International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
***Article Received:** 30 May 2023 **Revised:** 25 July 2023 **Accepted:** 18 August 2023*

A Review and Analysis on Various Cyber Security Threats on the Basis of Machine Learning Algorithms

Dr. Abhilash Maroju,

Ph.D. Research Graduate, Department of Information Technology, University of the Cumberlands , USA.
doctorabhilashmaroju@gmail.com

Dr. Srinivas A Vaddadi,

PhD Research Graduate, Department of Information Technology, University of the Cumberlands , USA. Vsad93@gmail.com

Dr. Rohith Vallabhaneni,

PhD Research Graduate, Department of Information Technology, University of the Cumberlands , USA.
rohit.vallabhaneni.2222@gmail.com

Sravanthi Dontu,

PhD Research Student, Department of Information Technology, University of the Cumberlands, USA.
sravanthi.dontu13@gmail.com

ABSTRACT

The field of Artificial Intelligence (AI) known as machine learning (or ML for short) helps develop systems that can learn from examples, spot trends, and make rational decisions with little to no human intervention. Cybersecurity methods provide up-to-date security answers for detecting and countering dangers. Security solutions that were formerly adequate are now insufficient since criminals may circumvent traditional security procedures. Cybersecurity refers to the practice of keeping computer systems, servers, mobile devices, networks, and the data associated with them safe from malicious attacks. The two most critical parts of combining ML with cyber security are (1) ensuring that cyber security is taken into consideration whenever ML is used and (2) utilising ML to facilitate cyber security. We hope this union will help us in many ways; for example, by making cyber security procedures more efficient, bolstering the safety of machine learning models, and allowing us to detect zero-day vulnerabilities with little human intervention.

1. INTRODUCTION

Protecting networks and systems against cyber-anomalies and other types of threats has become more important in recent days. Among these forms of assault are worms, botnets, malware, and denial-of-service (DoS) attacks. When these kinds of abnormalities occurred in large-scale computer networks, they caused damage that was irreparable and financial losses [1, 2]. As an illustration, a single ransomware infection that occurred in May of 2017 resulted in enormous losses for a variety of organisations and industries, such as the banking industry, the medical care industry, the power industry, and colleges, and it caused a loss of eight billion dollars [3]. The cybersecurity community has recently seen a rise in discussions surrounding the prevalence of security breaches and invasions. This is true whether we're talking about safeguarding an IoT application or a cyber-system. Although there are a number of traditional ways, such as firewalls, encryption, and so on, that are designed to deal with cyberattacks that are based on the Internet, the most effective way to deal with these problems is to have an intelligent system that can detect anomalies or attacks in an efficient manner. The understanding of artificial intelligence is the primary topic of this research. Machine learning security modelling is of particular interest to us because of the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
***Article Received:** 30 May 2023 **Revised:** 25 July 2023 **Accepted:** 18 August 2023*

automatic learning it can do with training security data, which could lead to better results in application.

Developing security models based on machine learning allows for the aim of delivering intelligent security services by analysing various cyberattacks or abnormalities and, in the end, detecting or forecasting the risks associated with them [4]. As a whole, the detection models can deal with a "binary-class" issue—the detection of anomalies—or a "multi-class" problem—the handling of numerous interconnected cyber-attacks. A plethora of recent studies have covered a wide range of topics in this field, from botnet attack detection [5] to anomaly and attack classification for use in intrusion detection systems [6], from the detection of unusual network connections to the classification of normal traffic and attacks [8], and many more. Despite the widespread use of machine learning techniques, most of them are either too narrow in scope to be useful for security intelligence modelling or too focused on comparing and contrasting the relative importance of various security features. These are briefly covered in Section 2, and Table 1 provides an overview of them. In addition, there are a number of security solutions that may be used in the case of unexpected attacks, aberrant behaviours that are thought to be anomalies, and the relevant model [1, 9]. Hence, intelligent modelling in cybersecurity is necessary for classifying the related attacks into many well-known types, including worm attacks, botnet attacks, malware attacks, and denial of service attacks. Classifying anomalies for unknown assaults coming from normal traffic is also important.

While all of these issues are being considered, different machine learning models may show different levels of effectiveness based on their ability to learn from security data. This is because the data's properties and the importance of the associated security features might affect how effectively a learning-based security model performs. A wide variety of known or undiscovered attack types, abnormalities, or a plethora of security features could all play a role in a real-world scenario's cybersecurity issues. Hence, a strong classification model plus an effective feature selection technique are often what make up an intelligent intrusion detection system.

2. LITERATURE REVIEW

The use of machine learning techniques is implemented in order to improve cybersecurity, recognise websites that are used for phishing, and recognise multiple automated new attacks at an early stage [10]. Machine learning techniques can be broadly grouped into three types: semi-supervised, unsupervised, and supervised. In supervised machine learning, the machine is already familiar with the targeted labelling or classes of data, and it makes use of these labelling and classes in order to train the computer. Unsupervised machine learning does not deliver the value that was intended. The discovery of data relationships is the major objective of unsupervised learning methodology. It does this by searching for patterns in the data, similar to how Clustering works. In the context of machine learning (ML), the term "semi-supervised" refers to a technique in which a portion of the data is tagged or in which human experience is required during the data collection process.

Threats are the possible hazards and dangers that are associated with all of the security breaches that have been mentioned, and assaults are the name given to attempts to cause security breaches [11]One possible interpretation of the term "cybersecurity" is the protection against malicious software and phishing, two of the most destructive types of online attacks.

Commonly referred to as "brand cloning," "phishing" involves impersonating a legitimate user in order to get access to sensitive information for the purpose of manipulation or misuse. This is done in order to get access to the information. One example of phishing is when a person pretends to be a con artist by exploiting the website of a real girlfriend in order to obtain personal information [12]. All forms of malicious software can be broken down into three primary categories: viruses, worms, and Trojan horses. The performance of a computer can be negatively impacted by a piece of software known as a virus, which operates without the user's knowledge. Viruses can harm your computer's operating system and data files. Elk Cloner was the first computer to be able to spread across floppy disc drives [13]. This advancement occurred in 1981.

On a computer, a worm is a piece of software that continuously duplicates itself while simultaneously consuming resources from the system or the network. Trojan horses are a type of malicious software that, in contrast to viruses and worms, are made to appear as legitimate software and are activated by certain procedures or activities rather than replicating themselves. An additional threat to cybersecurity is the sending of unwanted spam emails. These emails not only take a long time to accumulate in your inbox, but they also give rise to the Java applets that execute invisibly as you peruse your inbox. Phone calls, texts, and even video chats are all ways that spam can appear on mobile networks and devices [14]. Messages sent via text message and videos are regularly the targets of spammers on Twitter and YouTube. Network security systems consist of a number of different components, including firewalls, antivirus software, and intrusion detection technologies. The utilisation of intrusion detection systems (IDS) allows for the detection and identification of malicious illegal access as well as unauthorised invasions [15].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
***Article Received:** 30 May 2023 **Revised:** 25 July 2023 **Accepted:** 18 August 2023*

3. MACHINE LEARNING FOR AUTOMATION IN CYBERSECURITY

The digital age we are living in has both positive and negative aspects, just like any other era. An issue with security is the primary downside [16]. The frequency and severity of security breaches are on the rise as we move more and more of our sensitive data online. It is clear that fraudsters are getting better at escaping detection, since many newer malware kits already include creative techniques to do so. But cybersecurity is at a crossroads right now, and rather of focusing on mitigation and defensive measures, researchers should look at cyber-attack prediction systems that can foresee major events and the consequences of such attacks. There is an overwhelming need for solutions that are based on thorough, predictive analyses of cyber threats. Prediction, prevention, identification, detection, and incident response are essential cybersecurity functions that necessitate intelligent and automated performance. Artificial intelligence (AI), mostly ML-based, is the primary focus of this effort [17]. By seeing trends and making predictions based on past actions, AI can find and stop hazardous behaviour.

One of the most talked-about modern technologies in the fourth industrial revolution (4IR or Industry 4.0) is machine learning (ML), which allows systems to learn and get better over time without explicit programming [19]. Machine learning has the potential to revolutionise cyber security by enabling the extraction of important insights from data. Cybersecurity data may be structured or unstructured, and it can originate from many different places (as discussed in Section 3). You can use the data insights to build smart apps for intrusion detection, cyber-attack detection, phishing detection, malware detection, zero-day attack prediction, and more. In the past few days, there has been a meteoric spike in the need for cybersecurity solutions that can ward off cyber abnormalities and attacks like worms, phishing, malware, botnets, spyware, and denial-of-service (DoS). Because of this, intelligent data analysis methods and tools are necessary for practical cyber applications. The ability to quickly and intelligently glean insights or useful information from data is a prerequisite for these methods and technologies. Security researchers are confident that they can develop tools to detect and recognise attack patterns, which will allow them to protect against future attacks.

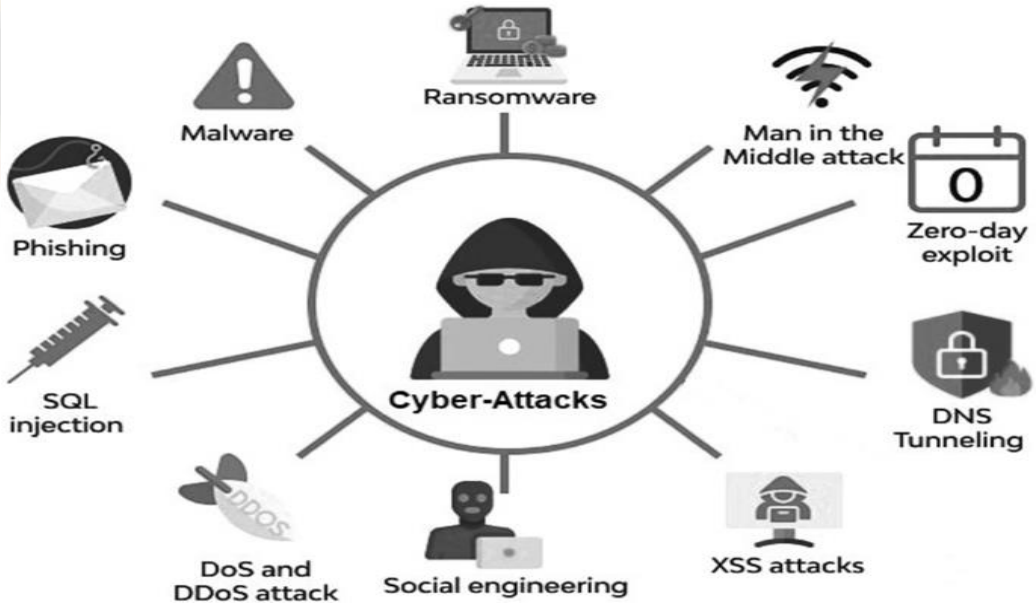

Fig 1: Several typical cybersecurity-related assaults or threats.

Overwhelmed security personnel are turning to automation as a critical tool in the fight against the increasingly broad, sophisticated, and targeted cyber attacks of today. Malware, phishing, ransomware, DoS, zero-day attacks, and other types of cyberattacks are common (Fig. 1). Reason being, no security solution is foolproof, and a lot of current detection methods depend on analysts to manually investigate and make decisions in order to find sophisticated threats, harmful user actions, and other significant related dangers. Machine learning is more effective than humans at identifying and forecasting particular trends. Security judgements and policy modifications in complex and ever-changing network systems have fallen short of expectations [20].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
***Article Received:** 30 May 2023 **Revised:** 25 July 2023 **Accepted:** 18 August 2023*

4. POTENTIAL USE CASES OF MACHINE LEARNING IN CYBERSECURITY

There has been a lot of success in applying machine learning approaches to various cybersecurity concerns across several application domains in the past few years. There are a plethora of typical applications for intrusion detection systems, as seen in Figure 2. Countless tasks fall under this category, including as malware analysis and detection, spam filtering, anomaly and fraud detection, zero-day attack detection, cyberbullying detection, threat analysis, and an extensive list of others.

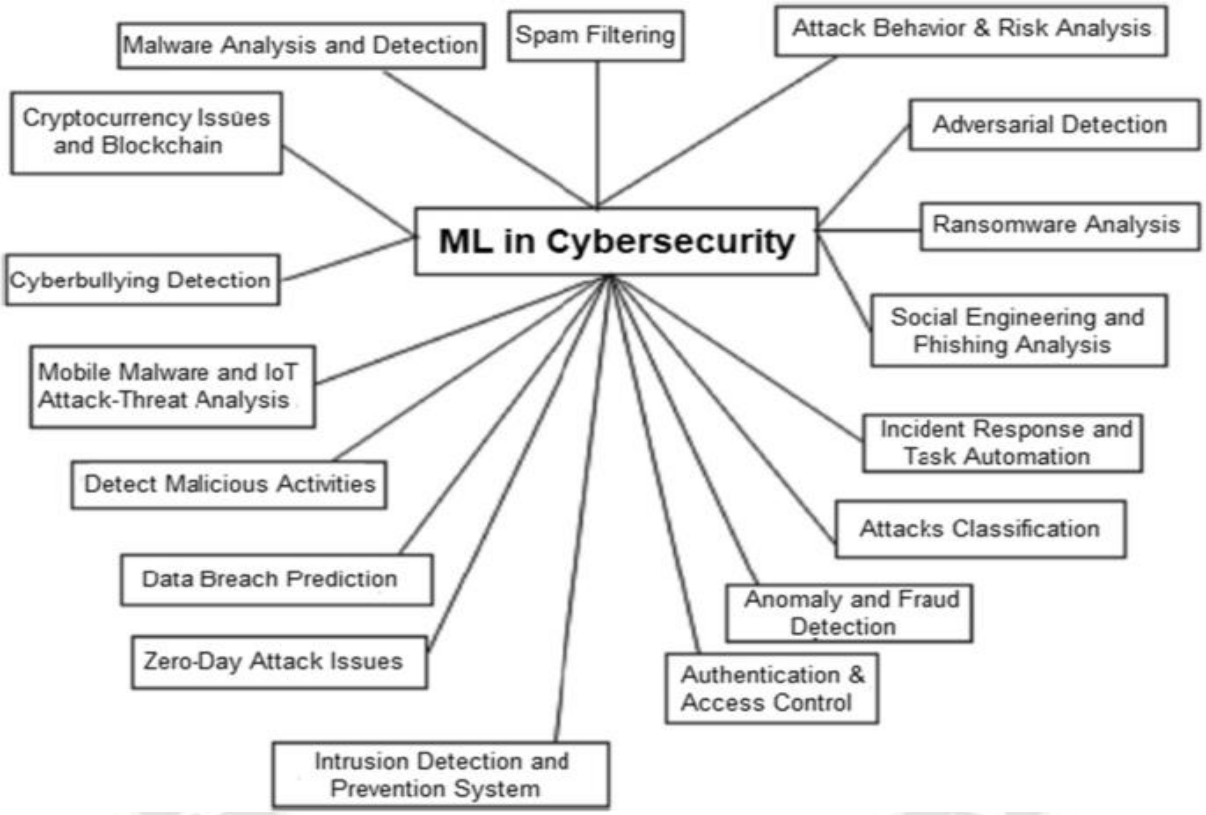

Fig2: Possible applications of AI in cybersecurity.

Machine learning studies can sift through large blockchain datasets in search of anomalies, evaluate market manipulations, and expose fraudulent users. Using machine learning techniques, smart contracts can automatically detect and locate vulnerabilities. One example is the use of classification models to give binary buy/sell trading recommendations in [21], while ML regression models are employed to predict the returns of the bitcoin-based dependent variable.

In the end, the efficacy and efficiency of a machine learning-based solution are determined by the learning algorithms' overall performance and the quality and type of the data used. It is quite difficult to standardise data acquired from endpoints, networks, and clouds and put it to good use in machine learning [22].

CONCLUSION

This article provides a comprehensive review of machine learning methods for automated cybersecurity and smart data analysis. Here, we have included a broad outline of the various machine learning techniques and how they could solve practical issues in the various cyber application fields covered in the essay. How well a machine learning model does depends on the data and how well the learning algorithms work. Before the system can enable intelligent decision-making and automation, the research explores the premise that it needs to train its learning algorithms with real-world cyber data and application-specific information. [23-24] Lastly, we discussed the challenges and potential areas for further research.

REFERENCES

1. M. Hasan *et al.*Attack and anomaly detection in IoT sensors in IoT sites using machine learning approachesInternet Things(2019)
2. I.H. Sarker *et al.*Abc-ruleminer: user behavioral rule-based machine learning method for context-aware intelligent servicesJ. Netw. Comput. Appl.(2020)
3. Y. Li *et al.*An efficient intrusion detection system based on support vector machines and gradually feature removal methodExpert Syst. Appl.(2012)
4. F. Amiri *et al.*Mutual information-based feature selection for intrusion detection systemsJ. Netw. Comput. Appl.(2011)
5. L. Koc *et al.*A network intrusion detection system based on a hidden naïve bayes multiclass classifierExpert Syst Appl(2012)
6. P. Sangkatsanee *et al.*Practical real-time intrusion detection using machine learning approachesComput. Commun.(2011)
7. M. Mazini *et al.*Anomaly network-based intrusion detection system using a reliable hybrid artificial bee colony and adaboost algorithmsJ. King Saud Univ. Comput. Inf. Sci.(2019)
8. I.H. SarkerA machine learning based robust prediction model for real-life mobile phone dataInt. Things(2019)
9. I.H. Sarker *et al.*Intrudtree: a machine learning based cyber security intrusion detection modelSymmetry (Basel)(2020)
10. . L. F. Maimo, A. L. P. Gomez, F. J. G. Clemente, M. G. Perez, and G. M. Perez, "A self-adaptive deep learning-based system for anomaly detection in 5G networks," IEEE Access, vol. 6, pp. 7700-7712, 2018.
11. 16. M. Islam and N. K. Chowdhury, "Phishing Websites detection using machine learning based classification techniques," in Proc. 1st Int. Conf. Adv. Inf. Commun. Technol., 2016, pp. 1-4. S. Marsland, Machine Learning: An Algorithmic Perspective. Boca Raton, FL, USA: CRC Press, 2014.
12. 17. S. R. Granter, A. H. Beck, and D. J. Papke, "AlphaGo, deep learning, and the future of the human microscopist," Arch. Pathol. Lab. Med., vol. 141, no. 5, pp. 619-621, May 2017.
13. 18. D. Yu and L. Deng, "Deep learning and its applications to signal and information processing [Exploratory DSP]," IEEE Signal Process. Mag., vol. 28, no. 1, pp. 145- 154, Jan. 2011.
14. 19. Y. Bengio, "Learning deep architectures for AI," in Foundations Trends Machine Learning, vol. 2, no. 1. Boston, MA, USA: Now, 2009.
15. 20. R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," J. Mach. Learn. Res., vol. 12 pp. 2493-2537, Aug. 2011.
16. Sarker IH (2022) Smart city data science: towards data-driven smart cities with open research issues. Internet Things 19:100528
17. Sarker IH, Asif IK, Yoosef BA, Fawaz A (2022) Internet of things (IoT) security intelligence: a comprehensive overview, machine learning solutions and research directions. Mobile Netw Appl 1–17
18. Sarker IH (2021) Machine learning: algorithms, real-world applications and research directions. SN Comput Sci 2(3):1–21
19. Sarker IH (2021) Cyberlearning: effectiveness analysis of machine learning security modeling to detect cyber-anomalies and multi-attacks. Internet Things 14:100393
20. Shi Y (2022) Advances in big data analytics: theory, algorithms and practices. Springer, Berlin
21. Sebastiao H, Godinho P (2021) Forecasting and trading cryptocurrencies with machine learning under changing market conditions. FinancInnov 7(1):1–30
22. Hagos DH, Yazidi A, Kure O, Engelstad PE (2017) Enhancing security attacks analysis using regularized machine learning techniques. In: 2017 IEEE 31st international conference on advanced information networking and applications (AINA). IEEE, pp 909–918
23. Ramya Manikyam, J. Todd McDonald, William R. Mahoney, Todd R. Andel, and Samuel H. Russ. 2016.Comparing the effectiveness of commercial obfuscators against MATE attacks. In Proceedings of the 6th Workshop on Software Security, Protection, and Reverse Engineering (SSPREW'16)
24. R. Manikyam. 2019.Program protection using software based hardware abstraction.Ph.D. Dissertation.University of South Alabama