NJESR/December-2022/Volume-4/Issue-12 E-ISSN-2582-5836
DOI-10.53571/NJESR.2022.4.12.32-49

A Study Of Cyber Security Using Machine Learning Techniques

Jiby P.Joseph
Lecturer
Department Of Computer Engineering
Government Polytechnic College
Perumbavoor

(Received:10November2022/Revised:20November2022/Accepted:30November2022/Published:16December2022)

Abstract

In the context of computing, cybersecurity is experiencing significant technological and operational shifts in recent days, and data science is driving these shifts. The key to making a security system automated and intelligent is to extract security incident patterns or insights from cybersecurity data and construct a corresponding data-driven model. Data science is the practice of using a variety of scientific methods, machine learning techniques, processes, and systems to comprehend and analyze the actual phenomena with data. Overall, our objective is to focus on the applicability of data-driven intelligent decision making to safeguarding systems from cyber attacks in addition to discussing cybersecurity data science and relevant methods. A subset of Artificial Intelligence (AI), machine learning (ML) focuses on the implementation of systems that can learn from past data, recognize patterns, and make rational decisions with little to no human intervention. The practice of preventing malicious attacks on digital systems like computers, servers, mobile devices, networks, and the data that goes with them is known as cyber security. Cyber security and machine learning can be combined in two main ways: taking into account cyber security when machine learning is used, and using machine learning to make cyber security possible. This unification can be beneficial to us in a number of ways, including by providing machine learning models with enhanced security, enhancing the effectiveness of cyber security methods, and facilitating the efficient detection of zero-day attacks with less human intervention. Cyber security and machine learning are brought together in this survey paper to bring together two distinct ideas. We also talk about the advantages, drawbacks, and difficulties of combining machine learning and cyber security. In addition, we provide a

comprehensive comparative study of various techniques in two distinct categories and discuss the various attacks. Finally, we suggest some directions for future research.

Keywords: Cyber Security; Machine Learning; Internet Of Things (IoT); Privacy; Security; Intrusion Detection

Introduction

Numerous security incidents, such as unauthorized access[2], malware attack, zero-day attack, data breach, denial of service (DoS)[2], social engineering or phishing, etc., have arisen as a result of the growing reliance on digitalization and the Internet of Things (IoT) [1]. in recent years have expanded at an exponential rate. For instance, the security industry was aware of fewer than 50 million distinct malware executables in 2010. According to the data provided by the AV-TEST institute in Germany, this number is expected to rise to over 900 million malicious executables in 2019, having doubled to around 100 million by 2012. Attacks and cybercrime can have devastating financial consequences for both individuals and businesses. That's what it's assessed, an information break costs 8.19 million USD for the US and 3.9 million USD on a normal , and the yearly expense for the worldwide economy from cybercrime is 400 billion USD . Over the next five years, the number of records breached will nearly triple annually, according to Juniper Research. As a result, in order to reduce the damage, businesses must adopt and implement a robust cybersecurity strategy. According to[1], a nation's national security is dependent on its business, government, and individual citizens having access to highly secure applications and tools and the ability to promptly identify and eliminate such cyber threats. As a result, a crucial problem that must be resolved as soon as possible is how to intelligently protect the relevant systems from various cyber attacks, whether they have been previously observed or not. A set of technologies and procedures aimed at safeguarding computers, networks, programs, and data from harm or unauthorized access is known as cybersecurity[1-2]. Data science (DS) is driving the change in cybersecurity, where machine learning (ML), a fundamental component of "Artificial Intelligence" (AI), can play a crucial role in discovering insights from data. In recent days, cybersecurity is undergoing massive technological and operational shifts in the context of computing. Data science is leading a new scientific paradigm, and machine learning has the potential to significantly alter the cybersecurity landscape [4].

Multiple layers of defense are distributed across the networks, computers, programs, or information that the strategy aims to keep safe. In order to build a solid defense against cyber

attacks and recover from them, a society's processes, people, and tools must all work together. A single threat management system can automate additions across specific Cisco Security products and accelerate key security processes: revelation, assessment, and remediation.

Customers need to understand and adhere to fundamental information security principles like using strong passwords, being wary of attachments in emails, and backing up data. Learn extra around fundamental online protection values.

How Does Cyber Security Make Working So Easy

There is no doubt that the cybersecurity tool simplifies our work by ensuring that restricted capitals can be obtained from any network. A business or society could look a colossal harm in the event that they are not genuine about the security of their web-based event. Everyone benefits from progressive cyber defense agendas in today's connected world. A cybersecurity outbreak can also cause everything from identity theft to attempts at extortion to the loss of important data like family photos. Everyone is dependent on dangerous structures like influence plants, hospitals, and financial services companies. For our civilization to function, it is essential to secure these and other societies. All of them also receive compensation for their work as cyberthreat investigators—similar to the 250 risk investigators at Talos—who investigate emerging threats and cyber security policies. They strengthen open source gears, educate the community about cybersecurity's position, and reveal new vulnerabilities. Their work ensures that everyone can use the Internet safely.

The majority of the computing devices we use today are connected to the Internet in an environment known as the Internet of Things (IoT). These gadgets share and send their information through the shaky (open) correspondence medium, likewise called as the Web. Healthcare data, banking data, insurance data, other financial data, and social security numbers are all examples of sensitive data. Online attackers (hackers) are always looking for that, where they can play with things and launch attacks like replay, man-in-the-middle, credential guessing, session key computation, malware injection, and data modification[1],[2]. As a result, various researchers periodically propose various security protocols to stop these attacks. There are many different types of security protocols or cyber security protocols: key management protocols, "blockchain enabled security protocols," "access control protocols," "intrusion detection protocols," and "authentication protocols" The following is a summary of these protocols.

• Methods For Authentication: The process of determining whether a person or a device is genuine (authentic) is known as authentication. It can be carried out with the help of some credentials or factors, such as a username, password, smartcard, or biometrics, that have a strong connection to the users or the device. We can have authentication between users, between users and devices, or between devices and users. User authentication protocols can again be broken down into three categories based on the factors that are available: one-factor user authentication protocols, two-factor user authentication protocols, and three-factor user authentication protocols.

• Methods For Controlling Access: The process of preventing unauthorized access to a person or a device is known as access control. After a user/device access control protocol has completed all of its steps, users or devices can gain secure access to other users or devices. There are two types of access control protocols:(1) client access control and(2) gadget access control. Device access control protocol can be used to control access to unauthorized devices, while user access control protocol can be used to control access to unauthorized users. Certificate-based or certificate-less access control are options. Approval is considered as a cycle through which a power (i.e., a server) decides whether a substance (i.e., a client) has consent to utilize the asset. Typically, it is carried out in conjunction with authentication so that the server can identify the client requesting access. Who has permission to access a resource and who does not is determined by this.

• Methods For Detecting Intrusions: An intentional act of something or someone constitutes an intrusion. This could be a malicious programming script or a system controlled by a hacker that is on the Internet. Typically, hackers attempt to infect online devices with malware to compromise their security or affect their performance. We require a subset of protocols that fall under the category of "intrusion detection protocols" for the purpose of detecting and mitigating intrusions. Signature-based intrusion detection, anomaly-based intrusion detection, or hybrid intrusion detection, which combines signature- and anomaly-based schemes, are all methods of intrusion detection. These days, intrusion detection that is based on machine learning or deep learning, also known as malware detection, is gaining a lot of popularity.

• Important Management Practices: Key management protocols are utilized for secure key management among various entities, including some users (smart home user, doctor, traffic inspector, and smart Internet of Things (IoT) devices, for example). Typically, a believed enlistment authority does the enrollment of all substances of the correspondence framework and

afterward stores the mysterious qualifications (i.e., secret keys) in their memory. For the purposes of creating new keys and storing them in devices, key establishment, and key revocation, we require a key management procedure. After establishing a shared secret key (also known as a session key) through the essential steps of an authenticated key agreement protocol, devices and users can securely exchange information.

• Security Protocols Made Possible By The Blockchain: One of today's emerging technologies is blockchain. Data is stored on a blockchain in the form of blocks that are linked together by some hash values. The distributed ledger technology (DLT) used in the blockchain keeps track of data in a distributed ledger. The DLT is accessible to any and all legitimate network participants, including miner parties. The information that we store over the blockchain safe and got against the different conceivable digital assaults. As a result, the various cyberattacks can be stopped by the security protocols that are enabled by blockchain [3].

Computer systems learn from data and use algorithms to carry out tasks without being explicitly programmed in a process called machine learning (ML). A subset of machine learning (ML) known as deep learning (DL) is part of artificial intelligence (AI). A complex set of algorithms that are based on the human brain serve as the foundation for DL. This makes it possible to process unstructured data like documents, images, and text. The term "ML" refers to a computer's capacity to think and act autonomously. On the other hand, DL usually only requires occasional human intervention. It is able to analyze images, videos, and unstructured data more effectively than conventional machine learning algorithms[4,5].

We can benefit in a number of ways from combining machine learning and cyber security. For instance, improved security for machine learning models, enhanced performance of cyber security methods, and efficient zero-day attack detection requiring less human intervention. However, it may have a number of issues and security challenges that must be carefully managed. Consequently, in this specific space, we really want some audit study connected with the "joining of network safety and AI" i.e., issues and difficulties, different assaults, different security plans with their near study and some future exploration bearings on which different analysts ought to work. As a result, we attempted to carry out such research in the proposed work [4,5].

Cyber Security In Machine Learning

Figure depicts the scenario of "cyber security in machine learning." 2, also known as machine learning security (ML security). Various phenomena can be analyzed and predicted with the help of ML models. However, the launching of certain attacks, such as dataset poisoning, model poisoning, privacy breach, membership inference, runtime disruption, and others, [] can have an impact on the performance of ML models[6]. These attacks may cause ML models to make incorrect predictions about the associated phenomena. An adversary inserts adversarial examples (updated values) into a dataset in the "dataset poisoning attack," which causes the ML model to make incorrect predictions. In addition, the "model poisoning attack" focuses on corrupting the models by altering their parameters and interfering with their internal operations. The attacker in a "privacy breach attack" works on exposing sensitive data and attempting to retrieve valuable model information. The privacy breach component is the membership inference attack. In addition, in the "runtime disruption attack," the attacker targets the model's execution process in order to influence accurate prediction results. Subsequently, there is a requirement for some network safety components (i.e., encryption procedures, signature age and confirmation methods, hashing systems) to safeguard against these assaults. The ML models and associated datasets become secure when these cyber security mechanisms are implemented, resulting in accurate predictions and outcomes.

Advantages Of Uniting Cyber Security And Machine Learning

Cyber security and machine learning can both benefit from one another and perform better together. The following are some of the benefits of joining forces:

• ML Models Complete Security Proof: The ML models are susceptible to a variety of attacks, as was previously mentioned. The ML models' operation, performance, and predictions may be affected by these attacks. However, certain cyber security mechanisms can be used to prevent these undesirable events. The functioning, performance, and input datasets of ML models are secured when cyber security mechanisms are implemented, resulting in accurate predictions and results[7].

• Techniques For Cyber Security That Perform Better: When ML algorithms are used in intrusion detection systems, which are part of cyber security plans, their performance is improved (more accurate detection rates and fewer false positives). Depending on the communication environment and associated systems, ML techniques like supervised learning, unsupervised learning, reinforcement learning, and deep learning algorithms can be utilized.

• Effective Zero-Day Attacks Detection: For zero day attacks (i.e., unknown malware attacks), the cyber security methods that use ML models to detect intrusions appear to be very effective. It happens because they detect using some deployed machine learning models. The ML models function by matching and collecting particular features; if a program's features match those of a malicious program, that program can be considered malicious. The ML models can carry out this detection function automatically. Hence, discovery of multi day assaults can be performed actually with the joining of digital protection and AI.

• Requirements For Human Intervention Are Minimal: The deployed ML models carry out the majority of the work in systems that are based on ML. When cyber security and machine learning are combined, the majority of the tasks for which these systems are deployed are carried out with little or no human intervention.

• Rapid Inspection And Mitigation: Because they operate through specific ML algorithms, the ML-based intrusion detection systems are able to detect the presence of attacks very effectively. As a result, when cyber security systems are combined with machine learning, they can quickly scan for intrusions and respond to any signs of intrusion. The best ML algorithm selection is the only thing we need to keep in mind.

Overview Of Various Threats And Attacks

The following attacks, which may occur in various computing environments, are detailed in this section.

• Listening In: This assault is uninvolved in nature which is otherwise called sniffing or sneaking around assault. An adversary attempts to listen in on a secret conversation between two parties in this attack.

• Analysis Of Traffic: This assault is detached in nature. In this attack, the adversary A intercepts the current conversation and then examines the messages to obtain information about the conversation's type, pattern, and behavior, location, and timing. A is able to launch additional attacks that are linked to the intercepted data.

• Counter Attack: A deliberately sends the previously exchanged captured messages back to the target in this attack. This is done by A to deceive or mislead the recipient and forces legitimate users to follow A's wishes.

•The Man-in-the-Middle (MiTM) Attack: A initiates independent connections with communicating entities and transmits messages to both ends in this active attack. In such

circumstances, the two communicating entities believe they are communicating directly with one another. As a result, A may unknowingly intercept, alter, transmit, or insert new information [8].

• Attack Of Impersonation: A imitates one of the legitimate parties in the network by deducing its identity and then sends modified or new messages to the other legitimate party on behalf of that party. This attack is also active in nature.

• A DoS (Denial Of Service) Attack: In a DoS attack, A floods the victim's computing resources with multiple fake requests (also known as HTTP flood messages). As a result, the legitimate user's service request cannot be processed. The legitimate user cannot access the network's services in such a scenario. Another type of DoS attack is the distributed denial-of-service (DDoS) attack, in which A uses a botnet of multiple machines to send multiple requests simultaneously to the victim's machine, consuming all of the system's computing resources quickly. Different flooding attacks, such as SYN flood, HTTP flood, UDP flood, and so on, can be used to carry out DoS or DDoS attacks.

• Attack By Malware: The execution of malicious script on the victim's machine is the method by which these attacks are carried out. The malware that is injected or installed is a file or a program that performs unauthorized activities in the systems, such as stealing data, encrypting the drive or stored data in an illegal way, modifying data, or deleting data. Keyloggers, spyware, viruses, ransomware, worms, trojan horses, and other types of malware include [6].

• Prearranging Assault: The disclosure of information from a web server-managed online database (such as an online banking database) is the subject of these attacks. Passwords, credit and debit card information, and other secret information can be obtained from the system by means of, for instance, "password cracking," "structured query language (SQL) injection attack," and "cross-site scripting (XSS) attack."

• Privileged Attack By An Insider: Any privileged system user with access to the registration information of multiple users and devices can carry out this attack. This attack becomes much more difficult to defend and has a greater negative impact because a privileged insider has access to the sensitive information.

• Stolen Smart Devices In Person: Nowadays, the majority of computing environments are controlled by smart devices, such as smart appliances for the home, smart healthcare devices, and smart manufacturing devices. There is no physical security provided when the smart devices are deployed. Power analysis attacks can be used to extract sensitive information from these

smart devices if they are physically stolen by an adversary A. Unauthorized tasks like illegal session key computation can be carried out following the removal of sensitive data [9].

• Affair On Birthday: A type of cryptographic attack known as a "birthday attack" exploits the probability theory-based mathematics of the "birthday problem." Birthday attacks can be used maliciously, such as to guess passwords or credentials. This attack is based on a fixed degree of permutations and the higher likelihood of collisions between random attack attempts, as described in the birthday paradox. The probability that some paired members of a group of n randomly selected individuals will share a birth date is the subject of the birthday paradox, also known as the birthday problem. The birthday attack, a well-known cryptographic attack that employs this probabilistic strategy to reduce the difficulty of cracking a hash function, was inspired by the mathematics of this problem[10].

• Attack On Dictionaries: A word reference assault is a sort of beast force assault on a cryptographic framework that is done perniciously. The attacker tries to break into the system's security by using every word in a dictionary as a password on a consistent basis. Additionally, a dictionary attack can be used to determine the key required to decrypt an encrypted communication or document. Through an up-to-date phrase or keyword library, the attacker tries to break the encryption or gain access. Automatic insertion into the target can be done with words from a dictionary or numerical sequences. Poor password usage, such as replacing passwords with sequential numbers, symbols, or characters, makes dictionary attacks easier. Some people use common words as passwords, so it works. Systems that employ multi-word passwords typically fall victim to these attacks. Also, it's hard for an attacker to break passwords made of uppercase and lowercase letters, numbers, and random combinations[10].

• Attack On A Stolen Verifier: In this malicious act, an attacker first tries to steal some devices, such as smart IoT devices, and then uses a power analysis attack to extract sensitive information, such as secret credentials and keys, from the memory of these devices. The attacker listens in on some of the messages that are being exchanged and then uses the information that is extracted to launch additional potential network attacks, such as unauthorized session key computation, password guessing, MiTM attacks, and impersonation attacks.

• A Session Key Computation Attack Without Authorization: An attacker attempts to compute the session key, which is established between the legitimate entities of the network, in this malicious act. The attacker employs a variety of strategies to complete this task, including the

stolen physical device attack, the privileged insider attack, and the stolen verifier attack. When calculating the session keys, it is always recommended to make use of both long-term secrets (such as secret keys and pseudo identities) and short-term secrets (such as random secret nonce values). This mechanism provides distinct keys to various entities across multiple sessions. Unfortunately, if an attacker discovers a session key, the remaining communication will remain secure because other session keys will remain safe.

• Attacks On Models For Machine Learning: The four broad categories of attacks on the ML model are as follows: a) a poisoning of the dataset; b) a poisoning of the model; c) a breach of privacy; and d) a disruption of the runtime [11].

• The Poising Attack On Data: A uses a variety of techniques to invade the training and testing data in this attack to disrupt the ML task's normal operation. A can attack the data server from which raw data must be extracted by employing adversarial examples. The compromising of the information sources serves to embeds the incorrect information, which perhaps modifies the working of the ML model. The ML-based system's output is further altered as a result[12].

• An Example Of Poisoning: In the model poisoning attack, A modifies the parameters, which causes A to produce incorrect output by interfering with the classifier. The parameters that the classifier uses to prepare the ML model change. A can alter the rate of accession, sensitivity limits, and under- or over-fitting, all of which have an impact on the normal execution of ML tasks[13].

• Breach Of Privacy: There are a variety of ways that sensitive user data and the model's internal working mechanism can be hacked. Data can leak when files aren't encrypted and there isn't one in the ML task during the training and deployment phases. That makes it even easier for the unauthorised user to alter the model. Because sensitive data may be compromised, it raises the privacy risks associated with the data[14]. Papernot and co.[15],[16] went over a variety of privacy-preserving strategies for model privacy. They also talked about using noise generation to "randomize model's behavior" to give the data and ML model different levels of privacy[17].

• Disruption Attack At Runtime: This task is used by adversary A to stop or delay the ongoing ML task. At the time of the deployment phase, A typically targets the server. Then, A tries to interfere with the ML process from a distance. As a result, the ML task's normal operation is disrupted, resulting in time and resource waste. A uses a variety of attacks, such as phishing, a denial-of-service (DoS) attack, and a SQL injection attack, to get into the run time server and

discover the weak points (vulnerabilities). The decentralization of ML work space has the potential to mitigate this attack. To further bifurcate and implement "distributed machine learning," the blockchain-based mechanisms can be used to safeguard the integrity and privacy of the user's data and the associated ML models.

Issues And Challenges Of Uniting Of Cyber Security And Machine Learning

However, there are numerous benefits to combining machine learning and cyber security. It also has some problems and challenges that need to be handled with extreme care. Below, we'll go over a few of them.

• Issues with compatibility: Algorithms for machine learning, such as clustering, classification, and convolutional neural networks (CNNs), as well as security techniques like encryption, signature generation, and verification, are included in the combination of cyber security and machine learning. In addition, IoT devices provide the data that serves as the primary input for the analysis process. Different methods of communication are used to operate these Internet of Things devices. There may be compatibility issues when these numerous algorithms are combined. As a result, we must choose carefully which algorithm and scheme complement each other best. As a result, compatibility-related issues must be handled with extreme caution [8].

• Over-burdening: We use a variety of algorithms to combine machine learning and cyber security, as was previously mentioned. We require additional resources to execute such algorithms. If not, the framework won't work as expected. As a result, the system may become overloaded as a result of the combination and application of various algorithms, which may further impair the system's actual operation. For instance, we are unable to devote all of the system's resources to security-related processes. Additionally, we require some resources for ML-related tasks. As a result, we should carefully select the algorithms and take into account the communication environment's resources. For the secure communication of IoT, for instance, we would prefer to use the symmetric-key based encryption algorithm known as the Advanced Encryption Standard (AES) rather than any public key cryptographic algorithm for encryption purposes. This is due to the fact that AES requires less computation, communication, and storage costs than public key cryptographic algorithms. We can also allocate the system's resources for important tasks in that circumstance.

• Precision: In the joining of network safety and AI, we utilize different ML components i.e., AI (ML) models to anticipate about a few actual peculiarities (i.e., chances of side of the road

mishap in the keen transportation framework). Because ML models are based on specific datasets, errors in the dataset or the ML model's settings can cause significant issues. For instance, the accuracy that was obtained is not entirely accurate [9].

• Security mechanisms that are flawed: We may employ various cyber security mechanisms in the integration of ML and cyber security. The system's security may be compromised if these mechanisms have a few flaws. The hackers attempt to find zero-day vulnerabilities and then take advantage of them the majority of the time. The system's sensitive data may be revealed, altered, or rendered inaccessible in such circumstances. As a result, when developing a new security protocol, the protocol developers ought to exercise extreme caution. Certain mechanisms, like the Automated Validation of Internet Security Protocols and Applications (AVISPA) [10], which uses formal security verification to check the security of the newly designed protocol against replay and man-in-the-middle attacks, can be used to test the security of the protocol. In addition, the "Burrows–Abadi–Needham (BAN) logic test"[11] can be used to look for "secure mutual authentication among the communicating entities." The Real-or-Random (ROR) model implementation, which identifies the possibility of an unauthorized session key computation attack on the designed authentication, access control, or key management protocol, also allows us to analyze the formal security of a security protocol[12]. In this manner, the designed protocol's security can be evaluated and analyzed.

Comparative Study

We have compared various methods under the headings "machine learning in cyber security" and "cyber security in machine learning" in this section. The specifics are provided below.

Comparison of machine learning-based intrusion (also known as malware) detection strategies' performance in cyber security protocols We conduct a comparative analysis.

A synopsis of the schemes Kumar et al.[13] proposed a framework to enhance IoT device intrusion detection by combining the advantages of ML models and blockchain. Through "clustering, classification, and blockchain," they have implemented it in a sequential manner. ML used clustering and classification algorithms to automatically extract the malware information, which was then stored over the blockchain network. Lei and others[11] came up with the name "EveDroid, a scalable and event-aware malware detection system" for a security plan. Their scheme, in contrast to other schemes, uses event groups directly to describe app behaviors, which may capture a higher level of semantics for detection purposes. Nguyen et al.[14] proposed

a method for identifying the Linux IoT botnet. The "PSI graph and CNN classifier" combination was the foundation for the detection. Dinakarrao and others[15] proposed a two-pronged approach to intrusion detection, utilizing a runtime malware detector (HaRM) and "Hardware Performance Counter (HPC)" values to distinguish between malicious software and legitimate applications. Su et al. proposed a simple method for identifying DDoS malware in IoT environments. They have removed the malware pictures and used a light-weight convolutional brain network for their grouping. Table 1 shows how these plans compare to one another.

We have only considered the most recent domain schemes for the comparative study portion. The various performance parameters, such as the F1-score, precision, recall, and accuracy, are utilized. We use parameters like true positive (TP), false positive (FP), true negative (TN), and false negative (FN) to calculate these parameters. The term "true negative (TN)" refers to the situation in which the intrusion detection scheme identifies a "normal program" as such; whereas a "false positive (FP)" occurs when an intrusion detection scheme identifies a "normal program" as a "malicious program." Similarly, a "true positive (TP)" occurs when an intrusion detection scheme identifies a "malicious program" as such; In contrast, a "false negative (FN)" occurs when an intrusion detection scheme identifies a "malicious program" as a "normal program."

• **Precision:** It is known as “positive predicted value” which is the fraction of the correctly identified intrusion cases to all the predicted positive cases of intrusions, where Precision=TPTP+FP.

• **Recall:** It is also known as “true positive rate or detection rate or sensitivity” which is treated as a fraction of correctly identified intrusion cases to the all real positive cases of intrusions, where Recall=TPTP+FN.

• **Accuracy:** It is one of the most important parameters measured as the all correctly identified cases, which is formulated as Accuracy=TP+TNTP+FP+TN+FN.

• **F1-score:** It is also known as F1-measure that is calculated through the harmonic mean of precision and recall. It gives the accurate estimate of the incorrectly classified cases than the accuracy and is formulated as F1-score=2(Precision×Recall)Precision + Recall.

Comparison Of Various Schemes

The comparison of the different schemes i.e., scheme of Kumar et al. , Lei et al. , Nguyen et al. , Dinakarrao et al. and Su et al. is given in Table 1. The schemes of Kumar et al. , Nguyen et al. , Dinakarrao et al. and Su et al. provide the accuracy value of 98%, 92%, 92.21% and 94%,

respectively. Contrary to that the schemes of Kumar et al. , Lei et al. and Nguyen et al. provide the F1-measure of 98%, 99% and 94%, respectively. From the comparison, it has been observed that Kumar et al.'s scheme provides better accuracy. However, the scheme of Lei et al. provides maximum F1-measure.

Table 1. Comparison Of Different ML-Based Intrusion Detection Schemes

Scheme	Method used	Accuracy	F1-measure
Kumar et al. [23]	"Blockchain enabled ML driven detection"	98.00%	98.00%
Lei et al. [11]	"EveDroid"	N/A	99.00%
Nguyen et al. [24]	"Graph-based convolution neural network (CNN)"	92.00%	94.00%
Dinakarrao et al. [25]	"HaRM malware detector"	92.21%	N/A
Su et al. [26]	"Light-weight convolutional neural network (CNN)"	94.00%	N/A

Performance Comparison Of Cyber Security In Machine Learning Protocols

We provide the details of different schemes that can secure the ML models.

Summary Of Considered Schemes

Jagielski, others suggested a defense mechanism that was able to withstand a variety of poisoning attacks. In addition, they offered formal assurances regarding its convergence and a limit on the impact of poisoning attacks. Peri and co. on the "CIFAR-10 dataset," a Deep k-NN defense mechanism against "collision and convex polytope clean-label attacks" was proposed. Chen and co. presented De-Pois, an attack-independent poisoning attack defense. Their plan's central concept was to train a mimic model. They have done this in order to imitate the target model's behavior. Phong and co. proposed "additively homomorphic encryption," a deep learning-based method for safeguarding gradients over the "honest but curious cloud server." The cloud server contained encrypted copies of the various gradients. Payman and Zhang proposed MPC-friendly alternatives to non-linear functions like "sigmoid and softmax," which were superior to the other existing schemes, and a mechanism to support secure arithmetic operations on shared decimal numbers. Chen and co. proposed a method for identifying and removing neural network backdoors. They demonstrated how well the plan for "neural networks

classifying text and images" worked. According to their claim, it was the first method for detecting poisonous data. Liu and co. proposed the efficient means of defense against backdoor attacks. Three backdoor attacks were used to investigate two promising defense strategies known as "pruning and fine-tuning." The "fine-pruning," which was a combination of "pruning and fine-tuning," was then evaluated. Weber and co. through the use of "randomized smoothing techniques," the unified framework was provided. It had demonstrated how it could be implemented to verify the system's resistance to "evasion and backdoor attacks." The "robust training process," or RAB, was then presented to smooth the trained model and demonstrate its resistance to backdoor attacks. The robustness bound for RAB-trained ML models was derived by them.

Conclusion

By combining ML and cyber security, we discussed two distinct concepts. The benefits, drawbacks, and difficulties of combining machine learning and cyber security were then discussed. The following are some benefits: "quick scanning and mitigation," "effective detection of zero-day attacks," "full proof security of ML models," and "improved performance of cyber security techniques" are a few examples. However, there are issues with this uniting as well, such as "compatibility issues," "overloading," "accuracy," and so on. Eavesdropping, traffic analysis, replay, MiTM, impersonation, DoS, malware insertion, scripting, birthday, physical stealing of smart devices, dictionary, dataset poising, model poisoning, and runtime disruption attacks were among the domain attacks that were discussed. After that, we conducted a comprehensive comparison of various techniques belonging to two distinct thought categories. Take, for instance, Kumar et al.'s plan. performed better in the "machine learning in cyber security" category, whereas Chen et al.'s plan performed better in the "cyber security in ML" category. In order for other researchers to provide some solutions for those issues, some future research directions (such as "secrecy of exchanged and stored data," "compatibility of different mechanisms and tools," "overloading and performance," and "improvement in accuracy of the system") were also provided. As a result, there is a trade-off between performance and the cost of learning. For instance, DL is more expensive than ML, but it has higher predictive scores. Additionally, we must increase our investments in the system's resources if we want to increase security. By combining cyber security and machine learning, we presented the specifics of two distinct concepts: AI in network safety" and "digital protection in AI". The benefits, drawbacks,

and difficulties of combining ML and cyber security were then discussed. In addition, we provided a comparison study of various techniques in two distinct categories and emphasized the various attacks. Finally, some directions for future research are provided.

References

[1]. Butun, I., Österberg, P., & Song, H. (2019). Security of the Internet of Things: Vulnerabilities, attacks, and countermeasures. IEEE Communications Surveys & Tutorials, 22(1), 616-644.

[2]. Lv, Z., Qiao, L., Li, J., & Song, H. (2020). Deep-learning-enabled security issues in the internet of things. IEEE Internet of Things Journal, 8(12), 9531-9538.

[3]. Wang, Y., Yu, J., Yan, B., Wang, G., & Shan, Z. (2020). BSV-PAGS: Blockchain-based special vehicles priority access guarantee scheme. Computer Communications, 161, 28-40.

[4]. Xu, Y., de Souza, R. W., Medeiros, E. P., Jain, N., Zhang, L., Passos, L. A., & de Albuquerque, V. H. C. (2022). Intelligent IoT security monitoring based on fuzzy optimum-path forest classifier. Soft Computing, 1-10.

[5]. Wazid, M., Das, A. K., Chamola, V., & Park, Y. (2022). Uniting cyber security and machine learning: Advantages, challenges and future research. ICT Express.

[6]. Kamruzzaman, M. M., Alrashdi, I., & Alqazzaz, A. (2022). New opportunities, challenges, and applications of edge-AI for connected healthcare in internet of medical things for smart cities. Journal of Healthcare Engineering, 2022.

[7]. Stofkova, K. R., & Janoskova, P. (2021). Perception of the concept of smart city from the perspective of cities and municipalities of the Slovak Republic. In SHS Web of Conferences (Vol. 129, p. 08018). EDP Sciences.

[8]. Zeng, H., & Zou, Q. (2022). Secure Analysis for IIOT Systems Using Hyperchaotic Image Encryption. Security and Communication Networks, 2022.

[9]. Parah, S. A., Kaw, J. A., Bellavista, P., Loan, N. A., Bhat, G. M., Muhammad, K., & de Albuquerque, V. H. C. (2020). Efficient security and authentication for edge-based internet of medical things. IEEE Internet of Things Journal, 8(21), 15652-15662.

[10]. Sun, Y., Bashir, A. K., Tariq, U., & Xiao, F. (2021). Effective malware detection scheme based on classified behavior graph in IIoT. Ad Hoc Networks, 120, 102558.

[11]. Yang, J., Bian, Z., Liu, J., Jiang, B., Lu, W., Gao, X., & Song, H. (2021). No-reference quality assessment for screen content images using visual edge model and adaboosting neural network. IEEE Transactions on Image Processing, 30, 6801-6814.

[12]. Zhao, Y., Yang, J., Bao, Y., & Song, H. (2021). Trustworthy authorization method for security in Industrial Internet of Things. Ad Hoc Networks, 121, 102607.

[13]. Messerges, T. S., Dabbish, E. A., & Sloan, R. H. (2002). Examining smart-card security under the threat of power analysis attacks. IEEE transactions on computers, 51(5), 541-552.

[14]. Soltanian, M., & Amiri, I. (2016). Problem Solving, Investigating Ideas, and Solutions. Theoretical and Experimental Methods for Defending Against DDOS Attacks, 33-45.

[15]. Lei, T., Qin, Z., Wang, Z., Li, Q., & Ye, D. (2019). EveDroid: Event-aware Android malware detection against model degrading for IoT devices. IEEE Internet of Things Journal, 6(4), 6668-6680.

[16]. Steinhardt, J., Koh, P. W. W., & Liang, P. S. (2017). Certified defenses for data poisoning attacks. Advances in neural information processing systems, 30.

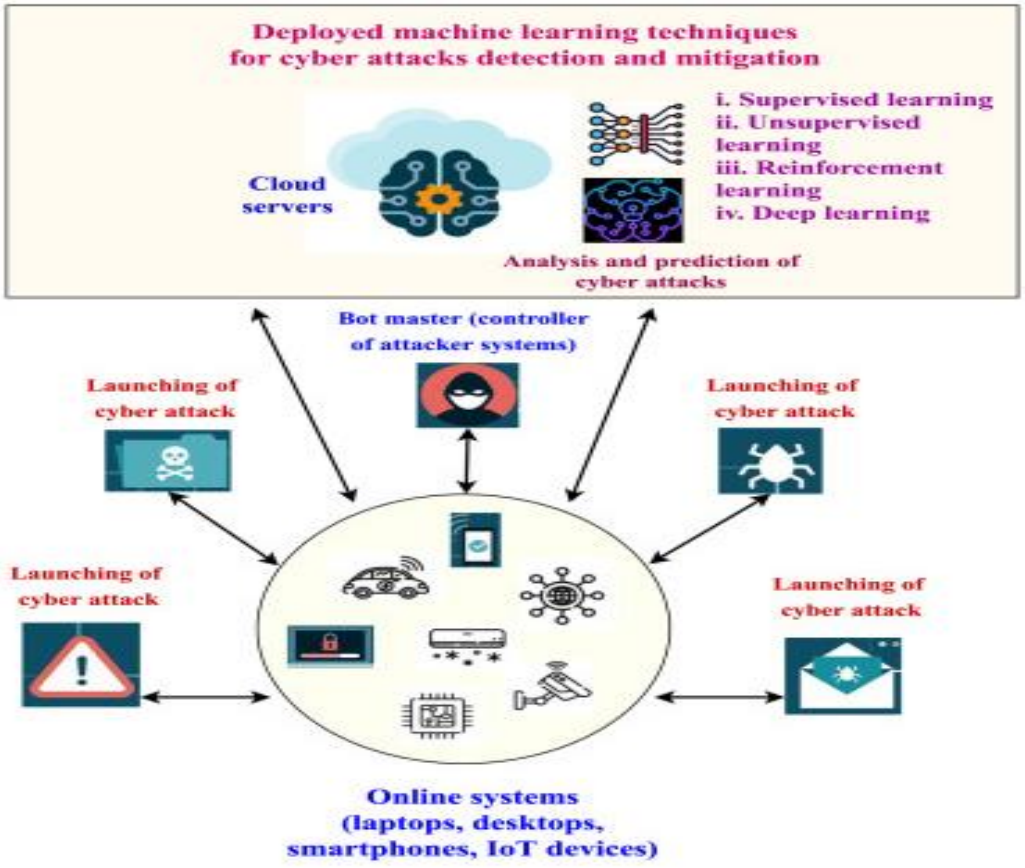

Fig. 1. Scenario Of Machine Learning In Cyber Security

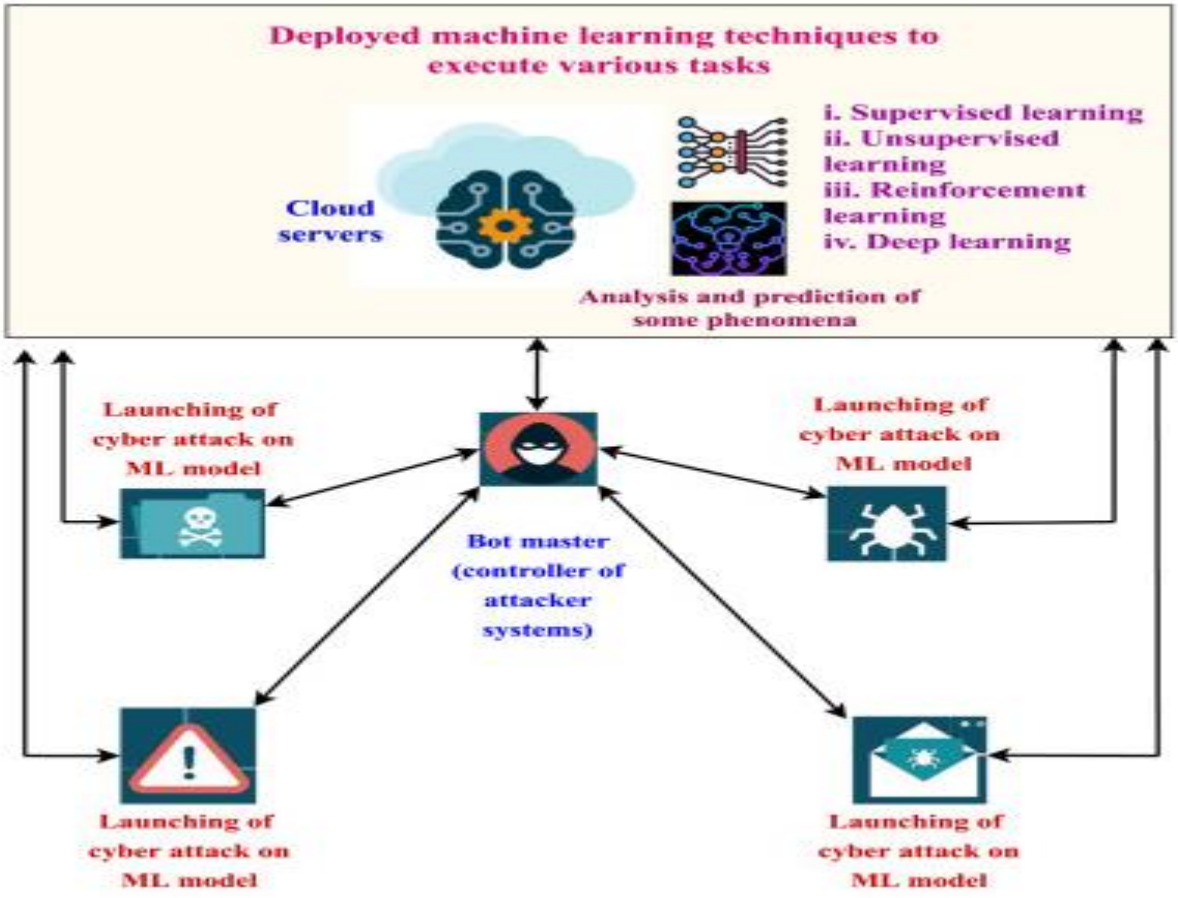

Fig. 2. Scenario Of Cyber Security In Machine Learning