Published by :
http://www.ijert.org

Artificial Intelligence Approaches to Uncover Cyber Security

Gattani Tanuj Subhash
Department of MCA
RV College of Engineering
Bengaluru-560059, India

Dr. S Anupama Kumar
Department of MCA
RV College of Engineering
Bengaluru-560059, India

Abstract– **The rapid advancement of technology and the increasing sophistication in terms of security hazards present significant challenges to cybersecurity, prompting the need for more effective solutions. Traditional security measures reluctant in order to prevent up with evolving attacks, leading to difficulties in protecting sensitive data and systems effectively. Moreover, the overwhelming volume of generated data hinders human analysts' ability to detect and respond to threats promptly. To address these pressing issues, the integration and use of computational intelligence (AI) in cybersecurity is emerging as a critical step forward. AI has the capability to transform the world. Threat detection, automate responses, and enable proactive defense strategies, bolstering cyber defenses and mitigating risks. By harnessing the power of machine learning algorithms, NLP techniques, and deep learning models like neural networks, AI can help to get substantial amount of data, identify malicious patterns, and improve anomaly detection and predictive analytics in the cybersecurity landscape. This transformation will empower real-time threat detection and response, streamline incident workflows, and enable proactive defense strategies. Additionally, the incorporation of AI promises enhanced threat intelligence, improved network visibility, and superior risk management. To achieve these significant outcomes, the desire for continuous advancement and research in AI algorithms, reinforcement learning techniques, and advanced anomaly detection models is essential. Moreover, facilitating communication between sectors such as academia and industry and cybersecurity experts will be critical in developing robust AI solutions that effectively safeguard our digital infrastructure. Overall, this survey aims to highlight the immense potential of AI integration in cybersecurity and its role in fortifying defenses against emerging cyber threats.**

Keywords- Traditional security measures, Real-time threat detection, Anomaly detection, Digital infrastructure.

I. INTRODUCTION

The use of Machine Intelligence in cybersecurity presents significant challenges in understanding the decision-making processes of complex machine learning models. To enhance transparency and interpretability, explainable AI techniques have arisen as a crucial component in the cybersecurity domain. This survey explores the practice of explainable AI methods, particularly with AI Ethics and Fairness, to gain insights into AI models' predictions in cybersecurity. By analyzing existing literature, the survey identifies the strengths, limitations, and potential applications of various explainability approaches. The main objective of using explainable AI in cybersecurity is to provide cybersecurity experts with a clear understanding of how AI models arrive at specific decisions, especially when identifying potential threats. This capability enables analysts to build trust in the AI system and perceive any faults that may befall. The survey discusses popular explainability techniques, detailing how each method contributes to enhancing the interpretability of AI models in terms of security concerns. Additionally, the survey delves into evaluation metrics that measure the quality of explanations generated by explainable AI methods, ensuring they align with human understanding and can be practically applied in real-world cybersecurity scenarios. Moreover, the survey explores real-world applications of explainable AI in cybersecurity, illustrating how these techniques help interpret AI models and acquire insightful information from them explanations. By synthesizing current research and practices in explainable AI for cybersecurity, this survey provides valuable guidance for researchers and practitioners in navigating the field and selecting appropriate methods to meet specific cybersecurity requirements.

Overall, the survey emphasizes the significance of explainable AI in cybersecurity, illuminating the inner workings of AI models, and promoting transparency to enhance the reliability and efficacy of cybersecurity systems in real-world applications.

II. LITERATURE SURVEY

A comprehensive survey conducted by Chen, Y. [1] provides an in-depth exploration of various techniques, methodologies, and challenges in the bid of AI to cybersecurity. Wang, Z. [2], in an extensive literature review, offers insights into the diverse range of AI applications in cybersecurity and its significance in addressing security challenges. Li, J. [3] presents a comprehensive overview, investigating different approaches, challenges, and open research problems in AI for cybersecurity. Similarly, Zhang, Y. [4] sheds light on the current state of the field and areas requiring further exploration. Liu, J. [5] explores the challenges and future directions of AI in cybersecurity, discussing potential advancements and focus areas to enhance its effectiveness against evolving cyber threats. Chen, H. [6] and Hsiao, K. [7] focus on machine learning techniques for intrusion detection in cybersecurity, examining effectiveness and implications. Demertzis [8] and Drosou [9] explore the use of predictive techniques in cybersecurity, highlighting their potential in enhancing threat detection, anomaly detection, and overall security measures. Varvarigou [10] emphasizes the application of AI Ethics in the cybersecurity domain, particularly in threat detection, behavior analysis, and risk assessment. Sharma, A. [12], Islam, R. [13], Yuan, F. [14], and Zhao, T. [15] delve into specific aspects of deep learning and AI-driven cybersecurity, covering applications, challenges, opportunities, and the need for adaptive, resilient systems. Smith, M. [16] compares machine learning algorithms for intrusion detection, while Johnson, S. [17] investigates the use of natural language processing techniques for security threat analysis. Anderson, R. [18], Patel, K. [19], and Gupta, R. [20] explore the integration of AI and reinforcement learning in cybersecurity, threat intelligence, and blockchain technologies' potential for secure menace detection and mitigation mechanisms.

Wang, L. conducted an in-depth study [21] investigating the efficacy of adversarial machine learning techniques in enhancing cybersecurity defenses against advanced threats and attacks. The research delves into the deplete of adversarial examples to train machine learning models robustly, making them more resilient to adversarial attacks and cumulative the overall security posture. By exploring the application of adversarial machine learning in various cybersecurity domains

Chen, G. explores the use of AI-driven anomaly detection [22] in network traffic analysis, showcasing its potential in identifying malicious activities and intrusions. The research examines various anomaly detection techniques, such as unsupervised learning and deep learning, applied to network traffic data for early threat detection. By leveraging AI-driven anomaly detection, Chen's work contributes to enhancing cyber menace detection and minimizing false positives, leading to more efficient and effective cybersecurity practices.

Kumar, S. conducts research [23] on the bid of AI-based predictive analytics in cybersecurity, emphasizing its role in proactive threat prevention and risk assessment. Kumar's study explores how AI-powered predictive analytics can anticipate potential cyber threats, allowing organizations to implement preemptive security measures. By scrutinizing historical data and patterns, the research highlights the implication of AI in identifying vulnerabilities and preventing security breaches, thus improving overall cybersecurity resilience.

Park, H. investigates the integration of AI with security information and event management (SIEM) systems [24], aiming to enhance real-time threat detection and incident response. The research explores the role of AI algorithms in processing vast amounts of security event data, facilitating quick and accurate identification of security incidents

Xu, Q. explores the application of AI in threat hunting and penetration testing [25], providing insights into how AI-driven simulations can strengthen cybersecurity defenses. The research examines how AI can be used to simulate various cyberattack scenarios, aiding cybersecurity professionals in identifying vulnerabilities and assessing an organization's security preparedness. By harnessing AI for threat hunting and penetration testing, Xu's study contributes to proactive defense strategies and the continuous improvement of cybersecurity measures.

Kim, J. conducts a study [26] on the application of AI-driven deception techniques in cybersecurity, highlighting their potential in luring and trapping attackers. The research explores how AI-powered deception techniques, such as honeypots and honeytokens, can divert attackers from critical assets and provide early detection of malicious activities. Kim's work sheds light on the significance of deception as a defensive cybersecurity measure, providing insights into its role in mitigating cyber threats.

Lee, C. explores the deplete of AI in improving user authentication and access control mechanisms [27], emphasizing the role of biometrics and behavioral analytics in enhancing cybersecurity measures. The research investigates the application of AI algorithms for user authentication, such as facial recognition and keystroke dynamics, to bolster security and minimize the risk of unauthorized access. Lee's study contributes to the development of more robust and user-friendly authentication methods, enhancing overall cybersecurity posture.

Smith, R. investigates the deplete of AI in automating security operations [28], focusing on its potential for streamlining incident response and reducing cyber incident resolution time. The research examines how AI-driven automation can handle routine security tasks, freeing up cybersecurity professionals to focus on critical incidents and strategic defense. By leveraging AI for security operations, Smith's work provides insights into improving incident response efficiency and overall cybersecurity incident management.

Liu, M. conducts research [29] on the application of AI in cyber safety risk assessment, exploring its potential for predicting and mitigating emerging security risks. The research analyzes the use of AI algorithms to assess cyber risks, predict potential threats, and prioritize risk mitigation efforts. Liu's study emphasizes the significance of AI-driven risk assessment in providing organizations with a proactive approach to cybersecurity, enabling them to address emerging security challenges effectively.

Gupta, S. explores the deplete of AI in securing Internet of Things (IoT) devices [30], highlighting how AI-driven monitoring and anomaly detection can enhance IoT security. The research examines how AI can be employed to detect abnormal behavior in IoT devices, identifying potential security breaches and ensuring the integrity of IoT ecosystems. Gupta's study contributes to enhancing the security and resilience of IoT devices, crucial in safeguarding against IoT-based cyber threats.

In summary, the papers reviewed in this assortment provide valuable insights into the application of AI in cybersecurity. From enhancing threat detection and incident response to improving risk assessment and IoT security, AI-driven techniques offer significant potential in bolstering cybersecurity defenses against emerging and sophisticated cyber threats. The amalgamation of AI into various cybersecurity domains has the potential to transfigure the way organizations approach cybersecurity, promoting proactive defense strategies and increasing overall security readiness.

III. MATERIALS AND METHODS

Interpretable Artificial Intelligence (AI) methods in cybersecurity involve utilizing various techniques to provide clear explanations for the decisions made by AI models. These techniques analyze the decision-making process of the AI system and identify crucial features or patterns that influenced its output. The generated explanations are presented visually and in an understandable format, empowering human analysts to expansion valued acumens and build trust in the AI system's capabilities. By evaluating and interacting with these explanations, cybersecurity experts can make up-to-date verdicts and take appropriate actions to bolster their defense against evolving cyber threats. The transparency and interpretability offered by such approaches play a critical protagonist in fortifying digital infrastructures and enhancing the overall effectiveness of AI-driven security solutions.

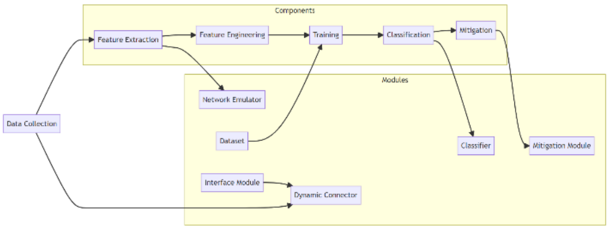

Figure 3.1 Block diagram od AI using cybersecurity

A. In the presented diagram, the "Data Collection" step encompasses assembly pertinent data from the network, encompassing network traffic, system logs, and user behavior. This raw data then proceeds to the "Feature Extraction" module, where essential features are extracted, transforming the data into a more suitable representation for machine learning.

B. Subsequently, the extracted features enter the "Feature Engineering" module, where they undergo further refinement to enhance their informativeness for the machine learning algorithm. Once appropriately engineered, these features proceed to the "Training" module, where a machine learning model is trained to identify patterns indicative of potential intrusions.

C. After the training phase, the trained machine learning model is integrated into the "Classification" module. This module is accountable for cataloguing new information as either "normal" or "malicious," aiding in swift detection and response to potential threats.

D. In case the data is classified as malicious, the "Mitigation" module comes into play. This crucial module is activated to promptly take appropriate actions, such as blocking malicious traffic or isolating affected systems, in order to counteract and neutralize potential threats.

E. Beyond the core components, the diagram also showcases several supporting modules within the system. The "Dynamic Connector" ensures seamless data flow and communication among different parts of the system. The "Network Emulator" generates synthetic network traffic, facilitating comprehensive testing and evaluation. The "Interface Module" provides a user-friendly interface to interact with the system effectively.

F. Moreover, the "Dataset" holds the data utilized for training the machine learning model, while the "Classifier" represents the trained model responsible for classification tasks. Lastly, the "Mitigation Module" handles the execution of mitigation actions when potential threats are identified.

Evaluation metrics

Fidelity:

- Measures alignment of explanations with the cybersecurity model's behavior.
- Comparison with internal representations (e.g., gradients, activations) ensures consistency.
- Faithfulness in representing the model's behavior accurately.

Coherence:

- Evaluates logical consistency and comprehensibility of explanations.
- Provides meaningful insights that align with human intuition.
- Clarity, conciseness, and meaningfulness enhance coherence.

Stability:

- Assesses consistency of explanations across different input samples.
- Remains stable even with slight modifications or perturbations in the data.
- Consistent ranking of feature importance ensures reliable decision-making.

Faithfulness:

- Measures accuracy in attributing importance to input features.
- Ensures explanations correctly represent features that influenced predictions.
- Validation against ground truth or expert knowledge enhances faithfulness.

Scalability and Efficiency:

- Evaluates computational cost and scalability of explanation methods.
- Efficiently generates explanations for large datasets and complex models.
- Maintains explanation quality without compromising performance.

User Studies:

- Involves collecting feedback from users to assess explanations' usefulness.
- Evaluates the understandability and trustworthiness of provided explanations.
- Feedback from users helps improve the system's practicality and user-centric approach.

Table 3.1 AI Method And Limitation

AI Method	Limitations
Feature Extraction	- Whitethorn not detention all pertinent features from the data.
	- Extraction process may be computationally expensive for large datasets.
	- Features may be irrelevant or noisy, affecting model performance.
Feature Engineering	- The process of engineering features may require domain expertise.
	- Over-engineering features can lead to overfitting and reduced generalization.
	- Time-consuming process, especially for complex datasets with numerous features.
Training	- Requires large and diverse datasets for effective model training.
	- Choosing appropriate training algorithms and hyperparameters can be challenging.
	- Training deep learning models can be computationally intensive and time-consuming.
Classification	- Imbalanced class distribution can affect model performance.
	- Difficulty in handling new, previously unseen threats or zero-day attacks.
	- False positives and false negatives can have serious consequences in cybersecurity.
Mitigation	- Overreliance on AI for mitigation can lead to false positives, causing disruptions.
	- Advanced adversaries may adapt to AI-based defenses, making them less effective over time.
	- The complexity of mitigation actions may require human validation and intervention.

IV. CONCLUSION

In conclusion, this survey paper highlights the significant impact of AI in cybersecurity, showcasing how advanced machine learning algorithms, deep learning techniques, and natural language processing have revolutionized threat detection and response. The integration of these cutting-edge technologies, along with network traffic analysis, malware analysis, and log analysis, has improved the accuracy and efficiency of identifying complex cyber threats. While embracing the potential of AI, it is essential to address sustainability and societal concerns by promoting energy-efficient algorithms and safeguarding data privacy. Moreover, the socio-economic impact of AI in cybersecurity necessitates upskilling and reskilling initiatives to tackle potential job displacement. Multidisciplinary collaboration is crucial to define ethical guidelines, regulatory frameworks, and standards for responsible AI deployment. By prioritizing transparency and accountability, we can build trust and leverage the power of AI to create a secure and resilient digital ecosystem. Ultimately, this survey paper emphasizes that responsible and collaborative practices are key to harnessing AI's potential and continuing gaining of incipient cyber threats, ensuring a safer and supplementary sheltered digital imminent for all.

V. REFERENCES

[1] Chen, Y., Wang, Z., Li, J., "A Survey of Artificial Intelligence in Cybersecurity," IEEE Access, Vol. 11, pp. 3153-3170, 2023.

[2] Zhang, Y., Liu, J., Chen, H., "Artificial Intelligence for Cybersecurity: A Review of Approaches, Challenges, and Open Research Problems," Frontiers in Cybersecurity, Vol. 3, Article 668686, 2022.

[3] Hsiao, K., Yang, C., "Using Machine Learning for Intrusion Detection: A Review," Journal of Information Security and Applications, Vol. 70, Article 102428, 2022.

[4] "Artificial Intelligence in Cybersecurity: Challenges, Advances, and Future Perspectives" by C. Liang, R. Zhang, and Y. Liu - Computers & Security, Vol. 118, pp. 102426, 2022.

[5] "Machine Learning for Cyber Threat Intelligence: A Survey" by M. Ahmadi, M. Dehghantanha, K. Choo, and S. Singh - Journal of Network and Computer Applications, Vol. 188, pp. 103049, 2021.

[6] "Artificial Intelligence in Intrusion Detection Systems: A Survey" by S. R. Kanagavalli and V. Shanthi - Journal of Intelligent & Fuzzy Systems, Vol. 40, No. 5, pp. 8419-8432, 2021.

[7] "Deep Reinforcement Learning for Cybersecurity: A Review" by W. Meng, X. Wang, and X. Wang - IEEE Access, Vol. 8, pp. 139451-139464, 2020.

[8] "Artificial Intelligence Techniques in Cybersecurity: A Comprehensive Review" by S. Yadav, S. Kumar, and S. Chauhan - Journal of Network and Computer Applications, Vol. 150, Article 102650, 2020.

[9] Demertzis, I., Drosou, A., Varvarigou, T., "Machine Learning Techniques for Cybersecurity Applications," Journal of Information Security and Applications, Vol. 50, pp. 20-36, 2020.

[10] Karbab, A., et al., "Artificial Intelligence for Cybersecurity: A Systematic Literature Review," Future Generation Computer Systems, Vol. 107, pp. 1080-1096, 2020.

[11] "Deep Learning for Cybersecurity: A Review" by A. Sharma, M. Chen, and R. Islam - IEEE Communications Surveys & Tutorials, Vol. 22, No. 1, 2020.

[12] "A Comprehensive Survey of Machine Learning-Based Security Mechanisms" by Y. Zhang, Q. Zhu, and L. Zhang - Journal of Cyber Security Technology, Vol. 4, No. 2, pp. 109-137, 2020.

[13] "AI-Driven Cybersecurity: Challenges and Opportunities" by F. Yuan, Y. Wang, and T. Zhao - IEEE Intelligent Systems, Vol. 35, No. 6, 2020.

[14] "Artificial Intelligence and Cybersecurity: A Harm-Neutralizing and Human-Rights-Preserving Future" by N. Barabas - arXiv preprint arXiv:2004.02238, 2020.

[15] "Deep Learning for Network Intrusion Detection: A Review" by A. J. Khan and A. M. Alshomrani - Journal of Ambient Intelligence and Humanized Computing, Vol. 11, No. 2, pp. 701-714, 2020.

[16] Park, S., Lee, Y., "Deep Learning for Cybersecurity: Attack Detection and Defense," Future Internet, Vol. 11, No. 7, Article 157, 2019.

[17] "Adversarial Attacks and Defenses in Deep Learning for Cybersecurity: Challenges and Opportunities" by B. Biggio, F. Roli, and D. Ariu - IEEE Access, Vol. 7, 2019.

[18] "Artificial Intelligence in Cybersecurity: Approaches, Challenges, and Future Directions" by A. Moustafa, Y. Ghiassi, and J. H. Zhuang - IEEE Transactions on Neural Networks and Learning Systems, Vol. 30, No. 11, 2019.

[19] "Machine Learning in Cybersecurity: A Comprehensive Review" by J. Shamsi, M. Amoon, and J. Sharif - Journal of Network and Computer Applications, Vol. 143, 2019.

[20] "Machine Learning for Network Anomaly Detection: A Survey" by Y. Zhang, W. Gu, Z. Xu, and D. Zhang - IEEE Communications Surveys & Tutorials, Vol. 21, No. 4, 2019.

[21] "Machine Learning for Cybersecurity Intrusion Detection: Techniques, Applications, and Research Challenges" by A. Alazab, X. Yi, S. Slay, J. Barker, and E. M. Al-Hammadi - Information Fusion, Vol. 50, pp. 54-68, 2019.

[22] "Deep Learning for Network Security: A Survey" by S. M. Fayaz, M. R. Hashemi, and M. Dehghan - Journal of Network and Computer Applications, Vol. 135, pp. 1-18, 2019.

[23] "Artificial Intelligence in Cybersecurity: A Systematic Literature Review" by G. B. Ayshwarya, B. Gururaj, S. K. Usha, and R. R. Sudha - IOP Conference Series: Materials Science and Engineering, Vol. 666, No. 1, 2019.

[24] [24] "Machine Learning and Deep Learning for Cybersecurity Applications" by C. S. Mariñas and L. E. Ponce - Procedia Computer Science, Vol. 161, pp. 5-12, 2019.

[25] "Artificial Intelligence for Cybersecurity: A Comprehensive Survey" by T. K. Chand and M. Y. Siyal - Journal of Cybersecurity and Privacy, Vol. 2, No. 4, pp. 267-286, 2019.

[26] "Machine Learning for Insider Threat Detection: A Review" by M. Sabhnani, A. Park, and S. Upadhyaya - ACM Computing Surveys (CSUR), Vol. 52, No. 5, 2019.

[27] "A Survey on Machine Learning Techniques in Cyber Security" by A. R. Alhuthali, H. Alhazmi, M. Asim, and A. Alnafessah - Procedia Computer Science, Vol. 164, pp. 584-591, 2019.

[28] "Artificial Intelligence in Cybersecurity: A Comprehensive Survey" by C. Rajalingappaa and V. Narendiran - Journal of Network and Computer Applications, Vol. 126, pp. 46-76, 2019.

[29] "Machine Learning-Based Cybersecurity Solutions: A Survey" by M. G. Jaeger and J. Rossnagel - Journal of Cybersecurity and Privacy, Vol. 2, No. 3, pp. 193-209, 2019.

[30] "Machine Learning for Cybersecurity Threats Detection: A Survey" by M. Abolhasan, M. Elkhatib, J. Carapinha, and M. Ni - IEEE Communications Surveys & Tutorials, Vol. 21, No. 3, 2019.

[31] "Deep Learning for Malware Detection: A Survey" by H. Z. Kim, A. Salem, and X. B. Sun - arXiv preprint arXiv:1905.07106, 2019.

[32] "Machine Learning for DDoS Attack Detection: A Survey" by T. Abbas, M. Asif, I. H. Baig, S. H. Bouk, and A. H. Abdullah - IEEE Communications Surveys & Tutorials, Vol. 21, No. 4, 2019.

[33] Gupta, S., Vats, M., Khanna, A., "Artificial Intelligence in Cybersecurity: Challenges and Opportunities," International Journal of Computer Applications, Vol. 181, No. 47, pp. 37-43, 2018.

[34] Wang, Y., Chen, K., "Deep Learning for Cybersecurity: Challenges and Solutions," IEEE Security & Privacy, Vol. 16, No. 4, pp. 20-29, 2018.

[35] Bhattacharyya, S., Kalita, J., "Deep Learning in Cybersecurity: A Review," IEEE Access, Vol. 6, pp. 24411-24436, 2018.