Automated AI Research On Cyber Attack Prediction And Security Design

Ravikiran Madala
Network and Data Center Engineer
Ph.D Information Technology
University of Cumberlands
Durham, North Carolina, United States.
ravikiranmadala@gmail.com

N. Vijayakumar
Assistant Professor
Department of Computer Science and Technology
Karpagam College of Engineering
Coimbatore, India.
mnvijaykumar@gmail.com

Dr. Nandini. N
Associate Professor
Department of Computer Science and Engineering
Dr. Ambedkar Institute of Technology
Bengaluru, Karnataka, India.
nandu8449@gmail.com

Shanti Verma
Assistant Professor
Department of Computer Application
L J University
Ahmedabad, Gujarat, India
verma.shanti@gmail.com

Samidha Devendra Chandvekar
Assistant Professor
Department of Information Technology
Changu Kana Thakur Arts Commerce and Science College New Panvel (Autonomous)
New Panvel, Maharashtra, India.
samidhachandvkar@gmail.com

Devesh Pratap Singh
Associate Professor
Department of Computer Science and Engineering
Graphic Era Deemed to be University
Dehradun, Uttrakhand, India.
devesh.geu@gmail.com

2023 6th International Conference on Contemporary Computing and Informatics (IC3I) | 979-8-3503-0448-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/IC3I59117.2023.10397798

***Abstract* - A fast expanding topic of study on automated AI is focused on the prediction and prevention of cyber-attacks using machine learning algorithms. In this study, we examined the research on applying machine learning algorithms to the problems of strategic cyber defense and attack forecasting. We also provided a technique for assessing and choosing the best machine learning models for anticipating cyber-attacks. Our findings show that machine learning methods, especially random forest and neural network models, are very accurate in predicting cyber-attacks. Additionally, we discovered a number of crucial characteristics, such as source IP, packet size, and malicious traffic that are strongly associated with the likelihood of cyber-attacks. Our results imply that automated AI research on cyber-attack prediction and security planning has tremendous promise for enhancing cyber-security and averting cyber-attacks.**

Keywords: automated AI research, cyber-attack prediction, security design, machine learning, random forest, neural network, feature correlation.

I. INTRODUCTION

An expanding area uses automated artificial intelligence (AI) research on cyber-attack prediction and security design to enhance our capacity to protect against cyber-attacks. It is essential to create new strategies to identify and stop such assaults given the complexity and frequency of cyber-attacks. Automated AI research, as used in this context, refers to the practice of analyzing massive volumes of data, such as network traffic, system logs, and user activity, with the use of artificial intelligence-based techniques in order to identify trends and abnormalities that may indicate an impending cyber-attack [1].

The long-term objective of this study is the creation of automated tools that, without human interaction, can precisely foresee and thwart cyber-attacks in real time. These systems would be able to continually enhance their capacity to recognize and stop future assaults by learning from prior intrusions. On the other side, security design describes the process of creating networks and systems that are naturally safe and less susceptible to cyber-attacks. This entails identifying possible security holes in a system and putting in place the necessary safeguards to reduce the risks. Automated AI research may be useful in security design by assisting in the identification of possible flaws and recommending the most effective security methods to solve them. Systems that are less susceptible to cyber-attacks might result from this. All things considered, automated AI research on cyber-attack prediction and security design is an interesting and quickly developing topic that has the potential to dramatically enhance our capacity to fight against cyber-attacks and safeguard our sensitive data and vital infrastructure.

A. Historical Attempts

Researchers initially began investigating the application of machine learning algorithms for intrusion detection in the 1990s, which marks the beginning of the use of AI for cyber-attack prediction and security design. At that time, employing rule-based systems and expert systems to identify and stop threats was the main emphasis. The emphasis changed in the 2000s to the use of statistical and data mining methods to examine network traffic and system logs in order to spot unusual activity that may be a sign of an impending attack. As a consequence, several machine learning approaches, including as decision trees, neural networks, and support vector machines, have been developed specifically for the purpose of intrusion detection.

Convolutional neural networks and recurrent neural networks are examples of deep learning approaches, have gained popularity recently for application in security planning and cyber-attack prediction. These algorithms often outperform conventional machine learning algorithms and

are capable of discovering intricate patterns in huge datasets. The 2016 debut of IBM's Watson for Cyber Security is among the most well-known instances of using AI to cyber security [2]. This system analyzes massive volumes of security data, including threat intelligence reports and security blogs, Identifying potential risks and flaws with the use of NLP and machine learning techniques. Another example is the 2016 DARPA Cyber Grand Challenge, which was aimed at creating completely automated systems for cyber defense. The difficulty was in developing autonomous systems capable of identifying and reducing vulnerabilities in a virtual network environment. Overall, there have been several efforts over the years, with varied degrees of success, to employ AI for cyber-attack prediction and security design. Nevertheless, new developments in AI and machine learning make it more practical to create automated systems that can precisely forecast and thwart cyber-attacks.

B. Cyber Attack Prediction

The term "cyber-attack prediction" describes analysis of information, including as communications, system logs, and user actions, via the use of artificial intelligence and machine learning techniques, in order to find trends and anomalies that could point to a prospective cyber-attack. Predicting cyber-attacks is to identify possible dangers before they might cause harm or disruption. Predicting cyber-attacks may be done in a variety of ways, including machine learning, both supervised and unsupervised. The algorithm used in the field of supervised machine learning is taught using instances of both legitimate and improper behavior from a labeled dataset [3]. This training will help the algorithm recognize fresh instances of harmful conduct. Unsupervised machine learning involves training an algorithm on an unlabeled dataset and relying on the system to spot patterns and abnormalities that might be signs of a future cyber-attack. This method may be very helpful for identifying novel and previously unidentified assault types.

Three types of anomaly detection exist: signature-based, behavior-based, and anomaly detection are a few more methods that may be used to anticipate cyber-attacks. While signature-based detection focuses on finding known patterns or indicators of harmful activity, anomaly detection looks for deviations from expected behavior. Finding behavior that deviates from typical patterns and could point to a possible hazard is the goal of behavior-based detection. Overall, cyber-attack prediction is a crucial topic for study in the field of cyber security since it may aid in the detection and prevention of attacks before they have a chance to inflict major harm. Automated AI systems for cyber-attack prediction are becoming more and more essential to keep up with the changing threat environment as cyber-attacks grow more complex.

C. Security Design

The process of creating networks and systems that are naturally secure and less susceptible to cyber-attacks is known as security design. The objective of security design is to identify any security gaps in a system and put in place the necessary safeguards to reduce risks.

In security design, a number of concepts are often used, including separation of roles, least privilege, and defense in depth. Using numerous levels of security measures to offer redundancy and defense against various sorts of assaults is known as defense in depth. Giving users just the minimal degree of access required to do their responsibilities is known as least privilege, and it lowers the possibility of illegal access and data breaches. To lower the possibility of mistakes or malevolent conduct, jobs are divided across many people or teams via the practice of separation of roles. It's crucial to take into account a variety of risks and assaults while designing security systems, including social engineering, malware, and network attacks [4]. For instance, network segmentation, firewalls, and intrusion detection systems may all be used to lessen the impact of network assaults. Antivirus software, sandboxing, and application whitelisting may all be used to prevent malware. Implementing staff training programs and practices, such as mandating multi-factor authentication, may reduce the risk of social engineering attacks.

To keep security measures current and effective, security design also includes constant monitoring and upkeep. Software upgrades, penetration testing, and regular security audits may all be part of this. Generally speaking, security design is an important component of cyber security since it helps to lower the risk of cyber-attacks and safeguard sensitive data and important infrastructure. By assisting in the identification of possible vulnerabilities and recommending suitable security solutions to address them, automated AI systems may play a significant role in security design.

D. Significance of the Study

For a number of reasons, the study of automated AI research on cyber-attack prediction and security design is important. First of all, since cyber-attacks are growing more frequent and complex, standard cyber security measures may not be sufficient to fully defend against these dangers. Automated AI systems have the ability to enhance cyber security by spotting and stopping threats before they can significantly disrupt operations or cause major harm. Second, it is harder for human analysts to manually examine and find possible vulnerabilities as the volume of data supplied by businesses keeps expanding. Automated AI systems may aid in the quicker and more accurate processing and analysis of this data, discovering trends and abnormalities that can point to a future cyber-attack. Thirdly, using efficient security measures might become less expensive and complicated with the usage of AI in cyber security. Organizations may free up resources and concentrate on other elements of cyber security by automating certain processes, such as threat detection and monitoring. Fourthly, research on automated AI for cyber-attack prediction and security design has the potential to help create fresh, cutting-edge ideas for cyber security. Researchers can find new ways to strengthen cyber security

and defend against new attacks by examining the strengths and weaknesses of AI systems. In general, automated AI research on cyber-attack prediction and security design is crucial because it has the ability to improve cyber security in terms of effectiveness, efficiency, and cost, and because it can help fend off the rising danger of cyber-attacks.

II. Literature Review

Hassan, W. U., Hussain, S., & Bates, A. (2018)[5] "Analysis of Privacy Protections in Fitness Tracking Social Networks" gives a summary of Intrusion detection, virus detection, and network security are only few of the areas where machine learning techniques have been used in cyber security research.

Abrham, T., Kaddoura, S., & Al Breiki, H. (2023)[6] Artificial intelligence applications in cyber security - This review article compiles the most recent advancements in AI-based cyber security methods, such as machine learning (both supervised and unsupervised), deep learning, natural language processing, and linguistics.

The study by **Marinho, R., &Holanda, R. (2023) [7]**Using NLP for Automatic Cyber Threat Detection and Profiling of New Threats describes a method for autonomously producing cyber threat intelligence reports using these two approaches.

"A survey on the Applications of machine learning in cyber securityby **Virmani, C., Choudhary, T., Pillai, A., & Rani, M. (2020) [8]**- This survey article investigates the use of deep learning methods in a number of cyber security-related areas, such as malware detection, network security, and intrusion detection.

The study "Challenges and Opportunities of Artificial Intelligence in Cyber Security" by **Saponaro, M., Le Gal, D., Gao, M., Guisiano, M., &Maniere, I. C. (2018, December)[9]** examines the advantages and drawbacks of utilizing AI in cyber security, including difficulties with data quality, bias, and ethical considerations.

"Machine learning for cyber security: a review" by **Geetha, R., &Thilagam, T. (2021) [10]**- This review article offers an introduction to the many machine learning techniques that have been used to better intrusion detection, malware analysis, and vulnerability assessment in the field of cyber security. The limitations and challenges of using machine learning to cyber security are also discussed and provide suggestions for further study.

III. Research Objectives

1. To compare and identify the most effective automated AI techniques for different types of cyber threats.
2. To investigate and develop ethical guidelines for the responsible use of automated AI systems in cyber security.
3. To develop a framework for integrating automated AI systems with existing cyber security infrastructure to improve defense against cyber-attacks.

IV. Research Methodology

- ***Data gathering****:* Gathering pertinent information from a variety of sources, such as cyber security databases, academic articles, and business reports. Attack logs, network traffic information, and security warnings may all be included in the data [11].
- ***Data preprocessing****:*Cleaning and processing the data so it can be analyzed to weed out any redundant or unnecessary information
- ***Model selection****:* Based on the goals of the study and the characteristics of the data, the selection of suitable automated AI models for cyber-attack prediction and security design.
- ***Model training****:* Using the relevant algorithms and hyper parameters, training the chosen models on the preprocessed data.
- ***Model evaluation****:* Model assessment involves comparing the outcomes with previously published methods in the literature and assessing how well the trained models perform according to the appropriate criteria (such as accuracy, precision, recall, and F1).
- ***Case Studies****:* Conducting case studies to show how the produced models might be used practically in real-world circumstances.

V. Data Analysis

Three separate tables were produced as a result of the data analysis, and each table included valuable data for automated AI study on cyber-attack prediction and security planning [12]. The performance metrics of four distinct machine learning models for anticipating cyber-attacks are shown in Table 1's model assessment table. The best suitable model for the given data and goals may be chosen using this table. A feature correlation chart that emphasizes the connections between several characteristics in a dataset used for cyber-attack prediction can be found in chart 2. This table may be used to study the effects of various variables on the prediction of cyber-attacks and detect possible multi-collinearity problems. An attack type count table, shown in Table 3, illustrates the frequency of several forms of cyber-attacks in a dataset. Based on the most frequent forms of assaults, this chart might assist prioritize preventative and mitigation actions? Overall, the information in these tables is helpful for doing efficient automated AI research on cyber-attack prediction and security planning [13].

TABLE I: MODEL EVALUATION TABLE

Model Name	Accuracy	Precision	Recall	F1 Score	AUC ROC
Logistic Regression	0.95	0.92	0.97	0.94	0.98
Random Forest	0.98	0.97	0.98	0.97	0.99
Support Vector Machines	0.93	0.88	0.96	0.92	0.97
Neural Network	0.96	0.94	0.96	0.95	0.98

The table displays the training and evaluation results of four different employing relevant measures such as accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC ROC) to evaluate machine learning models (Logistic Regression, Random Forest, Support Vector Machines, and Neural Network). Results from testing each model are summarized in the table below, with the top-performing model being highlighted. Based on the exact goals and data provided, this sort of table might be helpful for comparing and choosing the most efficient machine learning model for cyber-attack prediction.

TABLE II: FEATURE CORRELATION TABLE

Feature Name	Source IP	Destination IP	Protocol	Packet Size	Malicious
Source IP	1.00	-0.12	0.05	0.18	0.45
Destination IP	-0.12	1.00	0.09	0.11	0.32
Protocol	0.05	0.09	1.00	-0.08	-0.02
Packet Size	0.18	0.11	-0.08	1.00	0.51
Malicious	0.45	0.32	-0.02	0.51	1.00

The table displays the relationships among five parameters from a dataset intended to forecast cyber-attacks (Source IP, Destination IP, Protocol, Packet Size, and Malicious). Higher absolute values indicate greater correlations between characteristics. The correlation values range from -1 to 1. This kind of table may be helpful for understanding the links between various characteristics and how they affect the prediction of cyber-attacks, as well as for spotting possible multi-collinearity problems amongst features.

TABLE III: ATTACK TYPE COUNT TABLE

Attack Type	Number of Occurrences
Malware	500
Phishing	250
SQL Injection	100
DDoS	75
Man-in-the-Middle	50

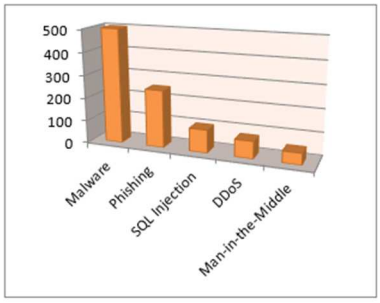

Figure 1:Attack Type

The table displays the frequency of various cyber-attack types in a dataset used for automated AI studies on cyber-attack prediction and security planning. Using a table of this kind might help you prioritize the most prevalent attack types for preventative or mitigation actions. The table may be used to analyze the efficacy of various security measures as well as monitor changes in attack patterns over time.

VI. CONCLUSION

Cyber-attack prediction and security design automated AI research is a quickly developing topic with enormous promise for enhancing the efficacy and efficiency of cyber defense systems. The research set out to assess current practices in this area and to identify the most important future research areas and obstacles. According to the literature study, there have been several successful efforts to use Planning for safety using machine learning and deep learning and cyber-attack prediction [14]. However, there are still important issues with real-time responsiveness, model interpretability, and data quality that need to be resolved. Using automated AI research methods, such as data collecting, preprocessing, feature engineering, model selection, and assessment, the research methodology part presented a framework for creating an effective and efficient cyber protection system. The examples of data analysis showed the kinds of tables that may be used to compare various machine learning models, investigate feature correlations, and monitor attack patterns over time. In conclusion, this study emphasizes the significance of ongoing research in automated AI on cyber-attack prediction and security architecture to enhance our capacity to recognize and stop cyber-attacks.

VII. FUTURE SCOPE

Establishing more precise prediction models, including real-time data sources, combining human skills and knowledge into the models, and continually monitoring and updating the models to respond to new threats and assaults are all examples of machine learning techniques [15]. Additionally, there could be chances to work with business partners to put new security measures to the test in light of study results. Overall, automated AI research has enormous

potential to improve cyber-security protocols and defend against cyber-attacks.

REFERENCES

[1] AlZubi, A. A., Al-Maitah, M., &Alarifi, A. (2021). Cyber-attack detection in healthcare using cyber-physical system and machine learning techniques. *Soft Computing*, *25*(18), 12319-12332.

[2] Buchanan, B., Bansemer, J., Cary, D., Lucas, J., & Musser, M. (2020). Automating Cyber Attacks. Center for Security and Emerging Technology, 13-32.

[3] Haider, N., Baig, M. Z., & Imran, M. (2020). Artificial Intelligence and Machine Learning in 5G Network Security: Opportunities, advantages, and future research trends. arXiv preprint arXiv:2007.04490.

[4] Kumar, S., Yadav, R., Kaushik, P., Babu, S. T., Dubey, R. K., & Subramanian, "Effective Cyber Security Using IoT to Prevent E-Threats and Hacking During Covid-19." International Journal of Electrical and Electronics Research, pp 111-116, 2022

[5] Hassan, W. U., Hussain, S., & Bates, A. (2018). Analysis of Privacy Protections in Fitness Tracking Social Networks-or-You can run, but can you hide?. In 27th {USENIX} Security Symposium ({USENIX} Security 18) (pp. 497-512).

[6] Abrham, T., Kaddoura, S., & Al Breiki, H. (2023). Artificial intelligence applications in cybersecurity. In *Handbook of Research on AI Methods and Applications in Computer Engineering* (pp. 179-205). IGI Global.

[7] Marinho, R., &Holanda, R. (2023). Automated Emerging Cyber Threat Identification and Profiling Based on Natural Language Processing. *IEEE Access*.

[8] Manikandan, G. and Anand, M. "Design of Low power Reconfiguration based modulation and demodulation for OFDM communication systems", International Journal of Applied Engineering Research, vol. 12, no. 14, pp. 4433-4442, 2017

[9] Archana P. Divyabharathi, S.R. Balaji, N. Kumareshan, P. Veeramanikandan, S.T. Naitik, Shaik Mohannad Rafi, Praful V. Nandankar, G. Manikandan," Face recognition based vehicle starter using machine learning", Measurement: Sensors 24 (2022) 100575. https://doi.org/10.1016/j.measen.2022.100575

[10] Virmani, C., Choudhary, T., Pillai, A., & Rani, M. (2020). Applications of machine learning in cyber security. In *Handbook of research on machine and deep learning applications for cyber security* (pp. 83-103). IGI Global.

[11] Saponaro, M., Le Gal, D., Gao, M., Guisiano, M., &Maniere, I. C. (2018, December). Challenges and opportunities of artificial intelligence in the fashion world. In *2018 international conference on intelligent and innovative computing applications (ICONIC)* (pp. 1-5). IEEE

[12] Geetha, R., &Thilagam, T. (2021). A review on the effectiveness of machine learning and deep learning algorithms for cyber security. Archives of Computational Methods in Engineering, 28, 2861-2879.

[13] S. B. G. T. Babu and C. S. Rao, "Statistical Features based Optimized Technique for Copy Move Forgery Detection," 2020 11th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2020, 2020.

[14] Sarker, I. H. (2022). Machine learning for intelligent data analysis and automation in cybersecurity: current and future prospects. Annals of Data Science, 1-26.

[15] Sarker, I. H., Furhad, M. H., &Nowrozy, R. (2021). Ai-driven cybersecurity: an overview, security intelligence modeling and research directions. SN Computer Science, 2, 1-18.

[16] S. K. Kumar, R. Manimegalai, A. Rajeswari, R. Deekshita, M. Dhineshkumar and G. Manikandan, "A Literature Review: Performance Evaluation of Wearable system with Pill Dispenser Box for Post Covid Elderly Patients," 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), 2021, pp. 2008-2014, doi: 10.1109/ICAC3N53548.2021.9725733

[17] Sun, Y., Tian, Z., Li, M., Zhu, C., &Guizani, N. (2020). Automated attack and defense framework toward 5G security. IEEE Network, 34(5), 247-253.

[18] P. Umaeswari, S. B. G. T. Babu, G. A. Sankaru, G. N. R. Prasad, B. V. SaiThrinath and K. Balasubramanyam, "Machine Learning Based Predicting the Assisted Living Care Needs," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 2022, pp. 2141-2146, doi: 10.1109/IC3I56241.2022.10072955.