SITJOURNAL
Security Intelligence Terrorism Journal

Security Intelligence Terrorism Journal (SITJ), Vol. 01 No. 01 (2024), pp. 1-11

Enhancing Security through Intelligent Threat Detection and Response: The Integration of Artificial Intelligence in Cyber-Physical Systems

Muhammad Nur Abdul Latif Al Waro'i[1]

[1]National Resilience Study, School of Strategic and Global Studies (SKSG) Universitas Indonesia, Jakarta, Indonesia
[1]latif.alwaroi46@gmail.com

Article Info
Received: 31-May-2024
Revised: 4-June-2024
Accepted: 11-Aug-2024

Keywords
Artificial Intelligence; Cyber-Physical Systems; Cyber Security; Deep Learning; Machine Learning

Abstract
Cyber-Physical Systems (CPS) play a crucial role in critical industries such as manufacturing, transportation, energy, and healthcare by integrating the physical and digital worlds. However, the complexity and interconnectivity of CPS with the global network increase their vulnerability to cyber-attacks. This research explores the benefits of implementing artificial intelligence (AI) in the context of cyber-physical systems (CPS) to detect and respond to security threats. This study uses machine learning and deep learning techniques to analyze sensor data and system threats. The data analysis methods encompass predictive modeling and evaluating AI algorithms' performance in detecting threats. The research data is obtained from relevant literature reviews and secondary data analysis. The research findings indicate that integrating AI in CPS can enhance the success rate of threat detection, prompt response, and accuracy in threat identification. By enabling the system to learn from previous experiences, AI can reduce the number of false positives and false negatives while providing automated real-time responses to threats without human intervention. The research concludes that AI has great potential to enhance security in CPS by providing more efficient and effective solutions to address increasingly complex cyber threats. The findings of this study are expected to provide insights and recommendations that can be applied in developing CPS security systems in the future.

1. Introduction

Cyber-Physical Systems (CPS) are systems in which software and hardware components are well-integrated and operate efficiently to accomplish their tasks. Furthermore, these tasks are performed automatically and distributed evenly (Panatagama, 2023). Industry 4.0 requires a strong understanding of Cyber-Physical Systems (CPS). The manufacturing environment that supports Industry 4.0 enabling real-time data collection, transparency, and analysis throughout the manufacturing process, is known as a cyber-physical system called cyber manufacturing. CPS has become the backbone of various critical industries such as manufacturing, transportation, energy, and healthcare, enabling the integration between the physical and digital worlds (Crackincode.com, 2024). However, the existence of CPS also brings significant security risks due to its complexity and connectivity to networks vulnerable to cyber attacks (Tyas Tunggal, 2023). In the current digital era, cybercrime not only impacts data damage and personal information but can also disrupt economic and business activities, infrastructure, and even the national security stability of a country (BPPTIK, 2023).

According to statistical data from the National Cyber and Crypto Agency (BSSN), in 2022, there were 370.02 million cyber attacks against Indonesia. This figure represents an increase of 38.72% compared to the previous year, which recorded 266.74 million cyber-attacks. The government administration sector is Indonesia's primary target of cyber attacks, reaching 284.09 million (BPPTIK,

*Corresponding author, email: latif.alwaroi46@gmail.com
doi: -

2023). Over the past ten years, the development of cyber-physical systems has undergone unexpected changes, bringing new opportunities and challenges. Threats, challenges, and critical issues have emerged, particularly in maintaining the security of CPS. The diversity of CPS essential components, whether in natural gas systems, transportation, or other automated domains, adds complexity to ensuring their security (Sectrio.com, 2024). The security threats to CPS systems are becoming increasingly complex and often difficult to detect with traditional methods, requiring innovative and effective detection and rapid response approaches. In this context, the rapid development of artificial intelligence (AI) integration in cyber-physical systems can be a promising solution to enhance the security and resilience of these systems.

The integration of AI in cyber-physical systems enables the utilization of techniques such as machine learning and deep learning to analyze data from various sensors and systems to detect undetected threat patterns by conventional methods. By harnessing AI, systems can learn from experience, enhance the success rate of threat detection, and identify threats with higher accuracy. AI integration can also improve the speed of response to threats, as systems can automatically respond to threats in real-time without human intervention. This allows for reducing response time to threats and minimizing potential negative impacts. Additionally, using AI in cyber-physical systems can aid in reducing the number of false positives (incorrect detection of threats) and false negatives (failure to detect actual threats), thereby enhancing system efficiency in managing security threats.

In this context, the research will investigate how implementing artificial intelligence (AI) in cyber-physical systems can enhance the capabilities to detect and respond to security threats. The study will analyze the success rate in threat detection, response speed, accuracy in identifying threats, and the reduction of false positives and false negatives. The research will utilize a literature review to examine how integrating artificial intelligence in cyber-physical systems can improve the capabilities to detect and respond to security threats. Data related to threat detection, response to threats, and the security performance of CPS systems will be collected and analyzed using appropriate data analysis methods. This research is expected to provide valuable insights and recommendations that can be applied in developing security systems for CPS in the future.

2. Literature Review

This research project was spurred by the realization that different people have different definitions or descriptions of digital transformation (DT), which has resulted in the creation of buzzwords and hype in academic and professional literature. On the other hand, the subject of exactly what DT is and how we should conceptualize it has received insufficient attention. In the long run, scholars and practitioners won't be able to "advance the theory and practice" of the field without the adoption of a single definition (Stock & Boyer, 2009). Although the literature now in publication indicates a new degree of interest in this field, there is evidence that this concept is not widely and thoroughly understood (Goerzig & Bauernhansl, 2018; Gray & Rumpe, 2017; Haffke et al., 2016; Morakanyane et al., 2017; Van Veldhoven & Vanthienen, 2022). Several gaps persist in the vast body of research on digital transformation. Although a large number of studies have concentrated on the technological components of digital transformation, a dearth of thorough evaluations that incorporate technology into business strategy and organizational culture exists. Second, there is a knowledge gap on the effects of DT on enterprises in emerging countries because the majority of research has been carried out in developed economies. Third, the long-term implications of digital transformation (DT) on small and medium-sized firms (SMEs) are not well studied, especially concerning maintaining innovation and competitive advantage. With an emphasis on SMEs in emerging economies, this study attempts to close these gaps by offering a comprehensive analysis of DT that takes into account its technological, strategic, and cultural aspects (Morakanyane et al., 2017; Van Veldhoven & Vanthienen, 2022).

3. Method

This study employs a qualitative methods approach to explore how Artificial Intelligence (AI) enhances the security of Cyber-Physical Systems (CPS). Data collection involves a comprehensive literature review of sources such as academic journals, conference papers, and industry reports, alongside an analysis of existing datasets on CPS security incidents and AI-based threat detection systems. For data analysis, machine learning and deep learning techniques are applied to CPS sensor data using supervised and unsupervised algorithms and neural networks. The performance of AI in threat detection is evaluated through metrics including accuracy, precision, recall, F1-score, and false positive/negative rates. Anomaly detection methods and big data analytics tools, like Apache Spark, are utilized to identify suspicious

activities within CPS. Predictive analytics are used to forecast potential security threats by analyzing historical data patterns in network traffic and system vulnerabilities. The research also investigates the integration of AI with traditional security systems, such as firewalls and IDS/IPS, to improve detection accuracy and response times. Validation and testing involve simulating CPS environments to evaluate AI-based threat detection systems under real-world cyber-attack scenarios and analyzing case studies from various industries, including manufacturing, transportation, energy, and healthcare, to assess AI's impact on security performance. Ethical considerations address data privacy, algorithmic bias, and potential AI-generated vulnerabilities, ensuring compliance with data protection regulations. The findings are presented in a structured format and disseminated through academic journals, conferences, and industry workshops, providing actionable insights for enhancing CPS security through AI integration.

4. Result and Discussion

This section will explore the capabilities of artificial intelligence (AI) that can be integrated into cyber-physical systems to enhance the ability to detect and respond to security threats. The discussion will cover the evaluation of threat detection success rates, rapid response, accuracy in threat identification, and the reduction of false positives and negatives.

4.1. Successful Threat Detection by AI

4.1.1. Machine Learning

The Machine Learning Algorithm utilizes supervised and unsupervised learning algorithms to detect suspicious patterns. In supervised learning, humans provide input, output, and feedback on prediction accuracy during training (Saravanan & Sujatha, 2018). Supervised learning can be found in various applications such as data classification, email spam detection, and face detection (Karthikeyan P. et al., 2021). This supervised learning algorithm can identify suspicious patterns by learning from labeled training data with known outcomes (Yasmin, 2024). This algorithm uses the underlying truth of samples in the training data to predict and classify new data points based on the learned patterns (Yasmin, 2024). Unsupervised learning is a method that does not require a training process and operates with unlabeled data (Saravanan & Sujatha, 2018). This method is often used in clustering, association rule mining, and anomaly detection (Karthikeyan P. et al., 2021). Unsupervised learning algorithms focus on identifying suspicious patterns in unlabeled data as references (Sen & Das, 2023). With this strategy, the algorithms search for significant patterns in the data that can provide new information and valuable knowledge (Sen & Das, 2023).

The main difference between supervised and unsupervised learning in identifying suspicious patterns lies in the availability of labeled data (Sen & Das, 2023). Supervised learning relies on labeled data to train the model, while unsupervised learning works with unlabeled data to discover relevant patterns (Sen & Das, 2023). Supervised and unsupervised learning algorithms can identify suspicious patterns by analyzing data without requiring explicit programming (Saravanan & Sujatha, 2018). Unsupervised learning methods, such as clustering, can group similar data points to detect anomalies that may contain suspicious patterns (Sen & Das, 2023).

Machine Learning (ML) algorithms, both supervised and unsupervised, are widely used to identify suspicious patterns in various domains, such as fraud detection, anomaly detection, and cybersecurity (Saravanan & Sujatha, 2018). The significance of these algorithms lies in their role in automated learning processes and performance improvement based on experience without the need for explicit programming (Saravanan & Sujatha, 2018). For instance, algorithms like Naive Bayes and Decision Trees are often employed in supervised learning, while for unsupervised learning, methods like K-means and Hierarchical Clustering are frequently chosen (Kumar et al., 2020). The appropriate selection of algorithms in the relevant context can significantly assist in effectively and efficiently identifying suspicious patterns in various machine learning (ML) applications.

4.1.2. Deep Learning

Deep Learning is a branch of machine learning that utilizes artificial neural networks to process data at various levels of abstraction (Noman, 2020). Neural networks consist of input, hidden, and output layers, resembling the human nervous system (Chaturvedi et al., 2022). Various types of neural networks, such as

CNN (Convolutional Neural Network) and RNN (Recurrent Neural Network), are commonly employed in deep learning applications (Noman, 2020).

Neural Networks have demonstrated outstanding performance in various machine learning tasks, including image and voice recognition, natural language processing, and data analysis for scientific and business purposes (Moraitis et al., 2018). Neural networks process inputs through linear and non-linear operations and use learning algorithms to optimize network parameters (Moraitis et al., 2018). With this capability, neural networks can be effectively implemented to detect complex threats and play a crucial role in cybersecurity and other applications requiring accurate and fast detection.

The application of neural networks for threat detection in cybersecurity involves the implementation of deep learning algorithms capable of extracting features from unstructured data, making them suitable for complex tasks such as threat detection (Dubey & Rasool, 2022). Neural networks can be utilized for prediction, decision-making, pattern recognition, and novelty detection with customized capabilities in threat detection through the transformation between input and output as the relationship between neurons in a sequence of layers (Kizilkan et al., 2022). Furthermore, neural networks are designed to learn from examples, making them suitable for data classification and prediction where statistical techniques and regression models have been used (ArulRaj et al., 2021). The adaptation of neural networks to evolving threats occurs by mimicking the artificial workings of the human brain, allowing for adaptation to emerging threats by learning from new examples (Kizilkan et al., 2022). The ability of neural networks to extract patterns and detect trends too complex for human observation or other computer techniques makes them highly suitable for detecting emerging threats (Macukow, 2016).

The challenges and considerations in implementing neural networks for threat detection are highly significant. The performance of neural networks in threat detection heavily relies on factors such as the number of layers, activation functions, and types of network layers (Dong et al., 2021). One of the main challenges faced in applying neural networks for threat detection is the "dimension curse," where increasing the input size and number of layers exponentially raises the complexity and learning time of the network (Zgurovsky et al., 2020). Furthermore, without expertise in a specific domain, neural networks may assist in estimating the functional structure and parameters that pose constraints in effectively implementing them for threat detection (Kizilkan et al., 2022).

4.1.3. Anomaly Analysis

Anomaly analysis is a technique used to detect unusual or suspicious activities within a system. A critical aspect of anomaly analysis is the utilization of big data. A study has concluded that high-frequency data (every five minutes) has an advantage in depicting stock market movements and can better reflect market volatility compared to low-frequency data (daily) (Aminuddin Jafry et al., 2020). In the context of the influence of analytical data on reporting quality, research has found a significant relationship between the use of analytical data and the quality of forensic audit reporting among auditors and practitioners in Malaysia (Suppiah & Arumugam, 2023). Essential methods and algorithms used in anomaly detection involve utilizing machine learning algorithms such as Support Vector Machines, Neural Networks, Random Forests, K-means, and Decision Trees (Dahiya, 2019). These algorithms have proven effective in detecting anomalies in large datasets, with a detection rate of up to 96.86% and an accuracy of 98.72% (Dahiya, 2019). Apache Spark has also emerged as a tool for analyzing big data analysis for anomaly detection, showcasing the efficiency of various machine learning algorithms in that context (Lighari & Hussain, 2017).

The utilization of analytical data has a significant contribution in detecting unusual activities. Although it does not directly discuss anomaly detection, a summary of the impact of analytical data on the quality of forensic audit reporting suggests that analytical data can assist in the collection and analysis of routine data, compiling reports, statistics, and trend analysis that are important for detecting unusual or suspicious activities (Suppiah & Arumugam, 2023). Anomaly detection in big data analytics significantly contributes to cybersecurity and fraud detection by enabling the identification of fraudulent activities and abnormal behavior within complex large datasets (Dahiya, 2019). Applying anomaly detection techniques supported by machine learning can help the detection of intrusions and unusual fraud in network security, making it an essential tool in addressing security threats (Tao & Chao, 2023). However, a significant challenge in anomaly detection in big data analytics is the "curse of big dimensionality" that affects the performance and accuracy of existing techniques due to the high volume and velocity of data generated by various sources (Thudumu et al., 2020). Traditional anomaly detection techniques may need to be more

exploratory in uncovering multiple aspects and perspectives of data analysis. They may need to fully understand the nature of anomalies, highlighting the need for specific anomaly detection techniques in particular domains (Vaishampayan et al., 2019).

Implementing anomaly detection in big data analysis has various applications in various industries, including finance, industrial systems, IoT, and critical infrastructure (Tao & Chao, 2023). Using anomaly detection techniques in these domains is highly beneficial for detecting financial fraud, identifying errors in industrial systems, and pinpointing unusual events in specific industries, such as the pulp and paper industry (Tao & Chao, 2023).

4.2. The Speed of AI in Threat Response

4.2.1. Response Automation

The utilization of artificial intelligence (AI) in threat detection and prevention has been elucidated in various studies, where AI acts as a virtual analyst alongside network intrusion detection systems to safeguard against evolving threats (Ramaiah et al., 2018). AI technology enables continuous monitoring of network activities and swift response to suspicious behaviors, thereby reducing the rate of alarm failures and false alarms (Ramaiah et al., 2018). Furthermore, AI plays a crucial role in facilitating automated response actions to address network security threats, including anomaly detection, real-time monitoring, and prompt handling of emerging threats (H. Wang et al., 2021). However, integrating AI into cybersecurity poses challenges and significant implications, including the potential emergence of new security vulnerabilities and the need for ethical considerations in developing and implementing innovative technologies (A. R. Sinha et al., 2023).

Artificial intelligence (AI) and automation are revolutionizing cybersecurity defense and incident response, overcoming traditional inefficiencies, and enabling faster incident resolution (Cid et al., 2023). A new framework advocates for synergy between AI and Open-Source Resources to enhance the effectiveness of Computer Security Incident Response Teams (CSIRT) in accelerating incident response and improving threat detection accuracy (Maezo & Echeverría Rey, 2023). AI and Digital Forensics enhance the detection and response to cloud-based security issues to improve proactive threat detection and optimized resource allocation (Mahajan et al., 2023). AI/ML technology in Security Orchestration, Automation, and Response (SOAR) solutions acts as a force multiplier, empowering SOC analysts further (Kinyua & Awuah, 2021). AI-driven planning has been applied to automate cyber incident response, demonstrating the potential for automating incident response and focusing on specific use cases such as Fake Data Injection Attacks (FDIA) (Choi et al., 2021). An integrated approach combining AI technologies such as machine learning, natural language processing, and behavioral analytics offers real-time threat identification and automated incident response in cloud security (Sinha et al., 2023). AI-assisted architecture from start to finish to detect Advanced Persistent Threats (APTs) shows a significant reduction in the amount of data that needs to be reviewed and innovative incident extraction and security-conscious assessment (Soliman et al., 2023).

4.2.2. Predictive Analytics

Artificial intelligence (AI) can predict potential attacks by analyzing various data sources, such as network traffic, user behavior, and system vulnerabilities (Affan, 2018). Machine learning algorithms detect patterns and anomalies in data that can indicate potential security threats (Affan, 2018). AI-based predictive analytics can identify potential attack vectors and vulnerabilities, enabling proactive security measures to be implemented (Affan, 2018). Using AI to prepare responses before an attack by automating the threat detection process allows for quick and targeted mitigation actions (Affan, 2018). AI-supported automated incident response systems can analyze and prioritize security alerts, enabling swift and effective responses to potential threats (Affan, 2018). AI-driven security orchestration and automation platforms facilitate the creation of pre-defined response playbooks for various types of security incidents, simplifying the response process (Affan, 2018).

Artificial intelligence (AI) techniques such as machine learning, deep learning, and natural language processing are utilized to predict attacks by analyzing historical attack data and identifying patterns that indicate potential threats in the future (Affan, 2018). Identifying anomalous behavior, predictive modeling, and threat intelligence analysis are some AI techniques used to predict potential attacks and enhance

security readiness (Affan, 2018). Using AI in security to identify attacks before they occur involves continuously monitoring and analyzing network and system behaviors against compromise indicators and potential threats (Affan, 2018). Behavior analytics supported by AI can detect deviations from standard patterns and identify early warning signs of potential attacks to enable preventive actions (Affan, 2018).

4.2.3. Integration with Existing Security Systems

The integration of AI with existing security systems, such as firewalls and IDS/IPS, can enhance the capabilities of these systems in detecting and preventing network attacks (Z. Wang & Deng, 2023). AI-driven firewalls have demonstrated higher accuracy in detecting network attacks by highlighting their effectiveness in improving network security (Z. Wang & Deng, 2023). The use of AI technology in threat detection can lead to the autonomous evolution of threat detection capabilities, thereby enhancing the overall performance of security systems (Z. Wang & Deng, 2023). When integrating AI with firewalls, IDS/IPS, and other security solutions, considerations need to be made regarding adaptability, network isolation, identification of industrial communication protocols, real-time performance, and the reliability of security systems (Z. Wang, 2021). The design and development of AI firewalls should focus on building learning models and ensuring the compatibility of AI technology with industrial communication protocols to meet the specific requirements of industrial control systems (Es-Salhi et al., 2019).

Integrating artificial intelligence technology with existing security systems presents benefits and risks. AI can play a crucial role in enhancing network protection through intelligent intrusion detection and reducing the rate of missed alarms (Huang et al., 2023). However, this integration also brings security challenges stemming from AI algorithms and computational data that have the potential to impact the security of various AI-based applications (Huang et al., 2023). Combining AI with security systems can create new vulnerabilities, including the emergence of AI-assisted network attacks and the need to address security risks associated with AI infrastructure, design, and development (Huang et al., 2023).

AI can significantly enhance the effectiveness of firewalls, IDS/IPS, and other security solutions in a network environment. AI-driven firewalls can be expanded to include mobile devices by strengthening their defense capabilities against evolving threats and offering comprehensive protection beyond conventional network boundaries (Cowan, 2019). By leveraging AI, intelligent defense systems and logical architectures can be developed for AI firewalls, integrating cloud intelligence, security centers, and layered structures to enhance threat detection and overall network security (Yuhong & Xiangdong, 2021).

4.3. Accuracy of Threat Identification by AI

4.3.1. Threat Modeling

Threat modeling is a process that involves systematically identifying and analyzing security threats within a system (Yskout et al., 2020). Threat modeling still needs a higher level of maturity and faces challenges in aligning research with industry practices (Yskout et al., 2020). One solution that can be used is artificial intelligence (AI) to automate threat detection and response and provide insights and predictions to enhance threat identification accuracy (Sinha et al., 2023). Integrating AI technologies, such as machine learning and natural language processing, can significantly improve risk assessment in various industries, including cybersecurity (Nagendiran et al., 2023). However, the development of AI technology in cybersecurity also brings negative consequences, such as the complexity of AI systems that are difficult for human analysts to understand (Shand et al., 2024). AI models to combat cybersecurity threats also face challenges due to the evolving nature of threats, vulnerabilities, and exploitations (Piplai et al., 2023). The use of AI in cybersecurity also poses technical challenges, such as data labeling and model interpretation, as well as broader social implications, including the potential creation of new security vulnerabilities (Sinha et al., 2023).

Artificial intelligence (AI) technology can analyze historical data and integrate knowledge from experts in the field, developing new paradigms and analyzing new events to enhance the accuracy of threat models (Gupta et al., 2019). The proposed AI model demonstrates excellent predictive performance under specific parameter settings, significantly increasing prediction accuracy by up to 95% (Panilov et al., 2024). Furthermore, the application of AI has evolved to forecast human behavior, such as insider threats, providing a statistical framework to effectively interpret and address such questions (Burns, 2020).

4.3.2. Dynamic Risk Assessment

The utilization of artificial intelligence (AI) for risk assessment and real-time classification faces several challenges that hinder its smooth integration into various industries (Tayal & Kulkarni, 2023). The complexity of data generated by these sectors exceeds the capabilities of traditional risk analysis methods, emphasizing the need for a more sophisticated approach to efficiently manage large amounts of data (Tayal & Kulkarni, 2023). Additionally, accurate representation of uncertainty is an important aspect of AI-based risk analysis, highlighting the importance of incorporating uncertainty characterization techniques in the risk assessment process (Stødle et al., 2024). Despite offering promising opportunities in risk analysis, concerns still arise regarding the limitations of automation and potential risks associated with granting autonomous decision-making capabilities to AI systems, which demand careful consideration and supervision in implementing AI technology in this context (Stødle et al., 2024).

The comparison between the AI-Based Dynamic Risk Measurement Method and Traditional Methods reveals the potential of AI in enhancing decision-making quality, operational efficiency, and accuracy in risk evaluation, ultimately strengthening business resilience (Ramachandran et al., 2023). Implementing AI technology, such as Natural Language Processing (NLP) and AI-based data analytics, has improved the precision and speed of risk evaluation processes, positively impacting business continuity and enabling risk prediction capabilities (Kalogiannidis et al., 2024). Furthermore, the potential application of AI-based real-time Risk Measurement and Risk Clustering highlights the strategic role of AI in enhancing business resilience by providing instant information and insights, enabling organizations to identify and address emerging risks effectively (Ramachandran et al., 2023). Moreover, the application of AI in public safety risk management can significantly contribute to anticipating and controlling exposure risks in the workplace, with positive impacts on workers' health, safety, and well-being (Pishgar et al., 2021). Meanwhile, the potential of AI in analyzing disaster risks offers a new paradigm for identifying potential hazards, constraints, and challenges in realizing the full potential of disaster risk analysis (Guikema, 2020).

4.4. Reduction of False Positives and False Negatives by AI

4.4.1. AI Model Calibration

In artificial intelligence (AI) systems, model calibration and reducing false positive and false negative errors are crucial (Chanda & Banerjee, 2022). Model calibration significantly minimizes prediction errors that can have adverse effects, such as false positive and false negative errors (Chanda & Banerjee, 2022). The primary calibration goal is to ensure that the confidence levels expressed in the model's decisions align with the actual accuracy levels, thereby enhancing the accuracy of correct assessments and reducing the likelihood of prediction errors (Chanda & Banerjee, 2022). Critical methods used to calibrate AI models to reduce false positive and false negative errors include the utilization of binning-based estimation with bins of equal mass rather than estimation with equally wide bins, as this method is known to have a lower bias (Roelofs et al., 2022). In addition, the unbiased estimator and the "ECE<inf>sweep</inf>" method have also been proven to be reliable calibration error estimation for improving the effectiveness of recalibration methods and detecting model miscalibration more effectively (Roelofs et al., 2022). Furthermore, new training strategies that consider uncertainty, such as the Confidence Weight method, have also been shown to enhance calibration performance, reduce Expected Calibration Error (ECE), and improve accuracy in clinical applications (Dawood et al., 2023).

Implementing periodic model adjustments based on feedback poses several challenges that must be addressed. One of the challenges is the difficulty in optimizing calibration parameters, such as temperature, concerning the objective function influenced by the level of calibration set difficulty (Zou et al., 2023). Another challenge is the potential degradation of calibration performance when the difficulty level of the test set differs drastically from the calibration set, resulting in suboptimal calibration parameters for Outside of Distribution (OOD) data (Zou et al., 2023). Furthermore, there needs to be more consistency in the optimal model when using different calibration methods, highlighting the need for careful consideration of performance metrics when training and selecting models for complex and high-risk applications in the healthcare field (Dawood et al., 2023).

Model adjustment driven by feedback reduces false positives and false negatives by improving the rates of true positive and true negative decisions of artificial intelligence systems (Chanda & Banerjee,

2022). This allows for refining uncertain parameters based on available information, leading to better model predictions and reducing errors in decision-making processes (Bergerson & Muehleisen, 2015).

4.4.2. Continuous Learning

The implementation of reinforcement learning to reduce false positives and false negatives in threat detection is an approach that can significantly contribute to improving the performance of threat detection systems (Jia & Wang, 2020). In reducing false positives in threat detection, reinforcement learning is used as a method that utilizes training information to evaluate the actions taken, guiding the selection of actions to maximize reward (Jia & Wang, 2020). Agents learn through experiments and constant feedback to find actions that yield maximum reward (Jia & Wang, 2020). Algorithms such as Q-Learning and dynamic programming are used in reinforcement learning (Jia & Wang, 2020). On the other hand, in reducing false negatives in threat detection, reinforcement learning involves agents interacting with the environment to achieve goals by receiving reward signals to determine successful actions (Moussaoui et al., 2023). Agents learn to select the appropriate actions in each environmental condition to optimize their objectives and maximize long-term rewards (Moussaoui et al., 2023). This approach encourages the development of more effective and reliable threat detection systems in the face of complex security challenges.

In recent years, research has focused on specific methods in continuous learning to enhance threat detection accuracy. One of the methods utilized is reinforcement learning, where agents learn actions through trial-and-error interactions in a dynamic environment (Chen & Liu, 2017). The success of reinforcement learning methods in various applications has led to increased research in this field (Chen & Liu, 2017). In the context of threat detection for continuous learning, implementing reinforcement learning methods involves using policy-based and value-based methods (Chen & Liu, 2017). The primary goal of reinforcement learning is to learn the optimal policy that can map states to actions to maximize the cumulative reward obtained (Chen & Liu, 2017).

5. Conclusion

Integrating Artificial Intelligence (AI) into Cyber-Physical Systems (CPS) offers various advantages in enhancing cybersecurity. AI has proven to be more accurate in detecting threats than conventional methods. Using machine learning and deep learning techniques enables AI to recognize patterns and anomalies in data that traditional systems may overlook. As a result, the success rate of threat detection can significantly increase, reducing the risk of undetected cyber attacks. AI-supported systems are capable of responding to threats more quickly. Real-time data analysis allows for identifying and mitigating threats quickly, reducing the potential damage that can be caused. This rapid response speed is crucial in operational environments that require quick reaction times to maintain service and operational continuity. Accuracy in identifying threats is a critical factor in maintaining CPS security. AI can minimize identification errors (false positives and false negatives) by learning from historical data and adapting to new threats. This ensures that real threats can be addressed appropriately while false threats do not disrupt system operations. One of the main challenges in cybersecurity systems is the high level of false positives and negatives that can disrupt system performance and user trust. AI can help reduce these occurrences by conducting deeper and more accurate analyses. This results in improved operational efficiency and greater confidence in the security system. In conclusion, implementing AI in CPS can provide an effective solution to enhance the detection, response, and accuracy of identifying cybersecurity threats while reducing the occurrence of false positives and false negatives. An AI system that continuously learns and adapts provides the flexibility to address evolving cybersecurity threats. Although this research has demonstrated various benefits of integrating AI in CPS, several gaps still need further exploration. Suggestions for further research include the development of more advanced and specific AI algorithms for different types of threats in CPS using hybrid machine learning and deep learning techniques. Additionally, case studies in particular industries such as energy, transportation, or healthcare can be conducted to tailor the integration of AI to the cybersecurity needs of those industries. It is also necessary to test the effectiveness of AI in diverse operational environments, including those with limited computational resources or high-latency networks. The development of AI-supported automated response systems still needs improvement to respond to threats proactively without human intervention. Furthermore, the security of AI algorithms should be considered to prevent AI from becoming a new attack vector. Evaluating the economic and operational impact of implementing AI in CPS is also essential to assess cost-benefit analysis and the overall effect on system operations.

References

Affan, M. (2018). The threat of is proxy warfare on Indonesian Millennial Muslims. Indonesian Journal of Islam and Muslim Societies, 8(2), 199 – 224. https://doi.org/10.18326/ijims.v8i2.199-223

Aminuddin Jafry, N. H., Razak, R. A., & Ismail, N. (2020). Dependence measure of five-minutes returns compared to daily returns using static and dynamic copulas | Ukuran Kebersandaran bagi Pulangan Lima-Minit Berbanding Pulangan Harian menggunakan Kopula Statik dan Dinamik. Sains Malaysiana, 49(8), 2023–2034. https://doi.org/10.17576/jsm-2020-4908-25

ArulRaj, K., Karthikeyan, M., & Narmatha, D. (2021). A View of Artificial Neural Network Models in Different Application Areas. In N. A., B. R., N. M., Y. S., T. D.C.W., A. N.A.B., & A. B.B. (Eds.), E3S Web of Conferences (Vol. 287). EDP Sciences. https://doi.org/10.1051/e3sconf/202128703001

Bergerson, J., & Muehleisen, R. (2015). Bayesian Large Model Calibration Using Simulation and Measured Data for Improved Predictions. SAE International Journal of Passenger Cars - Mechanical Systems, 8(2), 415 – 420. https://doi.org/10.4271/2015-01-0481

BPPTIK. (2023). BPPTIK Kementerian Komunikasi dan Informatika RI.

Burns, D. (2020). APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN CYBERSECURITY. In Software Engineering: Artificial Intelligence, Compliance, and Security.

Chanda, S. S., & Banerjee, D. N. (2022). Omission and commission errors underlying AI failures. AI and Society. https://doi.org/10.1007/s00146-022-01585-x

Chaturvedi, A., Apoorva, N., Awasthi, M., Jyoti, S., Akarsha, D., Brunda, S., & C S, S. (2022). Analyzing the Performance of Novel Activation Functions on Deep Learning Architectures (pp. 903–915). https://doi.org/10.1007/978-981-19-5482-5_76

Chen, Z., & Liu, B. (2017). Lifelong Reinforcement Learning. Synthesis Lectures on Artificial Intelligence and Machine Learning, 89 – 103. https://doi.org/10.1007/978-3-031-01575-5_6

Choi, T., Ko, R., Saha, T., Scarsbrook, J., Koay, A., Wang, S., Zhang, W., & Clair, C. (2021). Plan2Defend: AI Planning for Cybersecurity in Smart Grids. 1–5. https://doi.org/10.1109/ISGTAsia49270.2021.9715679

Cid, M., Pérez, M., Jorquera Valero, J. M., López Martínez, A., Maestre Vidal, J., Martinez Perez, G., García, L., Plaza, F., Nespoli, P., Pastor-Galindo, J., Ramo, P., López, F., Sánchez, P. M., & Sotelo Monge, M. (2023). European framework and proofs-of-concept for the intelliGent aUtomAtion of cybeR Defence Incident mAnagemeNt. 1–2. https://doi.org/10.23919/JNIC58574.2023.10205559

Cowan, E. J. (2019). Bringing defensive artificial intelligence capabilities to mobile devices.

Crackincode.com. (2024, February). Mendekati Era Teknologi Cyber-Physical Systems: Integrasi Dunia Digital dan Fisik - TheCrackincode.

Dahiya, P. (2019). An efficient anomaly detection technique based on optimal deep learning in big data. Journal of Advanced Research in Dynamical and Control Systems, 11(6 Special Issue), 518 – 524.

Dawood, T., Chen, C., Sidhu, B. S., Ruijsink, B., Gould, J., Porter, B., Elliott, M. K., Mehta, V., Rinaldi, C. A., Puyol-Antón, E., Razavi, R., & King, A. P. (2023). Uncertainty aware training to improve deep learning model calibration for classification of cardiac MR images. Medical Image Analysis, 88. https://doi.org/10.1016/j.media.2023.102861

Dong, S., Wang, P., & Abbas, K. (2021). A survey on deep learning and its applications. Comput. Sci. Rev., 40, 100379. https://api.semanticscholar.org/CorpusID:233547699

Dubey, A., & Rasool, A. (2022). RECENT ADVANCES AND APPLICATIONS OF DEEP LEARNING TECHNIQUE. Journal of Theoretical and Applied Information Technology, 100(13), 500 – 509.

Es-Salhi, K., Espès, D., & Cuppens-Boulahia, N. (2019). DTE Access Control Model for Integrated ICS Systems. Proceedings of the 14th International Conference on Availability, Reliability and Security. https://api.semanticscholar.org/CorpusID:199527304

Guikema, S. (2020). Artificial Intelligence for Natural Hazards Risk Analysis: Potential, Challenges, and Research Needs. Risk Analysis, 40(6), 1117–1123. https://doi.org/10.1111/risa.13476

Gupta, S., Sabitha, A. S., & Punhani, R. (2019). Cyber security threat intelligence using data mining techniques and artificial intelligence. International Journal of Recent Technology and Engineering, 8(3), 6133–6140. https://doi.org/10.35940/ijrte.C5675.098319

Huang, C., Zhang, Z., Mao, B., & Yao, X. (2023). An Overview of Artificial Intelligence Ethics. IEEE Transactions on

Artificial Intelligence, 4, 799–819. https://api.semanticscholar.org/CorpusID:251174044

Jia, J., & Wang, W. (2020). Review of reinforcement learning research. Proceedings - 2020 35th Youth Academic Annual Conference of Chinese Association of Automation, YAC 2020, 186 – 191.

Kalogiannidis, S., Kalfas, D., Papaevangelou, O., Giannarakis, G., & Chatzitheodoridis, F. (2024). The Role of Artificial Intelligence Technology in Predictive Risk Assessment for Business Continuity: A Case Study of Greece. Risks, 12(2). https://doi.org/10.3390/risks12020019

Karthikeyan P., Velswamy, K., Prithivirajkumar, P. H., Ramasamy, R., Venugopal, J., & Sarveshwaran, V. (2021). Machine Learning Techniques Application: Social Media, Agriculture, and Scheduling in Distributed Systems (pp. 1396–1417). https://doi.org/10.4018/978-1-7998-5339-8.ch068

Kinyua, J., & Awuah, L. (2021). Ai/ml in security orchestration, automation and response: Future research directions. Intelligent Automation and Soft Computing, 28(2), 527 – 545. https://doi.org/10.32604/iasc.2021.016240

Kizilkan, Z. B., Sivri, M. S., Yazici, I., & Beyca, O. F. (2022). Neural Networks and Deep Learning. Springer Series in Advanced Manufacturing, 127 – 151. https://doi.org/10.1007/978-3-030-93823-9_5

Kumar, V. U., Krishna, A., Neelakanteswara, P., & Basha, C. Z. (2020). Advanced Prediction of Performance of a Student in an University using Machine Learning Techniques. Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020, 121 – 126.

Lighari, S. N., & Hussain, D. M. A. (2017). Testing of algorithms for anomaly detection in Big data using apache spark. 2017 9th International Conference on Computational Intelligence and Communication Networks (CICN), 97–100. https://api.semanticscholar.org/CorpusID:3974844

Macukow, B. (2016). Neural networks-state of art, brief history, basic models and architecture. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9842 LNCS, 3 – 14. https://doi.org/10.1007/978-3-319-45378-1_1

Maezo, R., & Echeverría Rey, A. (2023). Boosted CSIRT with AI powered open source framework. 1–8. https://doi.org/10.23919/JNIC58574.2023.10205787

Mahajan, K., Madhavidevi, B., Supreeth, B., Lakshmi, N., Joshi, K., & Bavankumar, S. (2023). Detecting and Responding to Cloud Security Incidents based on AI and Forensic Approach. 1–6.

Moraitis, T., Sebastian, A., & Eleftheriou, E. (2018). The role of short-term plasticity in neuromorphic learning: Learning from the timing of rate-varying events with fatiguing spike-timing-dependent plasticity. IEEE Nanotechnology Magazine, 12(3), 45 – 53. https://doi.org/10.1109/MNANO.2018.2845479

Moussaoui, H., Akkad, N. E. L., & Benslimane, M. (2023). Reinforcement Learning: A review. International Journal of Computing and Digital Systems, 13(1), 1465 – 1483. https://doi.org/10.12785/ijcds/1301118

Nagendiran, S., Renugadevi, R., Kumar, R. P. A., Sasirekha, K., & Harini, R. (2023). Secure Sensitive Information on IoT using Machine Learning. International Conference on Sustainable Communication Networks and Application, ICSCNA 2023 - Proceedings, 360–366. https://doi.org/10.1109/ICSCNA58489.2023.10370157

Noman, N. (2020). A Shallow Introduction to Deep Neural Networks. Deep Neural Evolution. https://api.semanticscholar.org/CorpusID:219479212

Panatagama, A. (2023, June). Pengertian Cyber-Physical System dan Pemanfaatannya. https://terralogiq.com/cyber-physical-system/

Panilov, P., Tsibizova, T., & Voskresensky, G. (2024). Methodology of Expert-Agent Cognitive Modeling for Preventing Impact on Critical Information Infrastructure. In Communications in Computer and Information Science: Vol. 1986 CCIS. https://doi.org/10.1007/978-3-031-51057-1_21

Piplai, A., Joshi, A., & Finin, T. (2023). Offline RL+CKG: A hybrid AI model for cybersecurity tasks. CEUR Workshop Proceedings, 3433.

Pishgar, M., Issa, S. F., Sietsema, M., Pratap, P., & Darabi, H. (2021). Redeca: A novel framework to review artificial intelligence and its applications in occupational safety and health. International Journal of Environmental Research and Public Health, 18(13). https://doi.org/10.3390/ijerph18136705

Ramachandran, K. K., Karthick, K. K., Vinjamuri, L. P., Ramesh, R., Al-Taee, M., & Alazzam, M. B. (2023). Using AI for Risk Management and Improved Business Resilience. 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering, ICACITE 2023, 978–982.

Ramaiah, Ch., Charan, D. A., & Akhil, R. (2018). Secure automated threat detection and prevention (SATDP).

International Journal of Engineering and Technology, 7, 86.

Roelofs, R., Cain, N., Shlens, J., & Mozer, M. C. (2022). Mitigating Bias in Calibration Error Estimation. In C.-V. G., R. F.J.R., & V. I. (Eds.), Proceedings of Machine Learning Research (Vol. 151, pp. 4036 – 4054). ML Research Press.

Saravanan, R., & Sujatha, P. (2018). A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification. 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), 945–949.

Sectrio.com. (2024, April). Securing Cyber-Physical Systems | Challenges & Solutions in 2024. https://sectrio.com/blog/securing-cyber-physical-systems-in-2024/

Sen, R., & Das, S. (2023). Unsupervised Learning. Indian Statistical Institute Series, 305 – 318. https://doi.org/10.1007/978-981-19-2008-0_21

Shand, C., Fong, R., & Butt, U. (2024). How Explainable Artificial Intelligence (XAI) Models Can Be Used Within Intrusion Detection Systems (IDS) to Enhance an Analyst's Trust and Understanding. In Advanced Sciences and Technologies for Security Applications: Vol. Part F2308. https://doi.org/10.1007/978-3-031-47594-8_17

Sinha, A. R., Singla, K., & Victor, T. M. M. (2023). Artificial intelligence and machine learning for cybersecurity applications and challenges. In Risk Detection and Cyber Security for the Success of Contemporary Computing. https://doi.org/10.4018/978-1-6684-9317-5.ch007

Soliman, H., Sovilj, D., Salmon, G., Rao, M., & Mayya, N. (2023). RANK: AI-Assisted End-to-End Architecture for Detecting Persistent Attacks in Enterprise Networks. IEEE Transactions on Dependable and Secure Computing, PP, 1–18. https://doi.org/10.1109/TDSC.2023.3338136

Stødle, K., Flage, R., Guikema, S., & Aven, T. (2024). Artificial intelligence for risk analysis—A risk characterization perspective on advances, opportunities, and limitations. Risk Analysis. https://doi.org/10.1111/risa.14307

Suppiah, K., & Arumugam, D. (2023). Impact of data analytics on reporting quality of forensic audit: A study focus in Malaysian auditors. In P. D., N. K. B., & K. V. (Eds.), E3S Web of Conferences (Vol. 389). EDP Sciences.

Tao, Z., & Chao, J. (2023). Financial Fraud and Anomaly Detection Techniques: A Literature Review. Proceedings of the 2023 4th International Conference on Computer Science and Management Technology.

Tayal, V., & Kulkarni, P. (2023). Role of Artificial Intelligence (AI) in Risk Management. AIP Conference Proceedings, 2736(1). https://doi.org/10.1063/5.0170689

Thudumu, S., Branch, P., Jin, J., & Singh, J. (Jack). (2020). A comprehensive survey of anomaly detection techniques for high dimensional big data. Journal of Big Data, 7(1). https://doi.org/10.1186/s40537-020-00320-x

Tyas Tunggal, A. (2023, June). Apa itu Risiko Keamanan Siber? Definisi Menyeluruh | Penjaga Atas. https://www-upguard-com.translate.goog/blog/cybersecurity-risk?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=tc

Vaishampayan, S., Palshikar, G. K., Apte, M., & Attar, V. Z. (2019). Causality: An Overlooked Aspect in Anomaly Detection. IEEE Region 10 Annual International Conference, Proceedings/TENCON, 2019-October, 2413 – 2417. https://doi.org/10.1109/TENCON.2019.8929238

Wang, H., Wang, X., Wu, H., Wang, H., Zhang, L., Zhang, Q., & Wang, J. (2021). Design and research of network security threat detection and traceability system based on AI. Other Conferences.

Wang, Z. (2021). Research on Feature and Architecture Design of AI Firewall. 75–78.

Wang, Z., & Deng, Q. (2023). Research on the Application and Testing Method of AI Firewalls in Network Attack Detection. 753–757. https://doi.org/10.1109/ICCASIT58768.2023.10351578

Yasmin, J. K. (2024). Supervised Learning (pp. 69–87).

Yskout, K., Heyman, T., Van Landuyt, D., Sion, L., Wuyts, K., & Joosen, W. (2020). Threat modeling: From infancy to maturity. Proceedings - 2020 ACM/IEEE 42nd International Conference on Software Engineering: New Ideas and Emerging Results, ICSE-NIER 2020, 9–12. https://doi.org/10.1145/3377816.3381741

Yuhong, W., & Xiangdong, H. (2021). Industrial Internet Security Protection Based on an Industrial Firewall. 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), 239–247. https://api.semanticscholar.org/CorpusID:236919469

Zgurovsky, M. Z., Sineglazov, V., & Chumachenko, E. (2020). Development of Hybrid Neural Networks.

Zou, Y., Deng, W., & Zheng, L. (2023). Adaptive Calibrator Ensemble: Navigating Test Set Difficulty in Out-of-Distribution Scenarios. Proceedings of the IEEE International Conference on Computer Vision, 19276 – 19285.