, M. A., Maglaras, L., Janicke, H., & Smith, R. (2019, September). Deep learning techniques for cyber security intrusion detection: *A detailed analysis. In 6th International Symposium for ICS & SCADA Cyber Security Research 2019* (pp. 126-136). https://doi.org/10.14236/ewic/icscsr19.16

Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications*, 50, 102419. https://doi.org/10.1016/j.jisa.2019.102419

Garcia, S., & Uhlir, V. (2014). The CTU-13 dataset. a labeled dataset with botnet, normal and background *traffic*. S. Lab, Ed., ed. https://www.stratosphereips.org/datasets-ctu13.

Garcia, S., Grill, M., Stiborek, J., & Zunino, A. (2014). An empirical comparison of botnet detection methods. *computers and security*, 45, 100-123. https://doi.org/10.1016/j.cose.2014.05.011

Geluvaraj, B., Satwik, P. M., & Ashok Kumar, T. A. (2019). 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 (pp. 739-747). Springer, Singapore. https://doi.org/10.1007/978-981-10-8681-6_67

Ghurab, M., Gaphari, G., Alshami, F., Alshamy, R., & Othman, S. (2021). A detailed analysis of benchmark datasets for network intrusion detection system. *Asian Journal of Research in Computer Science*, 7(4), 14-33. https://doi.org/10.9734/ajrcos/2021/v7i430185

Gmcdouga, (2022). "Check Point Research: Cyber Attacks Increased 50Year." https://blog.checkpoint.com/2022/01/10/check-point-research-cyber-attacks-increased-50-year-over-year/

Goel, S., & Nussbaum, B. (2021). Attribution Across Cyber Attack Types: Network Intrusions and Information Operations. *IEEE Open Journal of the Communications Society*, 2, 1082-1093. https://doi.org/10.1109/OJCOMS.2021.3074591

Gordon, G. J. (1995). Stable function approximation in dynamic programming. In Machine learning proceedings 1995 (pp. 261-268). *Morgan Kaufmann*. https://doi.org/10.1147/rd.33.0210

Gümüşbaş, D., Yıldırım, T., Genovese, A., & Scotti, F. (2020). A comprehensive survey of databases and deep learning methods for cybersecurity and intrusion detection systems. *IEEE Systems Journal*, 15(2), 1717-1731. https://doi.org/10.1109/JSYST.2020.2992966

Guo, S., Zhao, J., Li, X., Duan, J., Mu, D., & Jing, X. (2021). A black-box attack method against machine-learning-based anomaly network flow detection models. *Security and Communication Networks*, 2021. https://doi.org/10.1155/2021/5578335

Gupta, B. B., Yadav, K., Razzak, I., Psannis, K., Castiglione, A., & Chang, X. (2021). A novel approach for phishing URLs detection using lexical based machine learning in a real-time environment. *Computer Communications*, 175, 47-57. https://doi.org/10.1016/j.comcom.2021.04.023

Hagos, D. H., Yazidi, A., Kure, Ø., & Engelstad, P. E. (2017, March). Enhancing security attacks analysis using regularized machine learning techniques. *In 2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA)* (pp. 909-918). IEEE. https://doi.org/10.1109/AINA.2017.19

Hamid, B., Jhanjhi, N. Z., Humayun, M., Khan, A., & Alsayat, A. (2019, December). Cyber security issues and challenges for smart cities: *A survey. In 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics* (MACS) (pp. 1-7). IEEE. https://doi.org/10.1109/MACS48846.2019.9024768

He, Z., Miari, T., Makrani, H. M., Aliasgari, M., Homayoun, H., & Sayadi, H. (2021, April). When machine learning meets hardware cybersecurity: Delving into accurate zero-day malware detection. *In 2021 22nd International Symposium on Quality Electronic Design (ISQED)* (pp. 85-90). IEEE. https://doi.org/10.1109/ISQED51717.2021.9424330

Hegde, M., Kepnang, G., Al Mazroei, M., Chavis, J. S., & Watkins, L. (2020, October). Identification of botnet activity in IoT network traffic using machine learning. *In 2020 International conference on intelligent data science technologies and applications (IDSTA)* (pp. 21-27). IEEE. https://doi.org/10.1109/IDSTA50958.2020.9264143

Hindy, H., Atkinson, R., Tachtatzis, C., Colin, J. N., Bayne, E., & Bellekens, X. (2020). Utilising deep learning techniques for effective zero-day attack detection. *Electronics*, 9(10), 1684. https://doi.org/10.3390/electronics9101684

Ho, S., Al Jufout, S., Dajani, K., & Mozumdar, M. (2021). A novel intrusion detection model for detecting known and innovative cyberattacks using convolutional neural network. *IEEE Open Journal of the Computer Society*, 2, 14-25. https://doi.org/10.1109/OJCS.2021.3050917

Hossain, M. D., Ochiai, H., Doudou, F., & Kadobayashi, Y. (2020, May). SSH and FTP brute-force attacks detection in computer networks: LSTM and machine learning approaches. *In 2020 5th International Conference on Computer and Communication Systems* (ICCCS) (pp. 491-497). IEEE. https://doi.org/10.1109/ICCCS49078.2020.9118459

Hu, W., & Tan, Y. (2017). Generating adversarial malware examples for black-box attacks based on GAN. arXiv preprint arXiv:1702.05983.

Huang, Y., Jiazhong, L., Tang, H., & Liu, X. (2021). A Hybrid Association Rule-Based Method to Detect and Classify Botnets. *Security and Communication Networks*, 2021. https://doi.org/10.1155/2021/1028878

Humayed, A., Lin, J., Li, F., & Luo, B. (2017). Cyber-physical systems security—A survey. *IEEE Internet of Things Journal*, 4(6), 1802-1831. https://doi.org/10.1109/JIOT.2017.2703172

Husák, M., Komárková, J., Bou-Harb, E., & Čeleda, P. (2018). Survey of attack projection, prediction and forecasting in cyber security. IEEE Communications *Surveys and Tutorials*, 21(1), 640-660. https://doi.org/10.1109/COMST.2018.2871866

Hussain, F., Abbas, S. G., Pires, I. M., Tanveer, S., Fayyaz, U. U., Garcia, N. M., ... & Shahzad, F. (2021). A Two-Fold Machine Learning Approach to Prevent and Detect IoT Botnet Attacks. *IEEE Access*, 9, 163412-163430. https://doi.org/10.1109/ACCESS.2021.3131014

Hwang, R. H., Peng, M. C., Huang, C. W., Lin, P. C., & Nguyen, V. L. (2020). An unsupervised deep learning model for early network traffic anomaly detection. *IEEE Access*, 8, 30387-30399. https://doi.org/10.1109/ACCESS.2020.2973023

Imamverdiyev, Y. N., & Abdullayeva, F. J. (2020). Deep learning in cybersecurity: Challenges and approaches. *International Journal of Cyber Warfare and Terrorism (IJCWT)*, 10(2), 82-105. https://doi.org/10.4018/IJCWT.2020040105

Iyer, P., Jadhav, T., & Pillai, A. (2021, June). Analysis of Modern Intrusion Detection Algorithms and Developing a Smart IDS. *In 2021 International Conference on Intelligent Technologies (CONIT)* (pp. 1-7). IEEE. https://doi.org/10.1109/CONIT51480.2021.9498519

Jayakumar, K., Revathi, T., & Karpagam, S. (2015). Intrusion Detection using Artificial Neural Networks with Best Set of Features. *International Arab Journal of Information Technology* (IAJIT), 12.

Cabaj, K., Kotulski, Z., Ksie˛zopolski, B., & Mazurczyk, W. (2018). "Cyberse-˙ curity: Trends, Issues, and Challenges," 2018. https://doi.org/10.1186/s13635-018-0080-0

Kadam, G., Parekh, S., Agnihotri, P., Ambawade, D., & Bhavathankar, P. (2020, November). An Approach to Reduce Uncertainty Problem in Network Intrusion Detection Systems. *In 2020 IEEE 15th International Conference on Industrial and Information Systems* (ICIIS) (pp. 586-590). IEEE. https://doi.org/10.1109/ICIIS51140.2020.9342634

Kambourakis, G., Kolias, C., & Stavrou, A. (2017, October). The mirai botnet and the iot zombie armies. *In MILCOM 2017-2017 IEEE Military Communications Conference (MILCOM)* (pp. 267-272). IEEE. https://doi.org/10.1109/MILCOM.2017.8170867

Kamoun, F., Iqbal, F., Esseghir, M. A., & Baker, T. (2020, October). AI and machine learning: A mixed blessing for cybersecurity. *In 2020 International Symposium on Networks, Computers and Communications* (ISNCC) (pp. 1-7). IEEE. https://doi.org/10.1109/ISNCC49221.2020.9297323

Karatas, G., Demir, O., & Sahingoz, O. K. (2018, December). Deep learning in intrusion detection systems. *In 2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT)* (pp. 113-116). IEEE. https://doi.org/10.1109/IBIGDELFT.2018.8625278

Karim, A., Salleh, R. B., Shiraz, M., Shah, S. A. A., Awan, I., & Anuar, N. B. (2014). Botnet detection techniques: *Review, future trends and issues. Journal of Zhejiang University Science C*, 15(11), 943-983. https://doi.org/10.1631/jzus.C1300242

Khonji, M., Iraqi, Y., & Jones, A. (2013). Phishing detection: A literature survey. *IEEE Communications Surveys and Tutorials*, 15(4), 2091-2121. https://doi.org/10.1109/SURV.2013.032213.00009

Kilincer, I. F., Ertam, F., & Sengur, A. (2021). Machine learning methods for cyber security intrusion detection: Datasets and comparative study. *Computer Networks*, 188, 107840. https://doi.org/10.1016/j.comnet.2021.107840

Kim, J., Sim, A., Kim, J., Wu, K., & Hahm, J. (2021). Improving Botnet Detection with Recurrent Neural Network and Transfer Learning. arXiv preprint arXiv:2104.12602.

Kim, K., & Aminanto, M. E. (2017, September). Deep learning in intrusion detection perspective: Overview and further challenges. *In 2017 International Workshop on Big Data and Information Security* (IWBIS) (pp. 5-10). IEEE. https://doi.org/10.1109/IWBIS.2017.8275095

Kim, T., Suh, S. C., Kim, H., Kim, J., & Kim, J. (2018, December). An encoding technique for CNN-based network anomaly detection. *In 2018 IEEE International Conference on Big Data (Big Data)* (pp. 2960-2965). IEEE. https://doi.org/10.1109/ICCNC.2018.8390278

Kolias, C., Kambourakis, G., Stavrou, A., & Gritzalis, S. (2015). Intrusion detection in 802.11 networks: empirical evaluation of threats and a public dataset. *IEEE Communications Surveys & Tutorials*, 18(1), 184-208. https://doi.org/10.1109/COMST.2015.2402161

Krishna, A., Lal, A., Mathewkutty, A. J., Jacob, D. S., & Hari, M. (2020, July). Intrusion detection and prevention system using deep learning. *In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 273-278). IEEE. https://doi.org/10.1109/ICESC48915.2020.9155711

Krishnaveni, S., Sivamohan, S., Sridhar, S. S., & Prabakaran, S. (2021). Efficient feature selection and classification through ensemble method for network intrusion detection on cloud computing. *Cluster Computing*, 24(3), 1761-1779. https://doi.org/10.1007/s10586-020-03222-y

Kumar, V., & Gupta, C. P. (2021, September). Cyber Security Issue in Smart Grid. *In 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)* (pp. 1-9). IEEE. https://doi.org/10.1109/GUCON50781.2021.9573600

Lakshmanarao, A., Rao, P. S. P., & Krishna, M. B. (2021, March). Phishing website detection using novel machine learning fusion approach. *In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)* (pp. 1164-1169). IEEE. https://doi.org/10.1109/ICAIS50930.2021.9395810

Lakshminarayana, D. H., Philips, J., & Tabrizi, N. (2019, December). A survey of intrusion detection techniques. *In 2019 18th IEEE International Conference on Machine Learning And Applications (ICMLA)* (pp. 1122-1129). IEEE. https://doi.org/10.1109/ICMLA.2019.00187

Le Jeune, L., Goedeme, T., & Mentens, N. (2021). Machine learning for misuse-based network intrusion detection: overview, unified evaluation and feature choice comparison framework. *Ieee Access*, 9, 63995-64015. https://doi.org/10.1109/ACCESS.2021.3075066

Lee, J., Kim, J., Kim, I., & Han, K. (2019). Cyber threat detection based on artificial neural networks using event profiles. *IEEE Access*, 7, 165607-165626. https://doi.org/10.1109/ACCESS.2019.2953095

Li, D., & Li, Q. (2020). Adversarial deep ensemble: Evasion attacks and defenses for malware detection. *IEEE Transactions on Information Forensics and Security*, 15, 3886-3900. https://doi.org/10.1109/TIFS.2020.3003571

Li, D., Li, Q., Ye, Y., & Xu, S. (2021). A framework for enhancing deep neural networks against adversarial malware. *IEEE Transactions on Network Science and Engineering*, 8(1), 736-750. https://doi.org/10.1109/AIPR50011.2020.9425293

Li, F. Q., Wang, S. L., Liew, A. W. C., Ding, W., & Liu, G. S. (2020). Large-Scale Malicious Software Classification with Fuzzified Features and Boosted Fuzzy Random Forest. *IEEE Transactions on Fuzzy Systems*, 29(11), 3205-3218. https://doi.org/10.1109/TFUZZ.2020.3016023

Li, J., Sun, L., Yan, Q., Li, Z., Srisa-An, W., & Ye, H. (2018). Significant permission identification for machine-learning-based android malware detection. *IEEE Transactions on Industrial Informatics*, 14(7), 3216-3225. https://doi.org/10.1109/TII.2017.2789219

Li, Y., Xiong, K., Chin, T., & Hu, C. (2019). A machine learning framework for domain generation algorithm-based malware detection. *IEEE Access*, 7, 32765-32782. https://doi.org/10.1109/ACCESS.2019.2891588

Li, Z., Rios, A. L. G., Xu, G., & Trajković, L. (2019, May). Machine learning techniques for classifying network anomalies and intrusions. *In 2019 IEEE international symposium on circuits and systems (ISCAS)* (pp. 1-5). IEEE. https://doi.org/10.1109/ISCAS.2019.8702583

Liang, F., Hatcher, W. G., Liao, W., Gao, W., & Yu, W. (2019). Machine learning for security and the internet of things: The good, the bad and the ugly. *IEEE Access*, 7, 158126-158147. https://doi.org/10.1109/ACCESS.2019.2948912

Lin, G., Wen, S., Han, Q. L., Zhang, J., & Xiang, Y. (2020). Software vulnerability detection using deep neural networks: a survey. *Proceedings of the IEEE*, 108(10), 1825-1848. https://doi.org/10.1109/JPROC.2020.2993293

Lippmann, R. P., Fried, D. J., Graf, I., Haines, J. W., Kendall, K. R., McClung, D., ... & Zissman, M. A. (2000, January). Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation. *In Proceedings DARPA Information Survivability Conference and Exposition. DISCEX'00* (Vol. 2, pp. 12-26). IEEE.

Lorenzen, C., Agrawal, R., & King, J. (2018, December). Determining viability of deep learning on cybersecurity log analytics. *In 2018 IEEE International Conference on Big Data (Big Data)* (pp. 4806-4811). IEEE. https://doi.org/10.1109/BigData.2018.8622165

Luiijf, E., Besseling, K., & De Graaf, P. (2013). Nineteen national cyber security strategies. *International Journal of Critical Infrastructures* 6, 9(1-2), 3-31. https://doi.org/10.1016/B978-0-12-800743-3.00002-5

Macas, M., & Wu, C. (2020, November). Deep Learning Methods for Cybersecurity and Intrusion Detection Systems. *In 2020 IEEE Latin-American Conference on Communications (LATINCOM)* (pp. 1-6). IEEE. https://doi.org/10.1109/LATINCOM50620.2020.9282324

Mahdavifar, S., & Ghorbani, A. A. (2019). Application of deep learning to cybersecurity: *A survey*. *Neurocomputing*, 347, 149-176. https://doi.org/10.1016/j.neucom.2019.02.056

Mahdavifar, S., Kadir, A. F. A., Fatemi, R., Alhadidi, D., & Ghorbani, A. A. (2020, August). Dynamic android malware category classification using semi-supervised deep learning. *In 2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)* (pp. 515-522). IEEE. https://doi.org/10.1109/DASC-PICom-CBDCom-CyberSciTech49142.2020.00094

Maimó, L. F., Gómez, Á. L. P., Clemente, F. J. G., Pérez, M. G., & Pérez, G. M. (2018). A self-adaptive deep learning-based system for anomaly detection in 5G networks. *Ieee Access*, 6, 7700-7712. https://doi.org/10.1109/ACCESS.2018.2803446

Mamun, M. S. I., Rathore, M. A., Lashkari, A. H., Stakhanova, N., & Ghorbani, A. A. (2016, September). Detecting malicious urls using lexical analysis. In International Conference on Network and System Security (pp. 467-482). Springer, Cham. https://doi.org/10.1007/978-3-319-46298-1_30\

MITLL. (2022). "DARPA Intrusion Detection Data Sets"." https://www.ll.mit.edu/r-d/datasets/1999-darpa-intrusion-detection-evaluation-dataset/

Malin, C. H., Gudaitis, T., Holt, T. J., & Kilger, M. (2017). Viral Influence: Deceptive Computing Attacks Through Persuasion. Deception in The Digital Age: *Exploiting and Defending Human Targets Through Computer-Mediated Communications*, 77-124. https://doi.org/10.1016/B978-0-12-411630-6.00007-4

Martínez, T. J., Iglesias Comesaña, C., & García-Nieto, P. J. (2019). Machine learning techniques applied to cybersecurity. International Journal of Machine Learning and Cybernetics, 10(10), 2823-2836. https://doi.org/10.1007/s13042-018-00906-1

Mishra, J., Sahay, S. K., Rathore, H., & Kumar, L. (2021, October). Duplicates in the Drebin Dataset and Reduction in the Accuracy of the Malware Detection Models. In 2021 26th IEEE Asia-Pacific Conference on Communications (APCC) (pp. 161-165). IEEE. https://doi.org/10.1109/APCC49754.2021.9609892

Nguyen, T. T., & Armitage, G. (2008). A survey of techniques for internet traffic classification using machine learning. *IEEE Communications Surveys & Tutorials*, 10(4), 56-76. https://doi.org/10.1109/SURV.2008.080406

Otoum, S., Kantarci, B., & Mouftah, H. (2018, May). Adaptively supervised and intrusion-aware data aggregation for wireless sensor clusters in critical infrastructures. *In 2018 IEEE international conference on communications* (ICC) (pp. 1-6). IEEE. https://doi.org/10.1109/ICC.2018.8422401

Otoum, S., Kantarci, B., & Mouftah, H. T. (2019). On the feasibility of deep learning in sensor network intrusion detection. *IEEE Networking Letters*, 1(2), 68-71. https://doi.org/10.1109/LNET.2019.2901792

Pantelidis, E., Bendiab, G., Shiaeles, S., & Kolokotronis, N. (2021, July). Insider threat detection using deep autoencoder and variational autoencoder neural networks. *In 2021 IEEE International Conference on Cyber Security and Resilience (CSR)* (pp. 129-134). IEEE. https://doi.org/10.1109/CSR51186.2021.9527925

Patil, P., Rane, R., & Bhalekar, M. (2017, January). Detecting spam and phishing mails using SVM and obfuscation URL detection algorithm. *In 2017 International Conference on Inventive Systems and Control (ICISC)* (pp. 1-4). IEEE. https://doi.org/10.1109/ICISC.2017.8068633

Pu, G., Wang, L., Shen, J., & Dong, F. (2020). A hybrid unsupervised clustering-based anomaly detection method. *Tsinghua Science and Technology*, 26(2), 146-153. https://doi.org/10.26599/TST.2019.9010051

Rao, P. U., Sodhi, B., & Sodhi, R. (2020, December). Cyber Security Enhancement of Smart Grids Via Machine Learning-A Review. *In 2020 21st National Power Systems Conference* (NPSC) (pp. 1-6). IEEE. https://doi.org/10.1109/NPSC49263.2020.9331859

Rashid, A., Siddique, M. J., & Ahmed, S. M. (2020, February). Machine and deep learning based comparative analysis using hybrid approaches for intrusion detection system. *In 2020 3rd International Conference on Advancements in Computational Sciences (ICACS)* (pp. 1-9). IEEE.

Ren, C., & Xu, Y. (2019). A fully data-driven method based on generative adversarial networks for power system dynamic security assessment with missing data. *IEEE Transactions on Power Systems*, 34(6), 5044-5052. https://doi.org/10.1109/TPWRS.2019.2922671

Ren, K., Zheng, T., Qin, Z., & Liu, X. (2020). Adversarial attacks and defenses in deep learning. *Engineering*, 6(3), 346-360. https://doi.org/10.1016/j.eng.2019.12.012

Ring, M., Wunderlich, S., Grüdl, D., Landes, D., & Hotho, A. (2017, June). Flow-based benchmark data sets for intrusion detection. *In Proceedings of the 16th European Conference on Cyber Warfare and Security. ACPI* (pp. 361-369).

Rodriguez, E., Otero, B., Gutierrez, N., & Canal, R. (2021). A survey of deep learning techniques for cybersecurity in mobile networks. *IEEE Communications Surveys & Tutorials*, 23(3), 1920-1955. https://doi.org/10.1109/COMST.2021.3086296

Rohith, C., & Batth, R. S. (2019, December). Cyber warfare: nations cyber conflicts, cyber cold war between nations and its repercussion. *In 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)* (pp. 640-645). IEEE. https://doi.org/10.1109/ICCIKE47802.2019.9004236

Roopak, M., Tian, G. Y., & Chambers, J. (2019, January). Deep learning models for cyber security in IoT networks. *In 2019 IEEE 9th annual computing and communication workshop and conference (CCWC)* (pp. 0452-0457). IEEE. https://doi.org/10.1109/CCWC.2019.8666588

Sajal, S. Z., Jahan, I., & Nygard, K. E. (2019, May). A survey on cyber security threats and challenges in modem society. *In 2019 IEEE international conference on electro information technology (EIT)* (pp. 525-528). IEEE. https://doi.org/10.1109/EIT.2019.8833829

Salah, A., Shalabi, E., & Khedr, W. (2020). A lightweight android malware classifier using novel feature selection methods. *Symmetry*, 12(5), 858. https://doi.org/10.3390/sym12050858

Salo, F., Nassif, A. B., & Essex, A. (2019). Dimensionality reduction with IG-PCA and ensemble classifier for network intrusion detection. *Computer Networks*, 148, 164-175. https://doi.org/10.1016/j.comnet.2018.11.010

Sapre, S., Islam, K., & Ahmadi, P. (2021, January). A comprehensive data sampling analysis applied to the classification of rare IoT network intrusion types. *In 2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)* (pp. 1-2). IEEE. https://doi.org/10.1109/CCNC49032.2021.9369617

Sarker, I. H. (2022). Ai-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems. *SN Computer Science*, 3(2), 1-20. https://doi.org/10.1007/s42979-022-01043-x

Sedik, A., Hammad, M., Abd El-Latif, A. A., El-Banby, G. M., Khalaf, A. A., Abd El-Samie, F. E., & Iliyasu, A. M. (2021). Deep learning modalities for biometric alteration detection in 5G networks-based secure smart cities. *IEEE Access*, 9, 94780-94788. https://doi.org/10.1109/ACCESS.2021.3088341

Shabut, A. M., Lwin, K. T., & Hossain, M. A. (2016, December). Cyber-attacks, countermeasures, and protection schemes-A state of the art survey. *In 2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA)* (pp. 37-44). IEEE. https://doi.org/10.1109/SKIMA.2016.7916194

Shahzad, H., Sattar, A. R., & Skandaraniyam, J. (2021, January). DGA domain detection using deep learning. *In 2021 IEEE 5th International Conference on Cryptography, Security and Privacy (CSP)* (pp. 139-143). IEEE. https://doi.org/10.1109/CSP51677.2021.9357591

Sharma, R. K., Kalita, H. K., & Borah, P. (2016). Analysis of machine learning techniques based intrusion detection systems. In Proceedings of 3rd International Conference on Advanced Computing, Networking and Informatics (pp. 485-493). *Springer, New Delhi*. https://doi.org/10.1007/978-81-322-2529-4_51

Shi, Y., & Sagduyu, Y. E. (2017, October). Evasion and causative attacks with adversarial deep learning. *In MILCOM 2017-2017 IEEE Military Communications Conference (MILCOM)* (pp. 243-248). IEEE. https://doi.org/10.1109/MILCOM.2017.8170807

Shibahara, T., Yagi, T., Akiyama, M., Chiba, D., & Yada, T. (2016, December). Efficient dynamic malware analysis based on network behavior using deep learning. *In 2016 IEEE Global Communications Conference (GLOBECOM*) (pp. 1-7). IEEE.

Shiravi, A., Shiravi, H., Tavallaee, M., & Ghorbani, A. A. (2012). Toward developing a systematic approach to generate benchmark datasets for intrusion detection. *Computers and Security*, 31(3), 357-374. https://doi.org/10.1016/j.cose.2011.12.012

Singh, C. (2020, March). Phishing website detection based on machine learning: *A survey. In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 398-404). IEEE. https://doi.org/10.1109/ICACCS48705.2020.9074400

Singh, K., Guntuku, S. C., Thakur, A., & Hota, C. (2014). Big data analytics framework for peer-to-peer botnet detection using random forests. *Information Sciences*, 278, 488-497. https://doi.org/10.1016/j.ins.2014.03.066

Singhal, S., Chawla, U., & Shorey, R. (2020, January). Machine learning & concept drift based approach for malicious website detection. *In 2020 International Conference on COMmunication Systems & NETworkS (COMSNETS)* (pp. 582-585). IEEE. https://doi.org/10.1109/COMSNETS48256.2020.9027485

Sohn, I. (2021). Deep belief network based intrusion detection techniques: A survey. *Expert Systems with Applications*, 167, 114170. https://doi.org/10.1016/j.eswa.2020.114170

Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018a). Toward generating a new intrusion detection dataset and intrusion traffic characterization. *ICISSp*, 1, 108-116. https://doi.org/10.5220/0006639801080116Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018b). Intrusion detection evaluation dataset (CIC-IDS2017). *Proceedings of the of Canadian Institute for Cybersecurity*. https://www. unb.ca/cic/datasets/ids-2017.html. Last Accessed 13-October-2021.

Statistics, (2021). Cyber-attack and Ransomware. Major Cyber Attacks December. https://konbriefing.com/en-topics/cyber-attacks-2021-12.html. Last Accessed 16-July2022.

Sokolov, S. A., Iliev, T. B., & Stoyanov, I. S. (2019, May). Analysis of cybersecurity threats in cloud applications using deep learning techniques. *In 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 441-446). IEEE. https://doi.org/10.23919/MIPRO.2019.8756755

Song, J., Takakura, H., Okabe, Y., Eto, M., Inoue, D., & Nakao, K. (2011, April). Statistical analysis of honeypot data and building of Kyoto 2006+dataset for NIDS evaluation. *In Proceedings of the first workshop on building analysis datasets and gathering experience returns for security* (pp. 29-36). https://doi.org/10.1145/1978672.1978676

Sood, A. K., & Zeadally, S. (2016). A taxonomy of domain-generation algorithms. *IEEE Security & Privacy*, 14(4), 46-53. https://doi.org/10.1109/MSP.2016.76

Starodubtsev, Y. I., Balenko, E. G., Vershennik, E. V., & Fedorov, V. H. (2020, October). Cyberspace: terminology, properties, problems of operation. *In 2020 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon)* (pp. 1-3). IEEE. https://doi.org/10.1109/FarEastCon50210.2020.9271282

Suresh, P. V., & Madhavu, M. L. (2021, July). Insider Attack: Internal Cyber Attack Detection Using Machine Learning. *In 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE. https://doi.org/10.1109/ICCCNT51525.2021.9579549

Taheri, R., Javidan, R., Shojafar, M., Vinod, P., & Conti, M. (2020). Can machine learning model with static features be fooled: An adversarial machine learning approach. *Cluster computing*, 23(4), 3233-3253. https://doi.org/10.1007/s10586-020-03083-5

Tang, L., & Mahmoud, Q. H. (2021). A survey of machine learning-based solutions for phishing website detection. *Machine Learning and Knowledge Extraction*, 3(3), 672-694. https://doi.org/10.3390/make3030034

Thakur, K., Hayajneh, T., & Tseng, J. (2019). Cyber security in social media: Challenges and the way forward. *IT Professional*, 21(2), 41-49. https://doi.org/10.1109/MITP.2018.2881373

Thamer, N., & Alubady, R. (2021, April). A Survey of Ransomware Attacks for Healthcare Systems: Risks, Challenges, Solutions and Opportunity of Research. *In 2021 1st Babylon International Conference on Information Technology and Science (BICITS)* (pp. 210-216). IEEE. https://doi.org/10.1109/BICITS51482.2021.9509877

Tsochev, G., Trifonov, R., Manolov, S., Nakov, O., & Spasov, S (2021). Analysis of Threats to a University Network Using Open Source Technologies. *In 2021 International Conference Automatics and Informatics (ICAI)* (pp. 366-369). IEEE. https://doi.org/10.1109/ICAI50593.2020.9311369

"TM. (2022 March.). Top 10 Malware. https://www.cisecurity.org/insights/ blog/top-10-malware-march-2022, April 2022. Last Accessed 15-May2022.

Uğurlu, M., & Doğru, İ. A. (2019, September). A survey on deep learning based intrusion detection system. *In 2019 4th International Conference on Computer Science and Engineering (UBMK)* (pp. 223-228). IEEE. https://doi.org/10.1109/UBMK.2019.8907206

Ullah, F., Naeem, H., Jabbar, S., Khalid, S., Latif, M. A., Al-Turjman, F., & Mostarda, L. (2019). Cyber security threats detection in internet of things using deep learning approach. *IEEE access*, 7, 124379-124389. https://doi.org/10.1109/ACCESS.2019.2937347

Urien, P. (2021, October). Innovative Countermeasures to Defeat Cyber Attacks Against Blockchain Wallets. *In 2021 5th Cyber Security in Networking Conference (CSNet)* (pp. 49-54). IEEE. https://doi.org/10.1109/CSNet52717.2021.9614649

Vanitha, K. S., Uma, S. V., & Mahidhar, S. K. (2017, December). Distributed denial of service: Attack techniques and mitigation. *In 2017 International Conference on Circuits, Controls and Communications (CCUBE)* (pp. 226-231). IEEE. https://doi.org/10.1109/CCUBE.2017.8394146

Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., & Venkatraman, S. (2019). Robust intelligent malware detection using deep learning. *IEEE Access*, 7, 46717-46738. https://doi.org/10.1109/ACCESS.2019.2906934

Wang, H., Han, B., Su, J., & Wang, X. (2018, October). A High-Performance Intrusion Detection Method Based on Combining Supervised and Unsupervised Learning. *In 2018 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation* (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI) (pp. 1803-1810). IEEE. https://doi.org/10.1109/SmartWorld.2018.00304

Xin, Y., Kong, L., Liu, Z., Chen, Y., Li, Y., Zhu, H., ... & Wang, C. (2018). Machine learning and deep learning methods for cybersecurity. *Ieee access*, 6, 35365-35381. https://doi.org/10.1109/ACCESS.2018.2836950

Xiujuan, W., Chenxi, Z., Kangfeng, Z., Haoyang, T., & Yuanrui, T. (2019, February). Detecting spear-phishing emails based on authentication. *In 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)* (pp. 450-456). IEEE. https://doi.org/10.1109/CCOMS.2019.8821758

Xu, S., Xia, Y., & Shen, H. L. (2020). Analysis of malware-induced cyber-attacks in cyber-physical power systems. IEEE Transactions on Circuits and Systems II: *Express Briefs*, 67(12), 3482-3486. https://doi.org/10.1109/TCSII.2020.2999875

Xu, W., Jang-Jaccard, J., Singh, A., Wei, Y., & Sabrina, F. (2021). Improving performance of autoencoder-based network anomaly detection on nsl-kdd dataset. *IEEE Access*, 9, 140136-140146. https://doi.org/10.1109/ACCESS.2021.3116612

Yamaguchi, S. (2020, January). Botnet Defense System: Concept and Basic Strategy. *In 2020 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1-5). IEEE. https://doi.org/10.1109/ICCE46568.2020.9043058

Yang, K., Liu, J., Zhang, C., & Fang, Y. (2018, October). Adversarial examples against the deep learning based network intrusion detection systems. *In MILCOM 2018-2018 ieee military communications conference (MILCOM)* (pp. 559-564). IEEE. https://doi.org/10.1109/MILCOM.2018.8599759

Yavanoglu, O., & Aydos, M. (2017, December). A review on cyber security datasets for machine learning algorithms. *In 2017 IEEE international conference on big data (big data)* (pp. 2186-2193). IEEE. https://doi.org/10.1109/BigData.2017.8258167

Yu, B., Pan, J., Gray, D., Hu, J., Choudhary, C., Nascimento, A. C., & De Cock, M. (2019). Weakly supervised deep learning for the detection of domain generation algorithms. *IEEE Access*, 7, 51542-51556. https://doi.org/10.1109/ACCESS.2019.2911522

Yu, D., & Deng, L. (2010). Deep learning and its applications to signal and information processing [exploratory dsp]. *IEEE Signal Processing Magazine*, 28(1), 145-154. https://doi.org/10.1109/MSP.2010.939038

Zargar, S. T., Joshi, J., & Tipper, D. (2013). A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks. *IEEE Communications Surveys and Tutorials*, 15(4), 2046-2069. https://doi.org/10.1109/SURV.2013.031413.00127

Zhang, S., Xie, X., & Xu, Y. (2020). A brute-force black-box method to attack machine learning-based systems in cybersecurity. *IEEE Access*, 8, 128250-128263. https://doi.org/10.1109/ACCESS.2020.3008433